

Navneeta Bharadvaja

Arpit thesis 2026

Document Details

Submission ID

trn:oid:::27535:139715622

Submission Date

May 20, 2026, 9:37 AM GMT+0

Download Date

May 20, 2026, 9:41 AM GMT+0

File Name

Arpit thesis 2026.docx

File Size

5.0 MB

40 Pages

9,398 Words

58,165 Characters

8% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Cited Text

Match Groups

-  **39 Not Cited or Quoted 8%**
Matches with neither in-text citation nor quotation marks
-  **0 Missing Quotations 0%**
Matches that are still very similar to source material
-  **3 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 7%  Internet sources
- 4%  Publications
- 6%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 39 Not Cited or Quoted 8%**
Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%**
Matches that are still very similar to source material
- 3 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 7% Internet sources
- 4% Publications
- 6% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Student papers	Krea University on 2026-05-01	1%
2	Internet	dspace.dtu.ac.in:8080	1%
3	Student papers	Delhi Technological University on 2026-04-29	<1%
4	Student papers	Delhi Technological University on 2026-04-30	<1%
5	Internet	discovery.researcher.life	<1%
6	Internet	www.mdpi.com	<1%
7	Internet	webthesis.biblio.polito.it	<1%
8	Student papers	University of Petroleum and Energy Studies on 2015-11-23	<1%
9	Internet	sear.unisq.edu.au	<1%
10	Internet	www.esp.org	<1%

11	Internet	peerj.com	<1%
12	Internet	www.nepjol.info	<1%
13	Publication	"The Brassica juncea Genome", Springer Science and Business Media LLC, 2022	<1%
14	Publication	Antimicrobial Peptides and Innate Immunity, 2013.	<1%
15	Student papers	Far Eastern University on 2026-05-19	<1%
16	Internet	ece.anits.edu.in	<1%
17	Internet	www.researchsquare.com	<1%
18	Internet	www.trjfas.org	<1%
19	Student papers	Chiang Mai University on 2023-10-04	<1%
20	Internet	link.springer.com	<1%
21	Internet	mdpi-res.com	<1%
22	Internet	tdx.cat	<1%
23	Publication	Kang Zhang, Yinqing Yang, Xin Zhang, Lingkui Zhang et al. "The genome of Orych...	<1%
24	Student papers	University of Westminster on 2026-04-23	<1%

25	Internet	assets-eu.researchsquare.com	<1%
26	Internet	mafiadoc.com	<1%
27	Publication	Yueyao Wang, Yongdong Xu, Zhidan Liu. "A review of plant antipathogenic consti...	<1%
28	Publication	Gulsah Karabulut, Busra Gultekin Subasi, Petya Ivanova, Gulden Goksen, Vesela C...	<1%
29	Publication	Khames, Salha Khames Mohamed Ali. "Isolation and Characterization of Small An...	<1%
30	Student papers	Queen's University of Belfast on 2019-01-11	<1%
31	Publication	R. K. Naresh, R. K. Gupta, Rajbir Garg, R. K. Sohane, J. P. Singh, O. P. Gupta, Arun K...	<1%
32	Internet	pmc.ncbi.nlm.nih.gov	<1%

**Genome-Wide Identification and *In Silico*
Characterization of Putative Antimicrobial
Peptides from Seed Storage and Defense
Proteins of *Brassica juncea***

**Thesis Submitted
in Partial Fulfillment of the Requirements for the
Degree of**

**MASTERS OF TECHNOLOGY
in
BIOINFORMATICS**

By

**ARPIT RATURI
(Roll No. 24/BIO/13)**

**Under the Supervision of
Dr. NAVNEETA BHARADVAJA**



**Department of Biotechnology
DELHI TECHNOLOGICAL UNIVERSITY**

**(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42**



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I Arpit Raturi hereby certify that the work which is being presented in the thesis entitled **“Genome-Wide Identification and *In Silico* Characterization of Putative Antimicrobial Peptides from Seed Storage and Defense Proteins of *Brassica juncea*”** in partial fulfillment of the requirements for the award of the Degree of Master of Technology, submitted in the Department of Biotechnology, Delhi Technological University is an authentic record of my own work carried out during the period from January 2026 to May 2026 under the supervision of Dr. Navneeta Bharadvaja.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiner in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daultpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR

Certified that **Arpit Raturi (24/BIO/13)** has carried out their search work presented in this thesis entitled “**Genome-Wide Identification and *In Silico* Characterization of Putative Antimicrobial Peptides from Seed Storage and Defense Proteins of *Brassica juncea***” for the award of **Master of Technology** from Department of Biotechnology, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student himself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Dr. Navneeta Bharadvaja
Supervisor
Department of Biotechnology
Delhi Technological University

Prof. Yasha Hasija
Head of Department
Department of Biotechnology
Delhi Technological University

Date:

ACKNOWLEDGEMENT

I would like to sincerely thank Dr. Navneeta Bharadvaja, my research supervisor, for her unwavering support, expert guidance, and encouragement throughout the duration of this work. Her deep knowledge, valuable suggestions, and timely feedback were instrumental in shaping the direction and quality of this thesis. I am truly grateful for the opportunity to work under her mentorship.

I also extend my gratitude to the Department of Biotechnology at Delhi Technological University for providing all the essential facilities and resources. Their support was crucial for the experimental work and overall progress of this study.

A special thanks goes to the PhD scholars Ms. Harshita, Ms. Anuradha and Ms. Pragati who supported throughout this journey and guided with everyday work. Their expertise and willingness to help were truly invaluable, and I am deeply grateful for their mentorship. I would also like to thank to Mr. Jitendra Singh, Mr. CB Singh, Mr. Lalit, Mr. Jaspreet, and Mr. Rajesh for their technical assistance and guidance through the project.

Lastly, I would like to thank my family and friends for their unwavering support and for making this journey more enjoyable. Their encouragement and camaraderie were essential in helping us support each other through this process.

ARPIT RATURI

24/BIO/13

ABSTRACT

Seed storage proteins not only provide nutrients for the germinating embryo but can also serve as precursors of antimicrobial peptides (AMPs), key components of plant innate immunity. In this study, we assessed the predicted antimicrobial potential of major seed storage proteins of *Brassica juncea*, cruciferin, napin, and oleosin, as well as defense proteins, defensin and BjCHI1 (chitinase). Reference protein sequences from related species were used to find homologous genes in the *Brassica juncea* genome using BLAST. A total of twenty-nine homologous genes were identified and further screened for AMP-like characteristics using the CAMP database. To better understand their molecular diversity and evolutionary relationships, we performed genome-wide analyses, including conserved domain identification, chromosomal localization, phylogenetic analysis, gene structure examination, and physicochemical characterization. In silico enzymatic hydrolysis was performed to simulate natural proteolytic cleavage, generating potential AMP fragments ranging from 10 to 60 amino acids in length. Using various bioinformatics tools, structural and therapeutic safety assessments including toxicity and allergenicity prediction, were conducted. This study provides a systematic genome-wide identification and multi-level characterization of seed storage and defense protein homologs in *Brassica juncea*, highlighting their potential as a conserved genomic reservoir of antimicrobial peptide precursors.

Keywords- Antimicrobial peptides, *Brassica juncea*, Seed storage proteins, Machine learning, Genome wide identification, in silico analysis

CONTENTS

Title	Page no.
Candidate Declaration.....	(ii)
Certificate.....	(iii)
Acknowledgement.....	(iv)
Abstract.....	(v)
Contents.....	(vi)
Figures	(ix)
Tables.....	(x)
Abbreviations	(xi)

CHAPTER. 1

INTRODUCTION	1-3
1.1. The background and global context of Anti-microbial- resistant	1
1.2. Plant-Derived AMPs: The Case for Brassica juncea.....	1
1.2.1 Seed Storage Proteins as Precursors.....	2
1.2.2 Defense Proteins: Chitinases and Defensins.....	2
1.3 Objective.....	3

CHAPTER. 2

Literature review	4-7
2.1. Importance of Antimicrobial peptides.....	4
2.2. Seed Storage and Defense Proteins in Brassica juncea.....	4
2.3. Structural Basis of Antimicrobial Activity.....	4-5
2.4. Mechanisms of Action.....	5
2.5. Evolutionary Conservation and Diversity	5
2.6. Genome-wide Discovery Approaches in Antimicrobial Peptides	5
2.6.1 Bioinformatics Pipelines and Tools.....	5-6
2.6.2 Experimental Validation of Predicted AMPs.....	6

2.7 In Silico Characterization Methods for AMPs.....	6
2.7.1 Computational Tools and Databases.....	6
2.7.2 Machine Learning and Predictive Modeling.....	6
2.7.3 Molecular Docking and Dynamics Simulations.....	7
2.8. Real-world Implications.....	7

CHAPTER – 3

METHODOLOGY 8-14

3.1. Retrieval of Cruciferin, Napin, Oleosin, Defensin, and Chitinase Family Homologs in the Brassica juncea Genome.....	9
3.2. Phylo analysis.....	9
3.3. Prediction of Structure, Motifs and Conserved Domain.....	10
3.4. Promoter region analysis.....	10
3.5. Collinearity Synteny analysis.....	11
3.6. Localization of genes on Chromosome.....	11
3.7. Gene Ontology and subcellular localization	12
3.8. Ka/ks and duplication events.....	12
3.9. In silico enzymatic hydrolysis.....	12
3.10. Antimicrobial capability prediction of released peptides.....	13
3.11. Predicting AMPs from their physicochemical characteristics.....	14
3.12. Structure analysis of the AMP sequences.....	14
3.13. Computational Analysis of AMP Allergenicity and Toxicity.....	14

