




Article

K-Nearest Neighbor and Random Forest-Based Prediction of Putative Tyrosinase Inhibitory Peptides of Abalone *Haliotis diversicolor*

Sasikarn Kongsompong¹, Teerasak E-kobon^{2,3,*} and Pramote Chumnanpuen^{3,4,*}

¹ Interdisciplinary Graduate Program in Bioscience, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand; sasikarn.k@ku.th

² Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

³ Omics Center for Agriculture, Bioresources, Food and Health, Kasetsart University (OmiKU), Bangkok 10900, Thailand

⁴ Department of Zoology, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

* Correspondence: teerasak.e@ku.th (T.E.-k.); pramote.c@ku.th (P.C.)

Abstract: Skin pigment disorders are common cosmetic and medical problems. Many known compounds inhibit the key melanin-producing enzyme, tyrosinase, but their use is limited due to side effects. Natural-derived peptides also display tyrosinase inhibition. Abalone is a good source of peptides, and the abalone proteins have been used widely in pharmaceutical and cosmetic products, but not for melanin inhibition. This study aimed to predict putative tyrosinase inhibitory peptides (TIPs) from abalone, *Haliotis diversicolor*, using k-nearest neighbor (kNN) and random forest (RF) algorithms. The kNN and RF predictors were trained and tested against 133 peptides with known anti-tyrosinase properties with 97% and 99% accuracy. The kNN predictor suggested 1075 putative TIPs and six TIPs from the RF predictor. Two helical peptides were predicted by both methods and showed possible interaction with the predicted structure of mushroom tyrosinase, similar to those of the known TIPs. These two peptides had arginine and aromatic amino acids, which were common to the known TIPs, suggesting non-competitive inhibition on the tyrosinase. Therefore, the first version of the TIP predictors could suggest a reasonable number of the TIP candidates for further experiments. More experimental data will be important for improving the performance of these predictors, and they can be extended to discover more TIPs from other organisms. The confirmation of TIPs in abalone will be a new commercial opportunity for abalone farmers and industry.

Keywords: anti-tyrosinase peptides; bioinformatics; machine learning; random forest; k-nearest neighbor; abalone



Citation: Kongsompong, S.; E-kobon, T.; Chumnanpuen, P. K-nearest Neighbor and Random Forest-based Prediction of Putative Tyrosinase Inhibitory Peptides of Abalone *Haliotis diversicolor*. *Molecules* **2021**, *26*, 3671. <https://doi.org/10.3390/molecules26123671>

Academic Editors: Sabina Podlowska, Rita Guedes and Stanisław Jastrzębski

Received: 17 May 2021

Accepted: 15 June 2021

Published: 16 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Melanin is the primary determinant of skin, hair, and eye color, and helps protect against UV radiation [1]. Two types of melanin are produced in mammals; eumelanin (brownish black) and pheomelanin (reddish yellow) [2]. The production of melanin is catalyzed by a key enzyme, tyrosinase, known to be associated with pigmentation disorders [3]. Accumulation of this enzyme results in an excess amount of melanin and can cause freckles, lentigo, age spots, and skin cancer [4].

Many tyrosinase inhibitors have been used as skin whitening agents in the cosmetic industry, e.g., hydroquinone, kojic acid (KA), or arbutin. Their use remains limited due to the side effects of skin irritation, low stability in oxygen and storage, cytotoxicity, and insufficient skin penetration ability [5,6]. Investigation of new natural compounds has provided alternative opportunity for the industry. Several non-cytotoxic natural tyrosinase inhibitory peptides have been identified in the last 10 years, including proteins and peptides from milk, wheat, honey, and silk [6–12]. A diverse array of peptides have tyrosinase inhibitory functions such as cyclic peptides [13,14], *N*-acetyl-pentapeptides [15],

mimosine-tetrapeptides [16], kojic acid-peptides [5,17], and dipeptides [10]. These tyrosinase inhibitory peptides (TIPs) are hypothesized to reduce the melanogenesis process by inhibiting tyrosinase activity. Free amino acids, like cysteine, are also known to be one of the best tyrosinase inhibitors [18]. Strong tyrosinase binding peptides normally contain at least one arginine or phenylalanine with valine, alanine, and/or leucine, as well as those with hydrophobic properties [7].

Abalone is a highly valued nutritious food and luxurious cuisine [19–21] and has valuable bioactive molecules with anti-thrombotic, anticoagulant, anti-inflammatory, antioxidant, and anticancer properties [22]. However, the anti-tyrosinase property of abalone peptides has never been reported. Experimental identification of TIPs is costly and time intensive. Computational methods thus provide a preliminary solution to this hypothesis. Nithitanakool et al. [23] used molecular docking to examine the effect of compounds from Thai mango seed kernel extract on tyrosinase binding and predict anti-tyrosinase ability. Many studies have used machine learning (ML) methods to create prediction tools for various peptide properties such as MLACP for anticancer [24], AIPpred for anti-inflammatory [25], anti-biofilm [26], AmPEP for antimicrobial [27], and cell penetrating [28]. These tools were developed by extracting properties including amino acid composition (AAC), dipeptide composition (DPC), and physiochemical composition (PCP), from the amino acid sequences and used them as input features to train random forest (RF), k-nearest neighbors (KNN), and support vector machine (SVM) algorithms. Currently there are no anti-tyrosinase prediction tools. This study aimed to develop the TIP prediction tool based on the integration of k-nearest neighbor (kNN) and random forest (RF), and to predict the TIPs from abalone peptides of *Haliotis diversicolor*. Our previous study identified thousands of the abalone peptides from proteomic experiments. Discovering abalone putative anti-melanogenesis peptides and further characterizing them will be interesting for application in the pharmaceutical and nutraceutical industries.