CHAPTER - 4

RESULT AND DISCUSSION 15-21

4.1. Retrieval of Cruciferin, Napin, Oleosin, Defensin, and Chitinase Homologs	15
4.2. Phylogenetic Analysis of the Gene Families	15
4.3. Identification of Conserved Motifs, Domains, and Gene Structures.....	15-16
4.4. Cis-analysis of the genes.....	16
4.5. Physical mapping of Chromosomes location and Ka/ks ratio analysis.....	16-17

4.6.	Synten analysis	17-18
4.7.	Gene Ontology and subcellular localization.....	18
4.8.	Computational enzymatic hydrolysis of the identified genes.....	18
4.9.	Prediction of Antimicrobial potential of released peptides.....	18-19
4.10.	Comparative physicochemical profiling of predicted AMP candidates	19-20
4.11.	Secondary Structure Prediction of AMP's.....	21
4.12.	Computational Analysis of AMP Allergenicity and Toxicity.....	21

CHAPTER - 5

CONCLUSION AND FUTURE SCOPE.....	22
----------------------------------	----

REFERENCES.....	23-28
-----------------	-------

PLAGARISM REPORT.....	29
-----------------------	----

FIGURES

FIG NO.'s	TITLE	PAGE
Fig.1	Triangle of U	2
Fig.2	Workflow of the study	8
Fig.3	GUI of MEGA software	9
Fig.4	GUI of NCBI CD Database	10
Fig.5	GUI of GSDS 2.0 server	10
Fig.6	GUI of Plant Care server	11
Fig.7	GUI of TBtools Dual Synteny Plot	11
Fig.8	GUI of BIOPEP-UWM database	13
Fig.9	GUI of CAMP R4 Database	13
Fig.10	GUI of CSM-Toxin Server	14

TABLES

TABLE's	TITLE	PAGE
Table. 1	Summary of the ka/ks ratio analysis for the candidate gene pair present in <i>Brassica juncea</i>	17
Table. 2	Physicochemical profiling of antimicrobial peptides using in silico tools	20

LIST OF ABBREVIATIONS

AA:	Amino Acid
ABA:	Abscisic Acid
ABRE:	Abscisic Acid-Responsive Element
AMP's:	Antimicrobial Peptides
AMR:	Anti-microbial Resistance
ANN:	Artificial Neural Network
APD:	Antimicrobial Peptide Database
BjuCHI:	<i>Brassica juncea</i> Chitinase
BjuCr:	<i>Brassica juncea</i> Cruciferin
BjuDe:	<i>Brassica juncea</i> Defensin
BjuNa:	<i>Brassica juncea</i> Napin
BjuOl:	<i>Brassica juncea</i> Oleosin
CAMP:	Collection of Antimicrobial Peptides (Database)
CDS:	Coding Sequence
DRE:	Dehydration-Responsive Element
ERE:	Ethylene-Responsive Element
EST:	Expressed Sequence Tag
GO:	Gene Ontology
GRAVY:	Grand Average of Hydropathicity
GSDS:	Gene Structure Display Server
iTOL:	Interactive Tree of Life
JA:	Jasmonic Acid
Ka:	Non-synonymous substitution rate
Ks:	Synonymous substitution rate
MBS:	MYB Binding Site
MDR:	Multi-drug resistant
MEME:	Multiple Em for Motif Elicitation
ML:	Machine Learning/Maximum Likelihood
NCBI:	National Center for Biotechnology Information
pI:	Isoelectric Point
QTL:	Quantitative Trait Loci
RF:	Random Forest
SD:	Segmental Duplication
SNP:	Single Nucleotide Polymorphism
SSPs:	Seed Storage Proteins
SVM:	Support Vector Machine
TD:	Tandem Duplication
GUI	Graphical User Interface
WGT:	Whole-Genome Triplication

CHAPTER. 1

INTRODUCTION

1.1 The background and global context of Anti-microbial- resistant (AMR)

Multi-drug resistant (M.D.R) pathogens are one of the greatest threats to public health in the 21st century. Infections by bacteria and fungi were once easy to cure with traditional antibiotics, but microbes are developing increasingly complex resistance. This emergency has led to the search for new drugs. A group of highly virulent and resistant bacteria, which are known as the ESKAPE pathogens (Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, and Enterobacter species) has been identified by the World Health Organization (WHO) and highlights the need for alternative interventions against these bacteria (Huan et al., 2020). They are ubiquitous short molecules, generally cationic, with amphipathic properties, which are part of the innate immune systems of all living organisms. (Zasloff, 2007; Ma et al., 2024). AMPs act primarily through biophysical interactions, as opposed to conventional antibiotics that typically interact with a metabolic pathway/enzymatic reaction which microbes can circumvent by a single point mutation. They target the basic architecture of the microbial cell membrane (Su et al., 2025), including models like the carpet model, barrel-stave pore or toroidal pore, which causes the membrane to become permeable and ultimately to lyse. Resistance is harder and more energy-intensive to evolve for pathogens in this mechanism, which doesn't involve disruption of biochemical pathways, and AMPs are a cornerstone of the "post-antibiotic era" therapeutics (Kumar et al., 2018).

1.2 Plant-Derived AMPs: The Case for Brassica juncea

Plants are sessile organisms so they have developed elaborate chemical and biological defence mechanisms to overcome environmental stress and pathogen attacks. A major oilseed crop of the Brassicaceae family, Brassica juncea (Indian Mustard) is known for its potent defense arsenal. Although a large amount of research has been conducted on the nutritional value of B. juncea, the proteome of these seeds has untapped potential for bioactive peptides, particularly defense-related proteins and proteins that are stored in the seed. B. juncea is an allotetraploid species that is the result of hybridization between B. rapa and B. nigra (Fig. 1), creating a distinct evolutionary environment that is the focus of this study. The duplication event and subsequent diversification of the genome have probably resulted in the expansion of gene families that are involved in stress tolerance and defense, which makes it an ideal candidate for bioprospecting new AMPs.

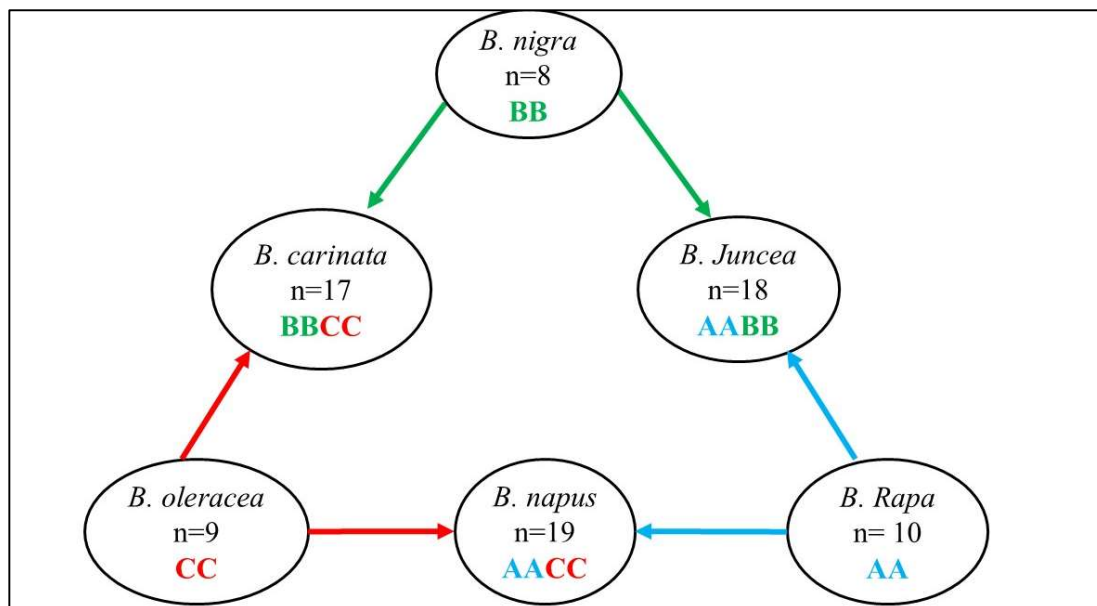


Fig. 1- Triangle of U

1.2.1 Seed Storage Proteins (SSPs) as Precursors

Seed storage proteins like 12S globulin, 2S albumin and Oleosin are produced during seed maturation to supply the germinating embryo with nitrogen and carbon sources. Recent work has indicated, however, that these proteins are not simply nutritive, but also that they carry "cryptic" peptides in the primary sequence. When they are present in the parent protein structure, they are biologically dormant and remain hidden in the protein structure. These pieces of protein can have strong antimicrobial, antioxidant, and antihypertensive effects when they are released by specific proteolytic processing either during natural seed germination or by in silico and in vitro enzymatic hydrolysis. The idea of "proteins as precursors" makes it possible to valorize common agricultural residues through a sustainable green pharmacology approach as a source of therapeutic molecules of high added value. (Kumar et al., 2011; Ayadi et al., 2022).