2. Results

One-hundred and thirty-three known anti-tyrosinase peptides were obtained from literature mining of 13 published research articles. These peptides were successfully used to develop kNN and RF-based TIP predictors. Performance measurement of these two predictors on the test dataset showed high accuracy, sensitivity, specificity, precision, recall, and receiver operating characteristic curve (ROC) scores (Table 1). However, the area under the curve (AUC) scores were 0.08 for the kNN model and 0.02 for the RF model. The kNN classifiers predicted 1075 TIPs and the RF classifiers predicted six TIPs from 8330 abalone peptides (Table S1). Fifty-eight peptides (5.4%) had a kNN predictive probability score of 1.0 and 758 peptides (70.5%) had scores more than 0.9. Two of six peptides identified from the RF classifier were also predicted by the kNN classifier with probability scores of 0.77 and 0.5. The first peptide (TIP1) had nine amino acids with double serine residues, two aromatic residues (tryptophan and tyrosine), one negatively charged aspartic acid, and one positively charged arginine (TASSDAWYR). The second peptide (TIP2) was 13-amino acids long with double phenylalanine residues, one negatively charged aspartic acid, and one positively charged arginine (SAPFMPDAFFRNV). Peptide sequence alignment of all predicted TIPs showed frequent patterns of positively charged residues (arginine and lysine) similar to those of the known TIPs, which showed frequent occurrence of arginine, lysine, cysteine, and serine (Figure 1). On the other hand, the non-TIPs had a frequent pattern of glycine in addition to the arginine and lysine appearance.

Table 1. Performance measurement of kNN and RF-based TIP predictors on the test dataset evaluated by the confusionMatrix() function of the caret R package.

Machine Learning Prediction Algorithms	Performance Measurement					
	Precision	Recall	Accuracy	Sensitivity	Specificity	ROC ¹
kNN	0.89	1.00	0.97	1.00	0.96	1.00
RF	0.97	1.00	0.99	1.00	0.99	1.00

¹ Receiver operating characteristic curve.

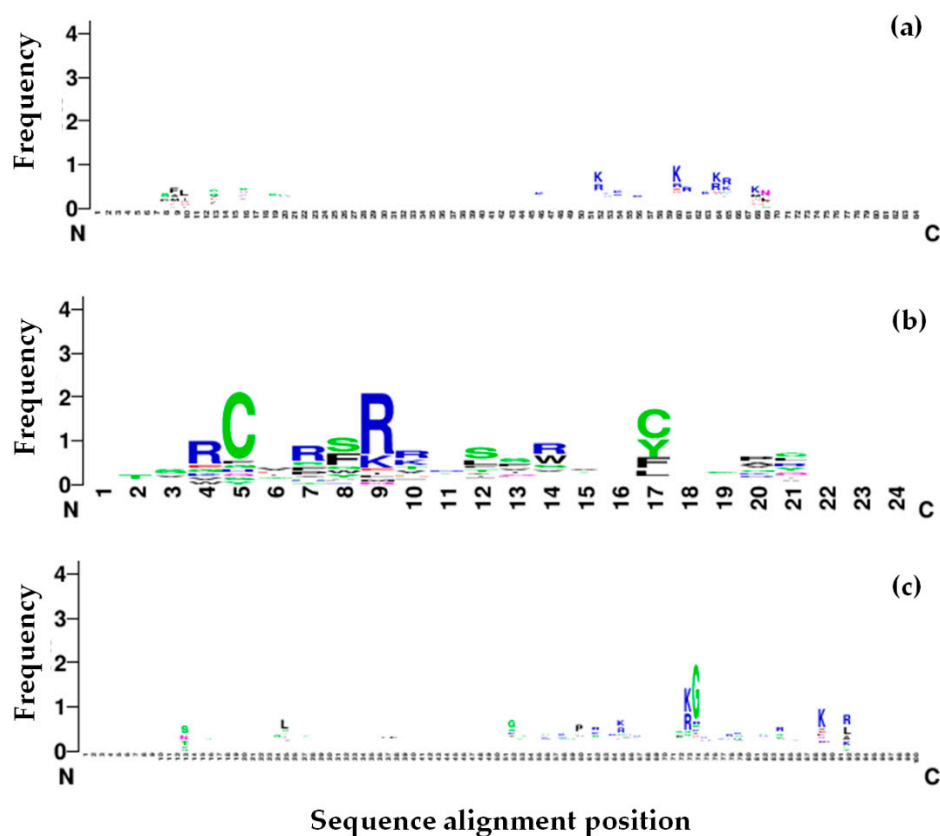


Figure 1. Multiple sequence alignment of known tyrosinase inhibitory peptides (TIPs), predicted TIPs and non-TIPs represent by the logo character plot. (a) predicted TIPs; (b) known TIPs; and (c) non-TIPs. Height of the amino acid characters showed the frequency that they appeared in the peptide sequences at a particular position.