1.2.2 Defense Proteins: Chitinases and Defensins

Brassica juncea also synthesizes defense proteins such as Chitinases and Defensins in addition to the storage proteins. Chitinases are involved in resistance against fungal pathogens, which break down chitin in the fungal cell walls, and are one of the most effective barriers to fungal infection. They have broad-spectrum activity against a wide variety of bacteria and fungi and are very heat and protein resistant. The identification of these gene families by a genome-wide approach will be crucial to understanding their evolutionary expansion and functional diversity in the *B. juncea* genome (Wang et al., 2025).

These genes can be mapped and specific gene isoforms identified that have evolved a specialized function in the plant for protection during sensitive growth periods, such as seed maturation and early seedling establishment.

1.3 Objectives-

Objectives proposed for this study are-

➤ **Genome-Wide Identification and Evolutionary Characterization**

To identify and analyze the structural, functional and evolutionary landscape of seed protein families involved in storage (Cruciferin, Napin and Oleosin) and defense (Chitinase and Defensin), in the genome of *Brassica juncea*. This is to map the distribution of genes, analyze duplication events and examine regulatory elements to understand the genetic basis of these potential peptide precursors.

➤ **In Silico Proteolysis and AMP Validation with machine learning**

To develop antimicrobial peptides (AMPs) by computer-based simulated enzymatic hydrolysis and to test their therapeutic potential with a multi-algorithm machine learning approach. This goal involves screening peptide fragments with high confidence in antimicrobial activity (ANN, SVM and RF), and comprehensive in silico profiling of toxicity, allergenicity and physicochemical stability of potential candidates for safe development and subsequent experiments.

CHAPTER. 2

LITERATURE REVIEW

2.1 Importance of Antimicrobial peptides

AMPs are key components of plant innate immunity and are the first line of defense against a wide variety of pathogens. For *Brassica juncea* (Indian mustard), a leading oilseed crop, seed storage and defense proteins, such as napin (2S albumin) and cruciferin (12S globulin), respectively, have shown to be attractive sources of AMPs. These proteins are not only a source of nutrients for the plant but also display antimicrobial and insecticidal properties, which help the plant survive. With the ever-growing risk of crop diseases and the need for sustainable biocontrol solutions, the systematic identification and characterization of seed-derived AMPs in *B. juncea* is important by genome-wide and in silico approaches (Rahman et al., 2021; Munir et al., 2025).

2.2 Seed Storage and Defense Proteins in *Brassica juncea*

- **Napin (2S albumin)**
Small protein (15kDa) that is of two subunits held together by disulfide bonds. The antibacterial, antifungal, and insecticidal activity of Napin is strong, which is why it is a potential candidate for AMP in *B. juncea* (Padovan et al., 2010).
- **Cruciferin (12S globulin)**
A bigger seed storage protein, similar in structure to napin, but with no known in vitro antimicrobial activity. The defense function of cruciferin is less significant, perhaps because it is quaternary structure and has lower membrane interaction potential (Padovan et al., 2010).
- **Vicilin**
Recognized as a key storage protein in other members of the Brassica genus, which could be used for future AMP mining (Rahman et al., 2021).

2.3 Structural Basis of Antimicrobial Activity

- **Napin**
Has a small structure with alpha-helices and disulfide bridges. A long hypervariable region and a unique repeat element within it is responsible for greater antimicrobial and insecticidal properties (Munir et al., 2025).
- **Cruciferin**
Structurally larger and less compact, less stable than disulfide bonds, more stable by noncovalent interactions. Although in silico predicted to have antimicrobial activity, cruciferin has no in vitro activity (Rahman et al., 2020).

2.4 Mechanisms of Action

- **Membrane Disruption:** Napin is cationic and amphipathic, and interacts with and permeabilizes microbial membranes (Wu et al., 2018).
- **Enzyme Inhibition:** Molecular docking analysis revealed that napin inhibits the growth of pathogens by binding to bacterial enzymes (Rahman et al., 2020).
- **Insecticidal Activity:** Napin and related peptides bind to receptors in the midgut of insects, resulting in cytotoxic effects and mortality (Munir et al., 2025).
- **Trypsin and Translation Inhibition:** Disulfide-stabilized structure allows napin to inhibit proteases and translation (Ngai & Ng, 2004).

2.5 Evolutionary Conservation and Diversity

- AMP gene families in *B. juncea* have expanded via segmental and tandem duplications (Li et al., 2021).
- While cysteine-rich motifs are conserved, sequence divergence is common, reflecting adaptation to diverse pathogens (Hammami et al., 2008).
- *B. juncea* AMPs share structural domains with other Brassicaceae and model plants, but unique hypervariable regions and copy number variations suggest lineage-specific adaptation (Li et al., 2021).

2.6 Genome-wide Discovery Approaches in Antimicrobial Peptides

- **EST and Transcriptome Mining:** The use of EST and transcriptome mining assists in the identification of genes which can be expected to produce antimicrobial peptides in *Brassica juncea*. These methods give important clues to gene expression in various stress and pathogen scenarios (Hammami et al., 2008).
- **Integration with AMP Databases:** Using the AMP database, scientists can compare newly discovered peptide sequences to existing antimicrobial peptides. This integration enhances the prediction accuracy and accelerates the identification of possible AMPs (Hammami et al., 2008).

2.6.1 Bioinformatics Pipelines and Tools

- **RGAugury Pipeline:** The RGAugury pipeline is a widely-used tool for finding resistance-related genes and defense proteins in plant genomes. It facilitates large-scale genome analysis and enhances screening of AMP candidates (Yang et al., 2020).
- **QTL Mapping and SNP Arrays:** SNP arrays and QTL mapping can be helpful in identifying key genomic regions linked to disease resistance traits. These are ways that the researchers can understand the genetic variation associated with the expression of AMP (Basnet et al., 2015).
- **Gene Expression Profiling:** Gene expression profiling is the study of gene expression during stress, infection or environmental changes. It enables the identification of the genes involved in the plant defense response that are activated in the plant (Wei et al., 2023).

2.6.2 Experimental Validation of Predicted AMPs

- **In Vitro Assays:** In vitro assays are laboratory experiments that test the antimicrobial activity of predicted peptides against microbes. The tests give direct proof of the effectiveness of the peptide (Rahman et al., 2020).
- **Correlation with Physicochemical Properties:** The antimicrobial activity is significantly affected by physicochemical properties like charge, hydrophobicity and peptide length. Researchers can use their study of these properties to predict the stability and efficiency of AMPs (Rahman et al., 2020).

2.7 In Silico Characterization Methods for AMPs

2.7.1 Computational Tools and Databases

- **Sequence Mining:** Tools such as BLAST are employed to compare peptide sequences to existing AMP databases to find a sequence that matches. It can rapidly and efficiently identify possible antimicrobial peptides (Hammami et al., 2008).
- **Structural Prediction:** Peptide structure and properties are studied by using tools like ExPASy, ProtParam, SWISS-MODEL and SOSUI. Such predictions aid in the understanding of the stability and biological role of peptides (Muruhan et al., 2024).
- **Proteomics Integration:** Proteomics integration is the technique of identification and analysis of peptides in biological samples with mass spectrometry-based methods. It gives information on the composition and activity of peptides in detail (Moyer et al., 2021).

2.7.2 Machine Learning and Predictive Modeling

- **Model Development:** Random Forest, SVM and deep learning are used for prediction of antimicrobial activity with machine learning models. These models enhance the efficiency and precision of AMP identification (Jaiswal et al., 2022).
- **Validation:** The reliability of the prediction models is guaranteed by the use of validation methods, for example, by using cross-validation and laboratory testing. It is important to have experimental tests to confirm computational results (Jaiswal et al., 2022).

2.7.3 Molecular Docking and Dynamics Simulations

- **Interaction Analysis:** Molecular docking and dynamics studies contribute to understanding the molecular symathethic mechanisms of action-reaction of AMPs. These methods give information about the binding patterns and the mechanisms of action (Chandrashekar et al., 2021).

- **Limitations:** Simulations may not accurately mimic actual biological systems. Experimental validation is sometimes needed to interpret results, if they are to be correct (Chandrashekar et al., 2021).

2.8 Real-world Implications

- **Crop Protection and Disease Resistance**

The antimicrobial peptides (AMPs) of *Brassica juncea*, such as napin, have demonstrated great potential for defence. It has been reported that napin is capable of inhibiting the growth of several plant pathogens by disrupting the microbial cell membrane, which not only helps in minimizing loss of crops, but also leads to the decrease of the use of chemical pesticides.

Transgenic crops with the use of AMP-based strategies can offer long-term and widespread resistance. It helps reduce the use of pesticides, environmental pollution, and overall productivity, which promotes sustainable agriculture (Munir et al., 2025).

- **Biocontrol and Therapeutic Applications**

Napin has strong antimicrobial and insecticidal activity against a variety of microorganisms and insect pests. Microbial membrane interaction and inhibition of pathogen growth by these peptides have been suggested based on molecular studies, in which the peptides interact with essential enzymes in microbial cells.

The natural origin and wide-spectrum activity of AMPs make them potential eco-friendly biocontrol agents for use in agriculture. Moreover, due to their low toxicity and bioactive properties, there has been an interest in their use as pharmaceutical and therapeutic agents (Munir et al., 2025).

- **Sustainability and Resistance Management**

The plant derived AMPs are environmentally friendly due to their specificity and biodegradability, making them an alternative to synthetic pesticides. Their application may help to sustain crop protection and minimize the spread of undesirable chemicals in soil and water systems.

One more big benefit of AMPs is that they have minimal risk of resistance development in pathogens as compared to conventional pesticides. This makes them useful tools for maintaining long-term resistance and in environmentally responsible agricultural production (Montesinos, 2007).

CHAPTER 3

METHODOLOGY

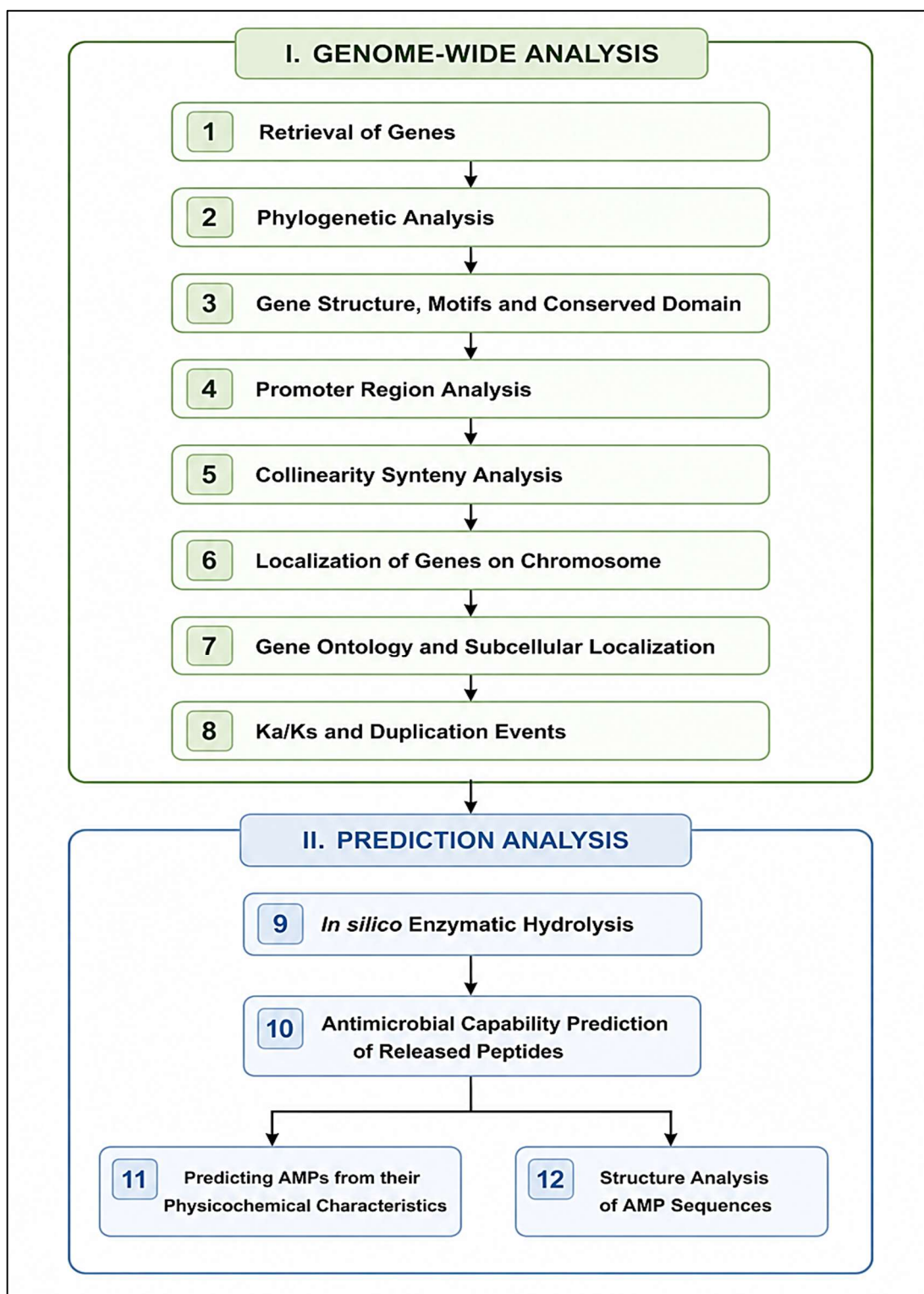


Fig. 2- Workflow of the study

3.1 Retrieval of Cruciferin, Napin, Oleosin, Defensin, and Chitinase Family Homologs in the *Brassica juncea* Genome

To identify members of the cruciferin, napin, oleosin, defensin, and chitinase gene families in *Brassica juncea*, well-annotated reference protein sequences were retrieved from the UniProt and NCBI databases. The sequences of cruciferin, napin, oleosin, defensin, and BjCHI1 chitinase were selected as query sequences. BLAST analysis was performed using the query sequences to identify homologous genes within the *Brassica juncea* genome using the *Brassica juncea* Information Resource database (Zhang et al., 2024). Finally, resulting in identification of homologous genes within the respective families. and their respective cDNA, genomic and peptide sequences were retrieved which were then used for further genome-wide characterization.

3.2 Phylogenetic analysis

The protein sequences were uploaded into the MEGA 12 software to find their evolutionary and phylogenetic relationships as seen in Fig.3. Multiple sequence alignment was executed by Clustal W. Phylogenetic analyses were performed using Jones-Talor-Thorton model and maximum likelihood approach with 1000 bootstrap replications. The MEGA12 software is used to build ML trees. (Kumar et al. 2024; Raturi et al., 2024). The iTOL v6 program is utilized for further annotation of the tree (Letunic & Bork, 2024).

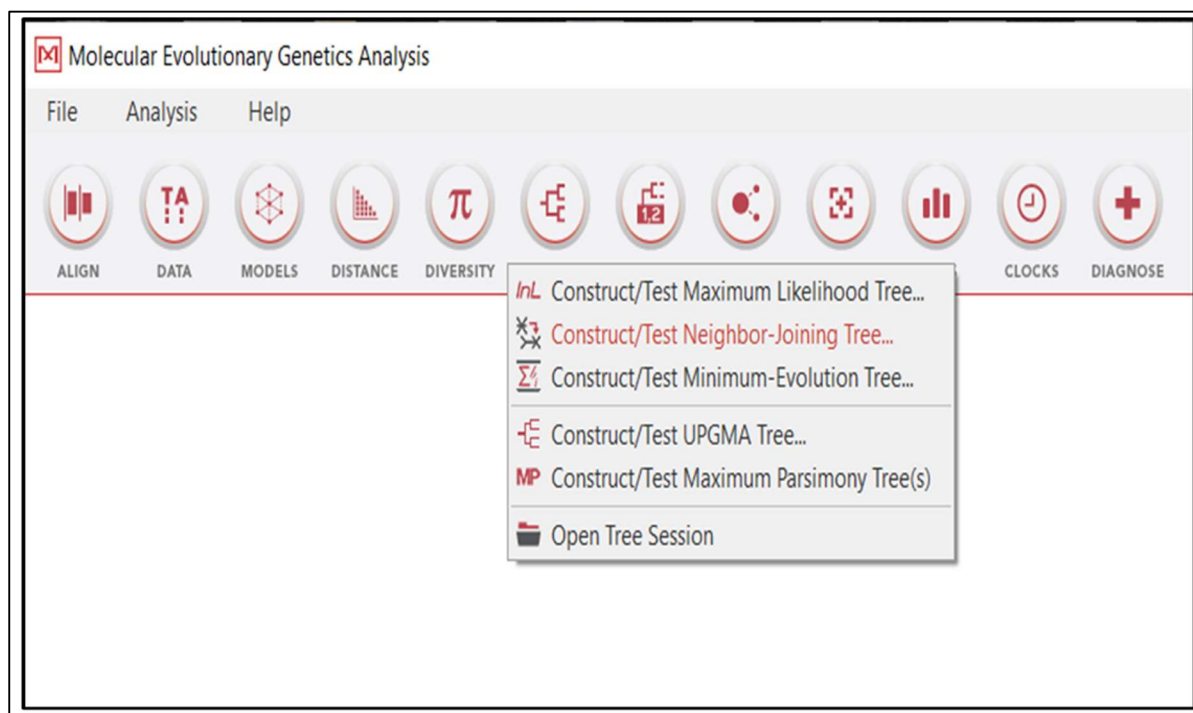


Fig. 3- GUI of MEGA software

3.3 Prediction Analysis of Gene Structure, Motifs and Conserved Domain

The Gene Structure Display Server (GSDS) 2.0 (Fig. 5) was utilized to identify and display the gene structures (Hu et al., 2014). The Genome and CDS sequences were uploaded to the server to get the results. Motifs were determined using the MEME suite by uploading the protein sequences (Bailey et al. 2015; Akbudak et al., 2021). Conserved domains were then predicted using the NCBI Conserved Domain Database (Fig. 4) and Pfam database, results were visualized using the TBtools (Chen et al., 2023).