Predicted structure of TIP1 and TIP2 showed alpha helical conformation. Molecular docking of these two TIPs to the predicted structure of mushroom tyrosinase demonstrated that TIP1 and TIP2 were similarly localized closer to the active site of the mushroom tyrosinase structure (as referenced by the position of the small ligand of the inhibitor tropolone in the structure) compared to those of the non-TIP, GKGLIAR (Figure 2). TIP1 could interact with eight residues near the active site of the mushroom tyrosinase by hydrogen bonding, while TIP2 interacted with six residues in the site (Figure 2 and Table 2). The non-TIP only interacted with six residues further out of the active site area (Figure 3). When compared with the interaction of the known TIPs (Seq_76, Seq_119, and Seq_125), the TIP1 and TIP2 showed binding positions in a similar manner to the known TIPs, and the non-TIP was clearly shown to be further away (Figure 3).

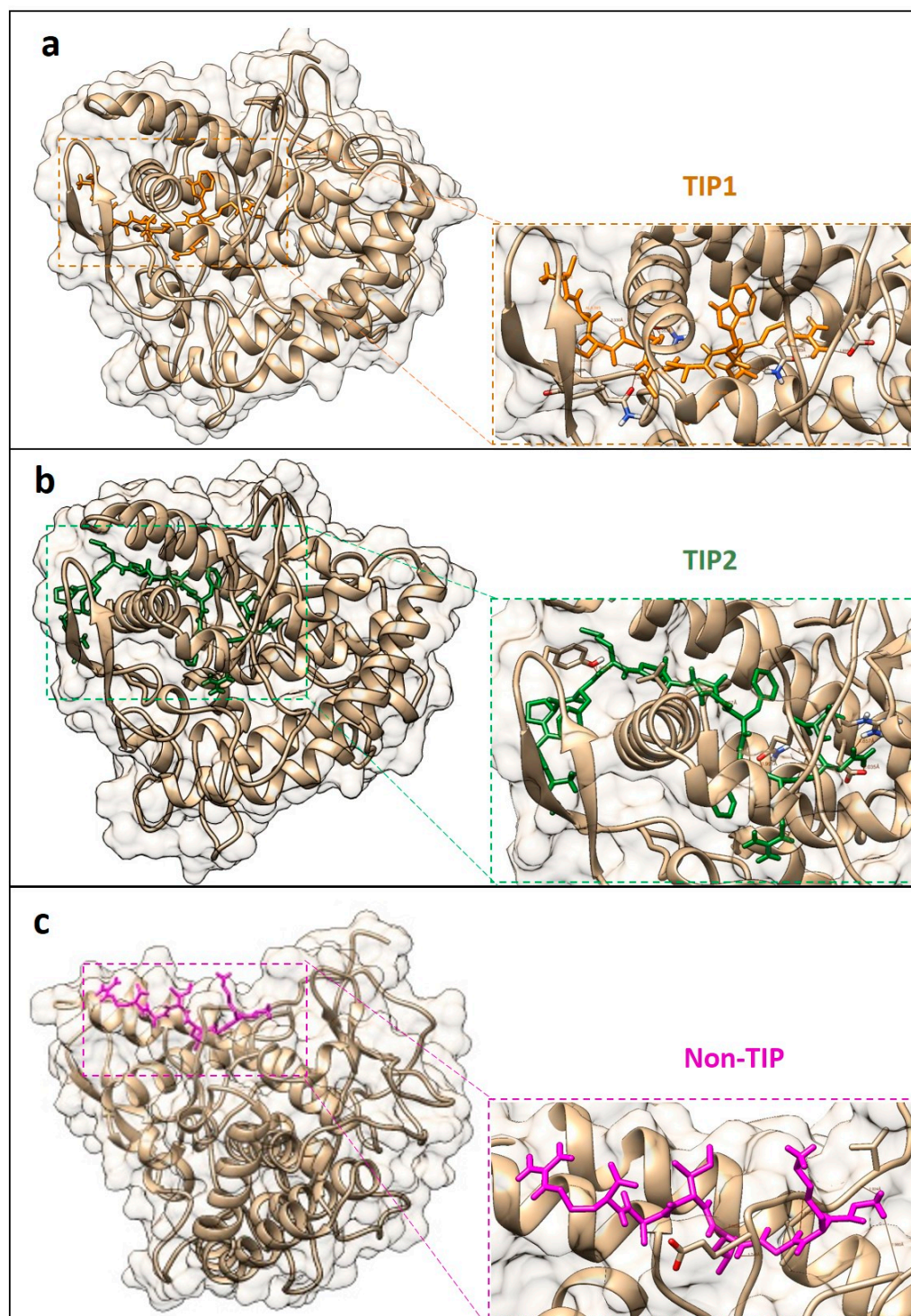
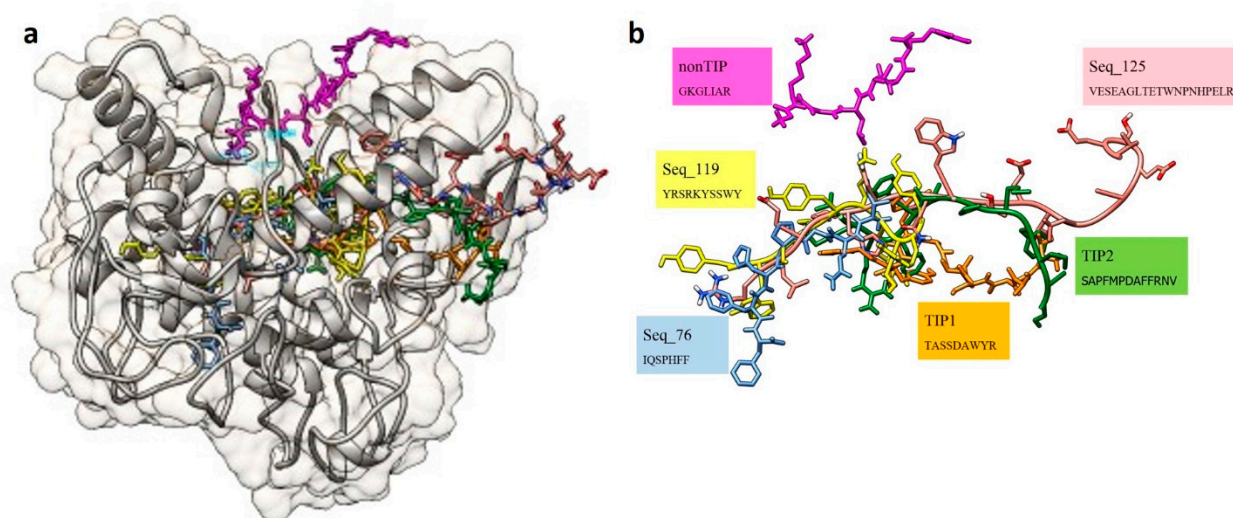


Figure 2. Molecular docking of two tyrosinase inhibitory peptides (TIP1 (a) and TIP2 (b)) and non-tyrosinase inhibitory peptide (non-TIP (c)) to the crystal structure of mushroom tyrosinase (PDB ID: 2Y9X). Structure of the mushroom tyrosinase is shaded in gray and the peptide sequences are colored as labeled above. The hydrogen bonds are shown as black lines.

Table 2. List of hydrogen bonds observed from molecular docking of three peptides (TIP1, TIP2, and non-TIP) to the crystal structure of mushroom tyrosinase (PDB ID: 2Y9X).

Peptides	Peptide Residues	Tyrosinase Residues	Distance (Å)
TIP1	SER 3	GLU 160	1.981
	SER 5	ASN 173	2.402
	SER 4	GLN 43	2.053
	TRP 7	GLN 132	1.901
	ARG 9	GLN 132	2.023
	ARG 9	GLN 132	1.939
	ARG 9	GLU 97	1.898
TIP2	ASP 7	LEU 34	1.952
	ARG 11	GLN 132	1.991
	ASN 12	GLN 132	2.096
	ASN 12	ARG 19	1.857
	ASN 12	GLU 97	2.035
Non-TIP	GLY 1	ILE 12	1.832
	GLY 1	GLY 11	1.924
	GLY 1	THR 359	1.980
	LYS 2	PRO 13	1.903
	LEU 4	ILE 16	1.746

**Figure 3.** Comparative molecular docking of three known tyrosinase inhibitory peptides (Seq_76, Seq_119, and Seq_125) with those of the non-tyrosinase inhibitory peptide (non-TIP) and the putative tyrosinase inhibitory peptides (TIP1 and TIP2) on the crystal structure of mushroom tyrosinase (PDB ID: 2Y9X) shown with (a) and without (b) the enzyme structure. Structure of the mushroom tyrosinase is shaded in gray and the peptides are labeled with different colors.

3. Discussion

Bioinformatics prediction of the tyrosinase inhibitory peptides is challenging due to lack of the TIP predictors available. This study has gathered the TIP information and found frequent amino acid patterns of these peptides consistent with those described by Schurink et al. [7]. This study observed that peptides containing at least one arginine, lysine, and phenylalanine would favor strong binding to the tyrosinase because of their charge properties, enabling the peptide–enzyme interaction (Figure 1). The finding could potentially support the predictive effort of the kNN and RF predictors, which attempted the preliminary prediction of the TIPs from limited amount of known data. The kNN algorithm is one of the simplest methods to classify the TIPs and non-TIPs by k nearest datapoints, in this case, the physicochemical properties and amino acid patterns. The peptides with similar feature patterns to those of the known TIPs would be expected

to be closer than the distinct ones. For the RF predictors, multiple decision trees were generated from our features. Some features could contribute to building a set of decision trees specific for classifying TIPs from non-TIPs. Combination of one simple and another complex machine learning algorithm allowed us to gather possible TIPs at first by the kNN classifier and then finely narrowed with the RF. A thousand of the kNN-predicted TIPs were reduced to a manageable number of six peptides by the RF classifier. The classification performance of these two algorithms were also recommended to use with various data types over Naïve Bayes algorithm by Singh et al. [29]. Despite the low AUC values of these two predictors, they might have been disregarded as giving more chance to the false positive predicted TIPs and deriving from an unbalanced dataset and effect of the oversampling method, leading to overfitting of one classifier to the data. However, the classifiers of this study had only the TIP and non-TIP group, and the performance for separating the TIPs was reasonably high based on other parameters, as shown in Table 1. Having bias towards one or another group would remain beneficial to the prediction because the authors have to interpret the predicted results by comparing the peptides with known properties from different experimental results. Fortunately, the predicted TIPs in this study had shared some sequence patterns with the known TIPs, although some might be lost as false negatives. These kNN and RF predictors proposed a few TIP candidates from nearly 8500 abalone peptides. These candidates will be easily examined by peptide synthesis and in vitro experimental treatment with tyrosinase or tyrosinase-producing cell lines. Further experimental results will assist the improvement of the TIP predictors. Successful detection of putative TIPs from the abalone peptides has raised a possible hypothesis on how the peptides performed the inhibitory function. Several organic compounds have been shown to be tyrosinase inactivators and inhibitors, scavengers of intermediate compounds, and denaturants [30]. From our molecular docking result in Figure 2, TIP1 and TIP2 were likely to be either competitive or non-competitive inhibitors of the tyrosinase, which bound to the external helices and perhaps affected conformational change of the enzyme during the reaction, resulting in a reduction in melanin production. This study also compared the predicted TIPs and non-TIP to the interaction of the known TIPs (seq_76 (IC₅₀ of 1.7 mM for monophenolase and 4.0 mM for the diphenolase activity) [12], seq_119 (IC₅₀ of 40 μM) [8], and seq_125 (IC₅₀ of 0.1 mM) [6]. Docking positions closer to the active site of the mushroom tyrosinase structure were very similar between the known TIPs and our predicted TIP1 and TIP2, compared to the non-TIP. The similar binding area of TIP1 and TIP2 to those of the three known TIPs could be further in silico evidence suggesting possible tyrosinase inhibition of our predicted peptides. Shen et al. [31] examined the inhibitory reaction of the similar peptide ECGYF (with two aromatic residues of tyrosine and phenylalanine) on tyrosinase activity, and their CD spectrometric analysis suggested that the peptide could bind to the non-active site of tyrosinase and alter the enzyme conformation, hence interfering with melanin synthesis. Therefore, the first version of the TIP predictors could suggest a reasonable number of TIP candidates for further experiments. More experimental data will be important for improving the performance of these predictors, and they can be extended to discover more TIPs from other organisms. The confirmation of TIPs in abalone will be a new commercial opportunity for abalone farmers and industry.