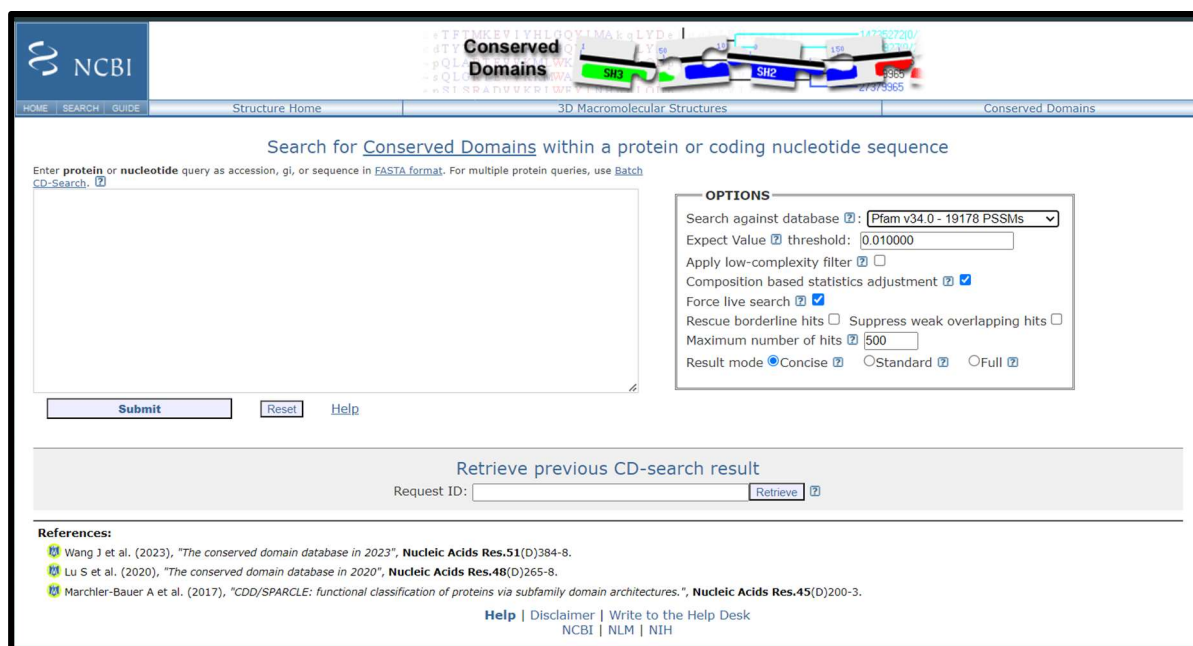


Fig. 4- GUI of NCBI Conserved-Domain Database

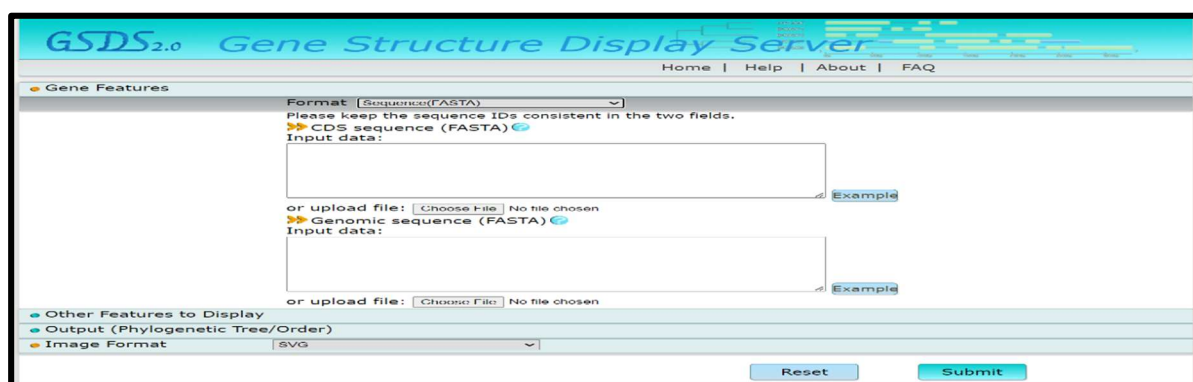


Fig. 5- GUI of GSDS 2.0 server

3.4 Promoter region analysis

To determine the cis-regulatory elements, 1500 base pairs upstream from the genomic region was manually selected from the Phytozome database (Shahbaz et al., 2023). PlantCARE database (Fig. 6) was then used to predict the cis-acting elements (Lescot et al., 2002). Using

the TBtools software a graph showing the locations of cis elements was created (Chen et al., 2023; Zia et al., 2024).

Menu

- Query CARE
- Search for CARE
- Other Queries
 - Classification...
 - Genes...
 - Name of Factor...
 - Name of Site...
 - Referencia...
- Motif Sampler
- Clustering
- Enter new data
- General References...
- Legal stuff...
- Our Home Page...
- Our References...

Search for CARE

email address to send the results back:

reference name or ID for the sequence (optional):

Sequence to submit:

Choose File No file chosen currently file size limited to 100Kb

NEW Search Reset Form Demo

★ To use our output you will need a modern browser for the implemented DHTML features. Earlier version might work partially (printing will not work properly)

★ Before using scripts dth=20 src='../html/GIF/star.gif' because of the increased usage of the site, I have been forced to implement a scheduling and results will be returned via email.

Scripts should not be used anymore as one can upload a multi-fasta file, but if you do, please be gentle, and do introduce a sleep of 60sec between each request in your scripts! Thanks. to submit a number of sequences through this page please contact [me](#).

(Those not respecting that will be blocked :-))

★ because of the increased usage of the site, I have implemented a scheduling and results will be returned via email.

This means that you might have to wait a little in order to get the results sent back to you, certainly when the load on the server is high. Please check you email spam folder if you haven't seen any email.

★ Please, use ASCII characters, roman alphabet, everything else will be trashed

Fig. 6- GUI of Plant Care server

3.5 Synteny analysis

A collinearity analysis was performed between *Brassica juncea* and three other species: *Arabidopsis thaliana*, *Brassica nigra* and *Brassica rapa* by downloading their complete genome sequence and gff3 files of all the species from the BjuIR database. Syntenic blocks were then identified and results were visualized using the MCScanX tool in TBtools (Fig. 7).

TBtools-II (Toolbox for Biologists) v2.472

Sequence Toolkit BLAST GO & KEGG Graphics Others About

?Report !Help Smart

Dual Synteny Plot for MCscanX

Set Ctl File
Drag and Drop an Input Ctl File Here, Which is the Same as That for Dual Plotter in MCscanX and Should be Manually Prepared.

Set Simplified GFF file
Drag and Drop an Input Simplified GFF File Here, Which is the Input File for MCscanX

Set the Input Collinearity File Here
Drag and Drop an Input Collinearity file, Which is the Resultant File from MCscanX

Gene List For Highlight (Optional)
Drag and Drop an ID List File Here

Other Parameters

- ☒ Filter Collinearity in the Up Genome
- ☒ Filter Collinearity in the Down Genome
- Filter Genes in Small Collinearity Block : 0

Start

Fig. 7- GUI of TBTools Dual Synteny Plot

3.6 Genes on Chromosome

The position of the identified genes and their respective chromosome number were downloaded from the BjuIR and the length of the chromosomes was fetched from NCBI Database. The graph was plotted using the chromosome position and length on Mapchart v2.2 (Voorrips, 2002).

3.7 Gene Ontology and subcellular localization

The identified transcript sequences were put into the Blast2Go software in order to interpret Gene Ontology (GO). With default parameters, operations such as functional annotation, BLAST search, InterProScan, annotation and mapping were executed. (Gotz and Conesa, 2008). The CELLO v 2.5 was used to perform protein subcellular localization using protein sequence data (Yu et al., 2006; Raturi et al., 2024).

3.8 Ka/Ks and duplication events

The TBtools software is used to calculate the Ka value and Ks value. These parameters were measured in order to assess the corresponding times of divergence and selective pressure (Wang et al., 2023). The first way of gene expansion is Tandem repeats that happens when two genes that are separated by fewer loci on the same chromosome and the other is segmental duplication which involves blocks of genes often located on different chromosomes (Laha et al., 2020; Wang et al., 2023).

3.9 In silico enzymatic hydrolysis

All the 29 identified genes protein sequence were subjected to hydrolyze under the enzyme action option present in the BIOPEP-UWM (Fig.8) (Minkiewicz et al., 2019). Digestive proteases chymotrypsin(A) (EC 3.4.21.1) and trypsin (EC 3.4.21.4) were the only one based on the findings of an earlier investigation on brassica genome (Duan et al., 2021). Using CAMP, the antimicrobial activity of released peptides with 10–60 AA residues was predicted (Duan et al., 2021).

Proteins (current number: 773)	Please cite the following paper if you are using BIOPEP-UWM database: Minkiewicz P, Iwaniak A, Darewicz M. BIOPEP-UWM Database of Bioactive Peptides: Current Opportunities. <i>International Journal of Molecular Sciences</i> . 2018; 20(23):5978. doi: 10.3390/ijms20235978
Bioactive peptides (current number: 5684)	
Allergenic proteins with their epitopes (current number: 136)	Please cite the following paper if you are using BIOPEP-UWM database of allergenic proteins and their epitopes: Dziuba M, Minkiewicz P, Dębek M. Peptides, specific proteolysis products as molecular markers of allergenic proteins – in silico studies. <i>Acta Scientiarum Polonorum Technologia Alimentaria</i> , 12, 2013, 101-112
Sensory peptides and amino acids (current number: 605)	Please cite the following paper if you are using BIOPEP-UWM database of sensory peptides and amino acids: Iwaniak A, Minkiewicz P, Darewicz M, Sieniewski K, Starowicz P. 2016. BIOPEP database of sensory peptides and amino acids. <i>Food Research International</i> , 95, 155-161. doi: 10.1016/j.foodres.2016.04.031
BIOPEP-UWM Virtual (current number: 682)	Our new database. Please cite the following paper if you are using BIOPEP-UWM Virtual database: Minkiewicz P, Iwaniak A, Darewicz M. 2022. BIOPEP-UWM Virtual—A Novel Database of Food-Derived Peptides with In Silico-Predicted Biological Activity. <i>Applied Sciences</i> , 12(14):7204. doi: 10.3390/app12147204
BIOPEP-UWM repository of amino acids and modifications (current number: 228)	BIOPEP-UWM repository of amino acids and modifications.
Submit new peptide sequence	Here you can send us the peptide sequence that is currently not present in our database(s). [Thank you.]