4. Materials and Methods

The overall workflow of this study is summarized in Figure 4. Peptides with known tyrosinase inhibitory properties were collected from previously published research [6–12,29,32] and the peptide sequences were prepared before using the predictor development. The peptide sequences were used as input for the in-house written R scripts to calculate amino acid (20 features) and di-amino acid composition (20 × 20 features), hydrophobicity, peptide length and mass, and numbers of positive charge and negative charge residues, and convert to a numeric matrix of 425 features.

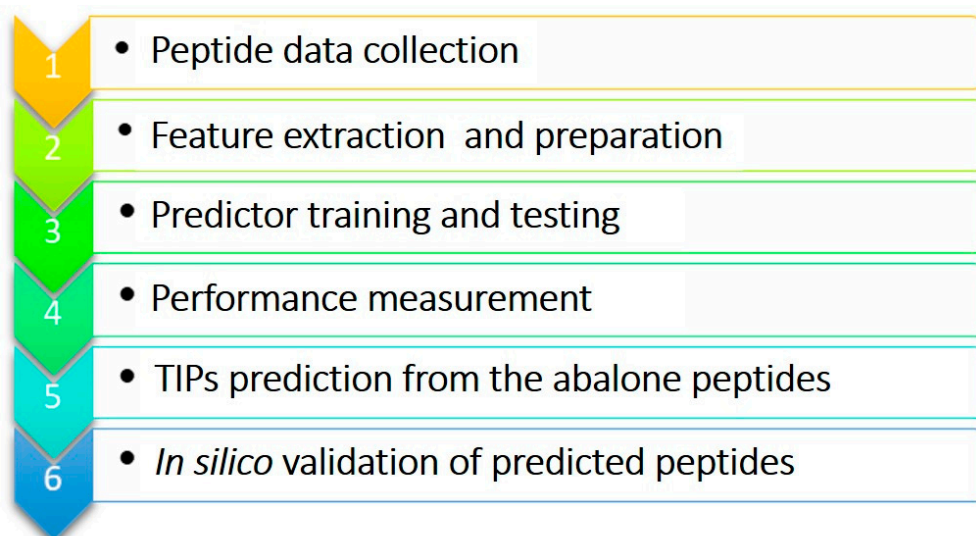


Figure 4. Workflow for bioinformatics prediction and in silico validation of the TIPs from abalone peptides.

The amino acid composition (AAC) of amino acid i was calculated in the equation below [26,33].

$$AAC(i) = \frac{\text{Amino acid frequency}(i)}{\text{Peptide length}} \quad (1)$$

Di-amino acid composition (DAA) represented the total number of dipeptides divided by 400 possible dipeptides in the given peptide sequence. The DAA of dipeptide i was calculated using the following equation [25,34,35].

$$DAA(i) = \frac{\text{Total number of dipeptides } (i)}{\text{Total number of all possible dipeptides}} \quad (2)$$

Physicochemical properties of the peptides were calculated from the percentage composition of hydrophobic (C, F, I, L, M, V, W), positively charged (K, R, H), and negatively charged (D, E) amino acid residues [24]. A new column was added to label the known TIPs as antimelanogenesis and non-TIPs as non-antimelanogenesis.

The k-nearest neighbor and random forest-based predictors were created by using the R scripts. The kNN performed a pairwise computation of a certain distance or similarity measure (k-value) for each unknown sample on every training sample [36]. This method classified the samples into groups by choosing the nearest group to the unknown samples based on the k-value [37]. The RF algorithm is suited for large datasets and has multi-model classification [34,35]. It consists of hundreds or thousands of decision trees that are called forests. Each forest randomly selects the feature at each node to determine the split and choose one or two features frequently given near the optimum results [25,38].