Number of visits from 2024.01.01 - 97128 [Open BIOPEP-UWM in new window](#) Last update - 2026-05-07

Financial Support of BIOPEP-UWM Database

1. University of Warmia and Mazury in Olsztyn, project No 17.610.014-110
2. Designated subsidy of the Minister of Science and Higher Education Republic of Poland, task entitled:
The Research Network of Life Sciences Universities for the Development of the Polish Dairy Industry –
Research Project.

Fig. 8- GUI of BIOPEP-UWM database

3.10 Antimicrobial capability prediction of released peptides

In the CAMP database (Fig. 9), the AMP activity of peptides was predicted using machine learning methods. Three algorithms—Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Random Forests (RF)—were implemented as described in Gawde et al. (2022) and Duan et al. (2021).

CAMP_{R4}
Collection of Anti-Microbial Peptides

<http://camp.bicnirrh.res.in>

Database architecture


AMP structures: 933


Patterns: 84
HMMs: 179


Patents: 2143


Natural AMPs: 11827
Synthetic AMPs: 12416

❖ Manually curated entries from PubMed, UniProt & NCBI protein

❖ 15425 antibacterial, 4783 antifungal, 1233 antiviral peptides

❖ Improved annotations for activity (anticancer, antiparasitic), peptide modifications (N & C-terminal) and nature of peptide (monomer, dimer, cyclic)

❖ Signatures for 8 new AMP family: Gurmarin, Macin, Magalinin, Nigrocin, Pardaxin, Piscidin, Ranacyclin, Stomoxyn


AMP prediction accuracy
Natural AMPs: 86%
Synthetic AMPs: 94%

❖ Novel prediction algorithms for **natural** and **synthetic** peptides

❖ **Antimicrobial region prediction** based on natural and synthetic peptides

❖ Tool for **rational design** of natural and synthetic peptides

Demonstrated utilities*

DATABASE


- ❖ Creation of other AMP databases – ADAM, InverPep, YADAMP, LAMP, C-PAMP, Hemolytik, ANTISTAPHYBASE
- ❖ Development of new algorithms – Joker, an algorithm for rational design of AMPs
- ❖ Identification of novel peptides from AMP families – antiviral peptide from *Acacia catechu*, differentially expressed peptides from serous ovarian cancer tissues

ALGORITHM


- ❖ Identification of peptides from natural sources – peptides from *Protaetia brevitarsis* larvae, Sichuan pepper, *Litopenaeus vannamei*, milk proteins, human sweat, Thermophilic *Geobacillus* sp. Strain ZGt-1,51, *Varanus komodoensis* (Komodo Dragon), American alligator plasma
- ❖ Rational design of AMPs - 17 peptides from 10,000 double mutants of Bacteneicin 2A

*Reference: Waghu, FH & Idicula-Thomas, S., 2020 <https://doi.org/10.1002/pro.3714>

Fig. 9- GUI of CAMP R4 Database

3.11 Predicting AMPs from their physicochemical characteristics

The Antimicrobe Peptide Database (APD) was used to determine the AMPs' physicochemical parameters—including molecular mass, hydrophobic ratio and Boman index—following the procedure of Duan et al. (2021). In addition, the aliphatic index, instability index, and isoelectric point were estimated with ProtParam on the ExPASy server (Akbulak et al., 2021; Gasteiger et al., 2005).

3.12 Structure analysis of the AMP sequences

The secondary structure of the AMPs was analyzed using the SOPMA program (Combet et al., 2000; Wang et al., 2023).

3.13 Computational Analysis of AMP Allergenicity and Toxicity

To check the toxicity and allergenicity of the identified AMPs to ensure their safety for future *in vitro* testing, AlgPred 2.0 and CSM-Toxin (Fig. 10) online platforms were used. For predicting the allergenicity, we use the Hybrid model (Morozov et al., 2023; Sharma et al., 2020).



Fig. 10- GUI of CSM-Toxin Server

CHAPTER 4

RESULT AND DISCUSSION

4.1 Retrieval of Cruciferin, Napin, Oleosin, Defensin, and Chitinase Homologs

A total of twenty-nine candidate genes belonging to the cruciferin, napin, oleosin, defensin, and chitinase families were identified through BLAST-based searches against the *Brassica juncea* genome within appropriate cutoffs.

4.2 Phylogenetic Analysis of the Gene Families

To investigate evolutionary relationships, *Brassica juncea* (AABB) with its parental counterparts *Brassica rapa* (AA) and *Brassica nigra* (BB), as well as with *Arabidopsis thaliana*, a model dicot plant. (Meinke et al., 1998; The *Brassica juncea* Genome, 2022).

Two phylogenetic trees were constructed. The first phylogenetic tree was constructed using the twenty-nine identified *Brassica juncea* candidate genes whereas the second tree was constructed using homologous genes of the cruciferin, napin, oleosin, defensin, and BjCHI1 (chitinase) families, which were obtained by BLAST searches against the genomes of *Brassica rapa* (AA), *Brassica nigra* (BB), and *Arabidopsis thaliana*, under the same stringent. (The *Brassica juncea* Genome, 2022).

Analysis of the internal phylogenetic tree analysis indicated that the oleosin family represented the most divergent lineage, whereas the defensin, napin and cruciferin families despite internal clustering, appeared to share a more recent common ancestor. The comparative phylogenetic tree further supported the allotetraploid origin of *Brassica juncea*, as its genes consistently clustered with orthologs from both diploid progenitors-*Brassica rapa* (A genome) and *Brassica nigra* (B genome) and with *Arabidopsis thaliana*, representing an evolutionarily conserved dicot lineage.

4.3 Identification of Motifs, Domains, and Structures

Conserved domain analysis, visualized using TBtools, revealed the presence of multiple domains, including Glyco_hydro_19, Chitin_bind_1, Cupin_1, Gamma-thionin, and Oleosin.

MEME analysis of ten conserved motifs revealed that identified gene families share distinct, conserved motif organization patterns. This structural conservation was most pronounced in the napin (BjuNa) family, where all twelve members displayed a nearly identical motif architecture. Although the defensin (BjuDe) and oleosin (BjuOl) families also exhibited high motif conservation, the cruciferin (BjuCr) and chitinase (BjuCHI) families exhibited greater structural divergence in their motif composition.

Most of the identified genes were within the 1 kb range in length. Among these BjuDe04 was the longest gene at 12 kb. BjuNa07, BjuNa08, BjuNa03, and BjuNa09 were identified to be intronless.

A similar observation in barley hevein genes—mostly intron-less or with a single intron—supports the hypothesis that intron loss may facilitate their rapid gene expression in response to biotic and abiotic stresses (Drikvand et al., 2024).

4.4 Cis-analysis of the genes

The identified cis-regulatory elements were primarily associated with hormone responsive pathways, including abscisic acid-responsive elements (ABRE, ABRE3a, ABRE4, AT~ABRE), ethylene-responsive elements (ERE), wound signaling (WUN-motif, WRE3), and pathogen defense related elements (Hobo et al., 1999; Nakashima et al., 2009).

The frequent occurrence of MYC elements leads to involvement in Jasmonic Acid (JA) signalling, whereas W-box motifs indicate regulation by WRKY transcription factors, which are key regulators of pathogen defense responses. Additional motifs like Dehydration-responsive elements (DRE) and MYB binding sites (MBS) suggest potential cross-talk with abiotic stress-responsive pathways. Overall, these results suggest that identified genes may transcriptionally regulated by jasmonic acid, ethylene, and ABA signaling networks that are crucial for AMP induction. (Eulgem & Somssich 2007; Lorenzo et al., 2004).

4.5 Physical mapping of Chromosomes location and Ka/ks ratio analysis

Gene duplication is a fundamental evolutionary mechanism contributes to gene family expansion and functional diversification (Ohno, 1970). The *Brassica juncea* genome has undergone multiple whole-genome duplication and chromosomal rearrangement events, which have contributed to gene-level evolution (Wang et al., 2025). Gene duplication patterns were analyzed to elucidate the genome expansion mechanism of the five defense and storage-associated gene families (Cruciferin, Napin, Oleosin, Defensin, and Chitinase) in *Brassica juncea* as shown in (Table 1). Nine duplicated gene pairs were identified across these families: seven were segmentally duplicated (SD)—namely BjuOl01/BjuOl02, BjuCr03/BjuCr04, BjuCHI02/BjuCHI03, BjuDe05/BjuDe06, BjuDe04/BjuDe02, BjuNa09/BjuNa11, BjuNa08/BjuNa07—and two were tandemly duplicated (TD): BjuNa01/BjuNa02 and BjuNa06/BjuNa04. These duplicated genes are distributed across both A and B subgenomes.

All ratios were less than 1, indicating that these gene families have predominantly evolved under strong purifying selection. This evolutionary pressure suggests that the structural integrity of these storage and defense proteins is critical for plant survival. Despite the expansion of the *Brassica* genome, the sequences have been conserved to maintain their essential physiological roles, this suggests that any mutations altering their bioactivity were likely deleterious and eliminated through purifying selection.

Table 1 Summary of the ka/ks ratio analysis for the candidate gene pair present in *Brassica juncea*.

Genes	Location	Duplicate	Ka	Ks	Ka/Ks
BjuOl01/BjuOl02	B05/A02	SD	0.0708	0.1678	0.4218
BjuCr03/BjuCr04	A02/B04	SD	0.0276	0.2368	0.1165
BjuCHI02/BjuCHI03	B07/A03	SD	0.0722	0.2479	0.2914
BjuDe05/BjuDe06	A09/B04	SD	0.0113	0.0803	0.1405
BjuDe04/BjuDe02	A02/B04	SD	0.0285	0.1025	0.2779
BjuNa09/BjuNa11	B05/B08	SD	0.0414	0.2266	0.1825
BjuNa01/BjuNa02	B05/B05	TD	0.0165	0.034	0.4865
BjuNa08/BjuNa07	A01/B05	SD	0.0255	0.1917	0.133
BjuNa06/BjuNa04	A01/A01	TD	0.005	0.0364	0.1368

4.6 Synteny analysis

To examine genomic conservation and chromosomal rearrangements, synteny maps were constructed between *Brassica juncea* and *Arabidopsis*, *Brassica*. The *Arabidopsis* and *Brassica juncea* genomes exhibited approximately 55% collinearity (16 syntenic pairs). For the *Brassica rapa* it showed 68.97% collinearity (20 syntenic pairs). Likewise, between *Brassica nigra* and *Brassica juncea* showed 62.07% collinear genes while sharing 18 syntenic gene pairs.

Collinearity analysis against the model *Arabidopsis thaliana* showed that most orthologs mapped to two or more syntenic blocks across the *Brassica juncea* A and B subgenomes, confirming the expansive effect of the ancestral whole-genome triplication (WGT) on these families (Schrantz et al., 2006).

Synteny with the diploid progenitors, *Brassica rapa* (A genome) and *Brassica nigra* (B genome), further supported the allotetraploid origin of *Brassica juncea* (AABB) via hybridization (Parkin et al., 2010).

4.7 Gene Onto and sub-localization

Gene Onto (GO) and InterProScan analysis revealed that the twenty-nine identified genes were primarily linked to seed storage protein families, such as oleosins (BjuOl), 2S albumins (BjuNa), and cupins (BjuCr) consistent with their roles in nutrient accumulation during seed development. (de Souza Cândido et al., 2014; Souza, 2020).

Although cupins (Cruciferin, BjuCr) primarily function as storage proteins, the superfamily to which they belong is known for its diversity of function, and some studies suggest an intrinsic antimicrobial potential or structural similarity to known AMPs (Rahman et al., 2020). Defensin genes (BjuDe), on the other hand, are annotated as γ -thionins, which are classical cysteine-rich plant antimicrobial peptides that disrupt microbial membranes and localise to

extracellular areas, representing a major class of AMPs (Nawrot et al., 2013; Ishaq et al., 2019). Chitinase genes (BjuCHI) exhibited chitinase activity and chitin binding. They map to chitin catabolic process and defence response to fungus. Chitinases are recognized as antimicrobial proteins that hydrolyze fungal cell wall components. Antifungal activities have been demonstrated for the *Brassica juncea* chitinase BjCHI1 (Guan et al., 2008). In contrast, oleosin genes (BjuOl) are primarily associated with lipid storage and oil body structure rather than being direct antimicrobial activity and may indirectly contribute to stress resilience through the stabilization of energy reserves during seed development (Shao et al. 2019).

4.8 Computational enzymatic hydrolysis of the identified genes

Proteolysis of the identified proteins was performed in silico using the digestive proteases chymotrypsin A (EC 3.4.21.1) and trypsin (EC 3.4.21.4) through the BIOPEP-UWM database (Minkiewicz et al., 2019). These enzymes were selected for their high AMP-generation potential based on the results of a prior study on brassica species (Duan et al., 2021). Following enzymatic hydrolysis, chymotrypsin(A) and trypsin generated a total of 1559 and 637 fragments, respectively. Chymotrypsin A digestion generated a higher number of predicted peptide fragments than trypsin, likely due to its greater number of cleavage recognition sites (3) than trypsin (2). Of these fragments, 505 derived from chymotrypsin(A) and 337 from trypsin were in the range of 10-60aa length. The peptide fragments were predominantly in the 10–60 amino acid range (average length: 33.26) (Huan et al., 2020). A total of 54 fragments were predicted to be AMP after removing the redundancy the final no. of unique AMP was 37 which were then used for further studies.

4.9 Estimation of AMP potential of released peptides

Three statistical models were utilized to assess antimicrobial peptides using the CAMP database. According to Gawde et al. (2022), the reported prediction accuracies for natural AMPs were 86.5%(RF), 84.1%(SVM) and 82.2%(ANN) respectively. Trypsin (42) was predicted to be the most efficient enzyme for generating AMPs compared to chymotrypsin(A)(12).

These amino acids(aa) have positive charges. At the termini of the peptides being released, they play an important role in determining the net cationic charge of the peptide (Zasloff, 2007). Cationic charge and amphipathicity are the two most defining characteristics of typical AMP, as they enable electrostatic interaction followed by disruption of negatively charged microbial cell membranes (Ma et al., 2024).

Notably, Defensin-derived peptide (BjuDe03) and Napin-derived peptides (BjuNa03) and (BjuNa06), all three cleaved by Trypsin, exhibited the highest average confidence scores across all three models. Their consistent high scoring across all models highlights them as high-priority candidates for future in vitro synthesis and experimental validation.

4.10 Comparative physicochemical profiling of predicted AMP candidates

AMP activity is influenced by factors such as net charge, peptide hydrophobicity and amphipathicity, and presence of cationic residues (P. Kumar et al., 2018; Duan et al., 2021).

All AMP candidates exhibit a typical molecular weight range of 911 Da to 4284 Da. The isoelectric point (pI) ranged from 4.53 to 12.02, with the prevalence of pI >7 promoting electrostatic attraction to microbial membranes, a hallmark of most AMPs (Jenssen et al., 2006). The charge was found between the -1 and +7. Although a strong net positive charge (+4 to +7) plays a significant role in the initial membrane binding, even peptides with a lower net positive charge (± 1) may also exhibit high antimicrobial potential, provided that the distribution of hydrophobic and hydrophilic residues is optimized to induce amphipathicity (Powers and Hancock, 2003).

A broad GRAVY index range indicated diverse membrane-interaction strategies among the predicted peptides. AMPs generally require a certain degree of hydrophobicity to partition into lipid bilayer (Huang, 2006). High aliphatic index values suggest thermal stability, and the Boman index variation suggests an ability to form different ordered structures (e.g., α -helical or β -sheet) upon membrane interaction, consistent with diverse AMP families (Bowman et al., 2005).

Table 2 Physicochemical profiling of antimicrobial peptides using in silico tools

S. No	Source	length	charge	iso	Wt	Bo index	Instability	Aliphatic	Hydrophobic %	G*
1	BjuCHI01	20	0	5.99	2003.32	0.1	51.89	112	35%	0.34
		16	1	8.2	1828	0.82	7.07	48.75	31%	-0.47
2	BjuCHI02	16	1	8.2	1802	0.71	33.94	55	38%	-0.256
		20	0	5.99	2017	0.06	45.72	117	40%	0.45
3	BjuCHI03	13	0	5.83	1452	-0.05	27.45	53	54%	0.6
		28	-1	4.53	2942	0.94	20.22	38.52	39%	-0.21
4	BjuCr01	12	1	8.59	1330	-0.29	41.27	73.33	42%	0.46
		22	0	5.81	2334	0.31	26.62	105.91	41%	0.36
5	BjuCr02	18	1	9.18	2045	1.17	63.42	65	33%	-0.48
		22	0	5.81	2332	0.03	30.05	115	45%	0.46
6	BjuCr03	44	1	8.59	4284	2.52	32.21	0	2%	-1.98
7	BjuCr04	21	1.25	8.76	2235	-0.7	9.71	148.57	52%	1.24
		22	0	5.81	2318	0.07	26.62	110.45	45%	0.48
8	BjuDe01	22	0.25	6.72	2196	1.25	30.2	30.91	32%	-0.3
9	BjuDe02	22	-0.75	5.4	2197	1.58	42.65	30.91	32%	-0.75
10	BjuDe03	22	0.25	6.72	2128	1.28	31.95	26.36	27%	-0.35
11	BjuDe04	22	-0.75	5.4	2193	1.47	43.08	30.91	32%	-0.55
12	BjuDe05	16	2.25	10.35	911	3.95	27.44	36.25	25%	-1.03
13	BjuDe06	10	4.25	9.8	1354	6.32	175.42	0	30%	-1.57
14	BjuNa01	25	4	11.26	2751	2.98	65.06	50.8	20%	-1.64
15	BjuNa02	21	1	8.75	2408	-0.82	-9.28	139.52	67%	1.52
16	BjuNa03	17	4	11.26	1872	2.74	41.21	45.88	29%	-1.26
		21	1	8.75	2408	-0.82	-5.69	139.52	67%	1.5
		14	1	8.03	1562	0.29	74.5	90	50%	0.46
17	BjuNa04	15	3	10.3	1599	2.21	50.4	58.67	27%	-1.28
		21	1	8.75	2394	-0.78	-9.28	134.76	67%	1.5
18	BjuNa05	22	4	11.26	2424	2.97	60.38	53.18	23%	-1.54
19	BjuNa06	14	1	8.03	1576	0.22	74.5	97.14	50%	0.49
20	BjuNa07	20	3.25	10.3	2177	2.68	59.56	44	20%	-1.63
21	BjuNa11	25	5	12.02	2712	3.24	74.48	31.2	20%	-1.64
		21	1	8.22	2316	-0.79	-9.28	153.33	67%	1.54
22	BjuNa12	19	4	11.26	2142	2.99	62.17	46.32	26%	-1.48
		21	1	8.75	2360	1.56	-9.28	153.33	67%	1.56
23	BjuOl01	30	0	6	2478	-1.06	80.61	86	67%	0.79
24	BjuOl02	40	1	8.75	3234	-1.28	81.82	82	68%	0.83
25	BjuOl03	14	3	10.29	1596	1.4	29.76	97.14	36%	-0.34
26	BjuOl04	14	7	11.75	1774	5.72	85.41	27.86	14%	-2.3