As the dataset was unbalanced, the oversampling method was used to balance the data with the `ovun.sample()` function of the ROSE package. The TIP/non-TIP dataset was split into 75% training and 25% test sets using the `createDataPartition()` function of the caret package. The training dataset was given to the `knn3()` function of the caret package with the optimized k value, $k = 2$ (from the optimization against the test dataset between $k = 2$ and $k = 10$), and `randomForest()` functions of the randomForest packages using `ntree = 1000`. The created predictors were tested against the test dataset using the `predict()` function and setting the type argument to "prob" for recording the predictive probability. Performance of the prediction was measured by calculating accuracy, sensitivity, specificity, precision, recall, receiver operating characteristic curve (ROC), and area under the curve (AUC) using

the confusionMatrix() function, and the twoClassSummary() and prSummary functions of the MLmetrics package. These scores were calculated by the following equations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Sensitivity/recall = \frac{TP}{TP + FN} \quad (4)$$

$$Specificity = \frac{TN}{FP + TN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

The ROC was a plot between false positive rate (FPR) as an X axis and true positive rate (TPR) as a Y axis. FPR and TPR were calculated by the following equations.

$$FPR = \frac{FP}{FP + FN} \quad (7)$$

$$TPR = \frac{TP}{TP + FN} \quad (8)$$

where *TP* = true positive, *TN* = true negative, *FP* = false positive, and *FN* = false negative. The AUC score was the measurement of the area underneath the ROC curve.

The TIP candidates were subjected to three-dimensional structure prediction using the PEP-FOLD program [39]. Peptide sequences of the known TIPs, predicted TIPs, and non-TIPs were multiply aligned using the ClustalW algorithm in the MEGA-X program version 10.2.2 [40]. The aligned sequences were visualized by plotting the logo graph using the WebLogo program version 2.8.2 [41]. Protein crystal structure of the mushroom tyrosinase (PDB ID: 2Y9X) was obtained from the PDB databank [42]. The predicted structures of known TIPs (Seq_76, Seq_119, and Seq_125), predicted TIPs and non-TIPs were docked to the mushroom tyrosinase enzyme using the GalaxyPepDock (<http://galaxy.seoklab.org/pepdock/> (accessed on 12 May 2021)) and the observed interactions were compared [43]. The docking results of the best model and hydrogen bond finding were visualized by the UCSF Chimera program to ascertain the putative predicted TIPs [44].

5. Conclusions

In conclusion, this study proposed using the first version of the kNN and RF-based TIP predictors to obtain two TIP candidates from 8330 abalone peptides for further experiments. TIP1 and TIP2 shared similar in silico binding activities to the known TIPs. The predictors can be extended to discover more TIPs from other organisms. The experimental validation of the abalone TIPs will provide novel commercial opportunity for abalone farmers and industry.

Supplementary Materials: The following are available online. Table S1: Abalone predicted anti-TIPs by kNN and RF-based predictors.

Author Contributions: Conceptualization, P.C. and T.E.-k.; methodology, P.C., T.E.-k. and S.K.; software, T.E.-k.; validation, P.C., T.E.-k. and S.K.; investigation, S.K.; resources, T.E.-k.; writing—original draft preparation, S.K.; writing—review and editing, P.C. and T.E.-k.; visualization, S.K. and T.E.-k.; supervision, P.C. and T.E.-k.; project administration, P.C.; funding acquisition, P.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science Achievement Scholarship of Thailand (SAST), the Interdisciplinary Graduate Program in Bioscience, Department of Genetics and Department of Zoology, Faculty of Science, Kasetsart University, Thailand.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank Department of Genetics and Department of Zoology, Faculty of Science, Kasetsart University, for support all research facilities and equipment.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Sample Availability: Not applicable.

References

1. Sun, C.L.; Chen, L.; Xu, J.; Qu, W.; Guan, L.; Liu, W.Y.; Akihisa, T.; Feng, F.; Zhang, J. Melanogenesis-inhibitory and antioxidant activities of Phenolics from *Periploca forrestii*. *Chem. Biodivers.* **2017**, *14*. [[CrossRef](#)]
2. Simon, J.D.; Peles, D.; Wakamatsu, K.; Ito, S. Current challenges in understanding melanogenesis: Bridging chemistry, biological control, morphology, and function. *Pigment Cell Melanoma Res.* **2009**, *22*, 563–579. [[CrossRef](#)]
3. Halaban, R.; Patton, R.S.; Cheng, E.; Svedine, S.; Trombetta, E.S.; Wahl, M.L.; Ariyan, S.; Hebert, D.N. Abnormal acidification of melanoma cells induces tyrosinase retention in the early secretory pathway. *J. Biol. Chem.* **2002**, *277*, 14821–14828. [[CrossRef](#)] [[PubMed](#)]
4. Oh, C.T.; Kwon, T.R.; Jang, Y.J.; Yoo, K.H.; Kim, B.J.; Kim, H. Inhibitory effects of *Stichopus japonicus* extract on melanogenesis of mouse cells via ERK phosphorylation. *Mol. Med. Rep.* **2017**, *16*, 1079–1086. [[CrossRef](#)] [[PubMed](#)]
5. Singh, B.K.; Park, S.H.; Lee, H.B.; Goo, Y.A.; Kim, H.S.; Cho, S.H.; Lee, J.H.; Ahn, G.W.; Kim, J.P.; Kang, S.M.; et al. Kojic acid peptide: A new compound with anti-tyrosinase potential. *Ann. Dermatol.* **2016**, *28*, 555–561. [[CrossRef](#)]
6. Karkouch, I.; Tabbene, O.; Gharbi, D.; Mlouka, M.A.B.; Elkahoui, S.; Rihouey, C.; Coquet, L.; Cosette, P.; Jouenne, T.; Limam, F. Antioxidant, antityrosinase and antibiofilm activities of synthesized peptides derived from *Vicia faba* protein hydrolysate: A powerful agents in cosmetic application. *Ind. Crops Prod.* **2017**, *109*, 310–319. [[CrossRef](#)]
7. Schurink, M.; van Berkel, W.J.; Wichers, H.J.; Boeriu, C.G. Novel peptides with tyrosinase inhibitory activity. *Peptides* **2007**, *28*, 485–495. [[CrossRef](#)]
8. Ubeid, A.A.; Zhao, L.; Wang, Y.; Hantash, B.M. Short-sequence oligopeptides with inhibitory activity against mushroom and human tyrosinase. *J. Investig. Dermatol.* **2009**, *129*, 2242–2249. [[CrossRef](#)]
9. Hsiao, N.W.; Tseng, T.S.; Lee, Y.C.; Chen, W.C.; Lin, H.H.; Chen, Y.R.; Wang, Y.T.; Hsu, H.J.; Tsai, K.C. Serendipitous discovery of short peptides from natural products as tyrosinase inhibitors. *J. Chem. Inf. Model.* **2014**, *54*, 3099–3111. [[CrossRef](#)]
10. Tseng, T.S.; Tsai, K.C.; Chen, W.C.; Wang, Y.T.; Lee, Y.C.; Lu, C.K.; Don, M.J.; Chang, C.Y.; Lee, C.H.; Lin, H.H.; et al. Discovery of potent cysteine-containing dipeptide inhibitors against tyrosinase: A comprehensive investigation of 20 × 20 dipeptides in inhibiting dopachrome formation. *J. Agric. Food Chem.* **2015**, *63*, 6181–6188. [[CrossRef](#)]
11. Ochiai, A.; Tanaka, S.; Tanaka, T.; Taniguchi, M. Rice bran protein as a potent source of antimelanogenic peptides with tyrosinase inhibitory activity. *J. Nat. Prod.* **2016**, *79*, 2545–2551. [[CrossRef](#)] [[PubMed](#)]
12. Nie, H.; Liu, L.; Yang, H.; Guo, H.; Liu, X.; Tan, Y.; Wang, W.; Quan, J.; Zhu, L. A novel heptapeptide with tyrosinase inhibitory activity identified from a phage display library. *Appl. Biochem. Biotechnol.* **2017**, *181*, 219–232. [[CrossRef](#)]
13. Morita, H.; Kayashita, T.; Kobata, H.; Gonda, A.; Takeya, K.; Itokawa, H.; Itokawa, H. Pseudostellarins A-C, new tyrosinase inhibitory cyclic peptides from *Pseudostellaria heterophylla*. *Tetrahedron* **1994**, *50*, 6797–6804. [[CrossRef](#)]
14. Morita, H.; Kayashita, T.; Kobata, H.; Gonda, A.; Takeya, K.; Itokawa, H. Pseudostellarins D-F, new tyrosinase inhibitory cyclic peptides from *Pseudostellaria heterophylla*. *Tetrahedron* **1994**, *50*, 9975–9982. [[CrossRef](#)]
15. Lien, C.Y.; Chen, C.Y.; Lai, S.T.; Chan, C.F. Kinetics of mushroom tyrosinase and melanogenesis inhibition by N-acetyl-pentapeptides. *Sci. World J.* **2014**, *2014*, 409783. [[CrossRef](#)] [[PubMed](#)]
16. Upadhyay, A.; Chompoo, J.; Taira, N.; Fukuta, M.; Gima, S.; Tawata, S. Solid-phase synthesis of mimosine tetrapeptides and their inhibitory activities on neuraminidase and tyrosinase. *J. Agric. Food Chem.* **2011**, *59*, 12858–12863. [[CrossRef](#)]
17. Kim, H.; Choi, J.; Cho, J.K.; Kim, S.Y.; Lee, Y.S. Solid-phase synthesis of kojic acid-tripeptides and their tyrosinase inhibitory activity, storage stability, and toxicity. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2843–2846. [[CrossRef](#)]
18. Kahn, V. Effect of Proteins, Protein Hydrolyzates and Amino Acids on o-Dihydroxyphenolase Activity of Polyphenol Oxidase of Mushroom, Avocado, and Banana. *J. Food Sci.* **1985**, *50*, 111–115. [[CrossRef](#)]
19. Iba, W. Nutrition Requirement of Cultured Abalone Post Larvae and Juveniles: A Review. *Indones. Aquac. J.* **2008**, *3*, 45–57. [[CrossRef](#)]
20. Lou, Q.M.; Wang, Y.M.; Xue, C.H. Lipid and fatty acid composition of two species of abalone, *Haliotis discus hannai* Ino and *Haliotis diversicolor* Reeve. *J. Food Biochem.* **2013**, *37*, 296–301. [[CrossRef](#)]
21. Latuihamallo, M.; Iriana, D.A.; Apituley, D. Amino acid and fatty acid of abalone *Haliotis squamata* cultured in different aquaculture systems. *Procedia Food Sci.* **2015**, *3*, 174–181. [[CrossRef](#)]
22. Venugopal, V.; Gopakumar, K. Shellfish: Nutritive value, health benefits, and consumer safety. *Compr. Rev. Food Sci. Food Saf.* **2017**, *16*, 1219–1242. [[CrossRef](#)]

23. Nithitanakool, S.; Pithayanukul, P.; Bavovada, R.; Saparpakorn, P. Molecular docking studies and anti-tyrosinase activity of Thai mango seed kernel extract. *Molecules* **2009**, *14*, 257–265. [[CrossRef](#)]
24. Manavalan, B.; Basith, S.; Shin, T.H.; Choi, S.; Kim, M.O.; Lee, G. MLACP: Machine-learning-based prediction of anticancer peptides. *Oncotarget* **2017**, *8*, 77121–77136. [[CrossRef](#)] [[PubMed](#)]
25. Manavalan, B.; Shin, T.H.; Kim, M.O.; Lee, G. AIPpred: Sequence-based prediction of anti-inflammatory peptides using random forest. *Front. Pharmacol.* **2018**, *9*, 276. [[CrossRef](#)] [[PubMed](#)]
26. Gupta, S.; Ansari, H.R.; Gautam, A.; Raghava, G.P. Identification of B-cell epitopes in an antigen for inducing specific class of antibodies. *Biol. Direct.* **2013**, *8*, 27. [[CrossRef](#)] [[PubMed](#)]
27. Bhadra, P.; Yan, J.; Li, J.; Fong, S.; Siu, S.W. AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest. *Sci. Rep.* **2018**, *8*, 1697. [[CrossRef](#)] [[PubMed](#)]
28. Sanders, W.S.; Johnston, C.I.; Bridges, S.M.; Burgess, S.C.; Willeford, K.O. Prediction of cell penetrating peptides by support vector machines. *PLoS Comp. Biol.* **2011**, *7*, e1002101. [[CrossRef](#)]
29. Singh, A.; Halgamuge, M.N.; Lakshminathan, R. Impact of Different Data Types on Classifier Performance of Random Forest, Naïve Bayes, and K-Nearest Neighbors Algorithms. *Int. J. Adv. Comp. Sci. Appl.* **2017**, *8*, 1–10. [[CrossRef](#)]
30. Zolghadri, S.; Bahrami, A.; Hassan Khan, M.T.; Munoz-Munoz, J.; Garcia-Molina, F.; Garcia-Canovas, F.; Saboury, A.A. A comprehensive review on tyrosinase inhibitors. *J. Enzym. Inhib. Med. Chem.* **2019**, *34*, 279–309. [[CrossRef](#)] [[PubMed](#)]
31. Shen, Z.; Wang, Y.; Guo, Z.; Tan, T.; Zhang, Y. Novel tyrosinase inhibitory peptide with free radical scavenging ability. *J. Enzym. Inhib. Med. Chem.* **2019**, *34*, 1633–1640. [[CrossRef](#)]
32. Pillaiyar, T.; Manickam, M.; Namasivayam, V. Skin whitening agents: Medicinal chemistry perspective of tyrosinase inhibitors. *J. Enzym. Inhib. Med. Chem.* **2017**, *32*, 403–425. [[CrossRef](#)]
33. Lata, S.; Sharma, B.K.; Raghava, G.P.S. Analysis and prediction of antibacterial peptides. *BMC Bioinform.* **2007**, *8*, 263. [[CrossRef](#)]
34. Sharma, A.K.; Gupta, A.; Kumar, S.; Dhakan, D.B.; Sharma, V.K. Woods: A fast and accurate functional annotator and classifier of genomic and metagenomic sequences. *Genomics* **2015**, *106*, 1–6. [[CrossRef](#)]
35. Gupta, S.; Sharma, A.K.; Jaiswal, S.K.; Sharma, V.K. Prediction of biofilm inhibiting peptides: An in silico approach. *Front. Microbiol.* **2016**, *7*, 949. [[CrossRef](#)]
36. Maillo, J.; Ramírez, S.; Triguero, I.; Herrera, F. kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data. *Knowl.-Based Syst.* **2017**, *117*, 3–15. [[CrossRef](#)]
37. Thanh Noi, P.; Kappas, M. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors* **2018**, *18*, 18. [[CrossRef](#)]
38. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
39. Beaufays, J.; Lins, L.; Thomas, A.; Basseur, R. In silico predictions of 3D structures of linear and cyclic peptides with natural and non-proteinogenic residues. *J. Pept. Sci.* **2011**, *18*, 17–24. [[CrossRef](#)]
40. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)] [[PubMed](#)]
41. Crooks, G.E.; Hon, G.; Chandonia, J.M.; Brenner, S.E. WebLogo: A sequence logo generator. *Genome Res.* **2004**, *4*, 1188–1190. [[CrossRef](#)]
42. Ismaya, T.; Rozeboom, J.; Weijn, A.; Mes, J.; Fusetti, F.; Wichers, J.; Dijkstra, W. Crystal Structure of Agaricus Bisporus Mushroom Tyrosinase: Identity of the Tetramer Subunits and Interaction with Tropolone. *Biochemistry* **2011**, *50*, 5477–5486. [[CrossRef](#)] [[PubMed](#)]
43. Lee, H.; Heo, L.; Lee, M.S.; Seok, C. GalaxyPepDock: A protein-peptide docking tool based on interaction similarity and energy optimization. *Nucleic Acids Res.* **2015**, *43*, W431–W435. [[CrossRef](#)] [[PubMed](#)]
44. Pettersen, F.; Goddard, D.; Huang, C.; Couch, S.; Greenblatt, M.; Meng, C.; Ferrin, E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)] [[PubMed](#)]