4.11 Secondary Structure Prediction of AMP's

The α -helical structures, a common functional motif in antimicrobial peptides (Huang et al., 2010), was predicted in only fourteen AMP sequences. The α -helical AMPs are critical for forming the amphipathic structures, possessing distinct hydrophilic and hydrophobic faces that mediate their broad-spectrum membrane disruption through lytic mechanisms such as the carpet or toroidal pore models (Wang et al., 2015). The predominant structural patterns observed were non-ordered conformations, with nineteen sequences predicted to contain β -turns and twenty-three sequences exhibiting extended strand conformations.

Additionally, eleven AMP candidates consisted exclusively of a random coil conformation. The very high occurrence of random coil, β -turn, and extended conformations may suggest a strong preference for non-lytic mechanisms or structures with induced-fit folding. Such induced-fit mechanism facilitates interactions with lipid bilayers, particularly in fungal pathogens, where structural adaptability is often required to penetrate the cell wall matrix (Akbulak et al., 2021; Jenssen et al., 2006).

4.12 Computational Analysis of AMP Allergenicity and Toxicity

Assessing the safety of antimicrobial peptides derived from *Brassica juncea* protein is essential, as peptide toxicity is a common associated with AMPs. (Tu et al. 2018). According to predictions from the CSM-Toxin platform (Morozov et al., 2023), almost all peptide sequences were predicted to be non-toxic, except for one derived from BjuCHI03 which was predicted to be potentially toxic. The overall low predicted toxicity suggests that most candidates are suitable for further experimental investigation. In contrast, allergenic potential—particularly relevant for food-derived peptides—remains an important concern (Lin et al., 2018).

AlgPred2.0 Hybrid model identified some potentially allergenic AMPs, with BjuCHI03 was predicted to be the most allergenic. These candidates require in vitro verification before any application, as computational predictions must be experimentally validated to confirm the safety and non-immunogenicity of *Brassica juncea*-derived AMPs for application.

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

In the study, a comprehensive genome and computational framework was applied to identify and characterize antimicrobial peptide (AMP) precursors derived from seed storage and defense protein families in *Brassica juncea*. A total of thirty-seven unique antimicrobial peptide candidates were identified from 5 different gene families 14 from napin, 7 from cruciferin, 4 from oleosin and 6 each from chitinase and defensin. These genes were systematically analyzed using a genome-wide identification framework, followed by in silico enzymatic hydrolysis and machine learning based prediction to evaluate the antimicrobial activity of potential peptide fragments. Most candidates exhibited characteristic physicochemical features of functional AMPs, including a positive net charge essential for membrane interaction, and all contained a random coil conformation, suggesting high structural flexibility.

Furthermore, computational safety assessments indicated that out of the thirty-seven candidates, only one AMP candidate was found to be toxic and ten were predicted to exhibit allergenic potential. Collectively, these findings illustrate that the seed storage and defense protein families of *Brassica juncea* represent a potential genomic reservoir of putative AMP precursors for the discovery and development of novel plant-derived AMPs for future biotechnological applications.

REFERENCES

- Zhang, L., Xiao, J., Liang, C., Chen, Y., Yu, C., Zhao, X., Li, J., Yan, M., Yang, Q., Chen, H., Liu, Z., Wan, Z., Yang, Z., & Yang, Q. (2024). BjuIR: A multi-omics database with various tools for accelerating functional genomics research in *Brassica juncea*. *Plant Communications*, 5(8), 100925. <https://doi.org/10.1016/j.xplc.2024.100925>
- Kumar, S., Stecher, G., Suleski, M., Sanderford, M., Sharma, S., & Tamura, K. (2024). MEGA12: Molecular Evolutionary Genetic Analysis version 12 for adaptive and green computing. *Molecular Biology and Evolution*, 41(12). <https://doi.org/10.1093/molbev/msae263>
- Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, 52(W1), W78–W82. <https://doi.org/10.1093/nar/gkac268>
- Hu, B., Jin, J., Guo, A., Zhang, H., Luo, J., & Gao, G. (2014). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 31(8), 1296–1297. <https://doi.org/10.1093/bioinformatics/btu817>
- Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. *Nucleic Acids Res* 43: W39–W49. <https://doi.org/10.1093/nar/gkv416>
- Wang, G., Li, X., & Wang, Z. (2015). APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Research*, 44(D1), D1087–D1093. <https://doi.org/10.1093/nar/gkv1278>
- Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., Liu, Y., Feng, J., Chen, H., He, Y., & Xia, R. (2023). TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Molecular Plant*, 16(11), 1733–1742. <https://doi.org/10.1016/j.molp.2023.09.010>
- Lescot, M. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, 30(1), 325–327. <https://doi.org/10.1093/nar/30.1.325>
- Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity*, 93(1), 77–78. <https://doi.org/10.1093/jhered/93.1.77>
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *Humana Press eBooks* (pp. 571–607). <https://doi.org/10.1385/1-59259-890-0:571>
- Yu, C. S., Chen, Y. C., Lu, C. H., & Hwang, J. K. (2006). Prediction of protein subcellular localization. *Proteins*, 64(3), 643–651. <https://doi.org/10.1002/PROT.21018>
- Conesa, A., & Götz, S. (2008). Blast2GO: a comprehensive suite for functional analysis in plant genomics. *International Journal of Plant Genomics*, 2008, 1–12. <https://doi.org/10.1155/2008/619832>
- Gawde, U., Chakraborty, S., Wagh, F. H., Barai, R. S., Khanderkar, A., Indraguru, R., Shirsat, T., & Idicula-Thomas, S. (2022). CAMPR4: a database of natural and synthetic antimicrobial peptides. *Nucleic Acids Research*, 51(D1), D377–D383. <https://doi.org/10.1093/nar/gkac933>
- Laha, S. D., Dutta, S., Schäffner, A. R., & Das, M. (2020). Gene duplication and stress genomics in Brassicas: Current understanding and future prospects. *Journal of Plant Physiology*, 255, 153293. <https://doi.org/10.1016/j.jplph.2020.153293>

- Morozov, V., Rodrigues, C. H. M., & Ascher, D. B. (2023). CSM-Toxin: A Web-Server for Predicting protein toxicity. *Pharmaceutics*, 15(2), 431. <https://doi.org/10.3390/pharmaceutics15020431>
- Duan, X., Zhang, M., & Chen, F. (2021). Prediction and analysis of antimicrobial peptides from rapeseed protein using in silico approach. *Journal of Food Biochemistry*, 45(4). <https://doi.org/10.1111/jfbc.13598>
- Combet, C., Blanchet, C., Geourjon, C., & Deléage, G. (2000). NPS@: Network Protein Sequence Analysis. *Trends in Biochemical Sciences*, 25(3), 147–150. [https://doi.org/10.1016/s0968-0004\(99\)01540-6](https://doi.org/10.1016/s0968-0004(99)01540-6)
- Meinke, D. W., Cherry, J. M., Dean, C., Rounsley, S. D., & Koornneef, M. (1998). *Arabidopsis thaliana*: A Model Plant for Genome Analysis. *Science*, 282(5389), 662–682. <https://doi.org/10.1126/science.282.5389.662>
- Drikvand, R. M., Sohrabi, S. M., Sohrabi, S. S., & Samiei, K. (2024). Molecular Identification and Characterization of Hevein Antimicrobial Peptide Genes in Two-Row and Six-Row Cultivars of Barley (*Hordeum vulgare* L.). *Biochemical Genetics*, 62(6), 5092–5114. <https://doi.org/10.1007/s10528-024-10695-8>
- Finn, R. D. (2005). Pfam: the protein families database. *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*. <https://doi.org/10.1002/047001153x.g306303>
- Lorenzo, O., Chico, J. M., Saénchez-Serrano, J. J., & Solano, R. (2004). JASMONATE-INSENSITIVE1 Encodes a MYC Transcription Factor Essential to Discriminate between Different Jasmonate-Regulated Defense Responses in Arabidopsis[W]. *The Plant Cell*, 16(7), 1938–1950. <https://doi.org/10.1105/tpc.022319>
- Eulgem T, Somssich IE. Networks of WRKY transcription factors in defense signaling. *Curr Opin Plant Biol*. 2007;10(4):366–371. <https://doi.org/10.1016/j.pbi.2007.04.020>
- Hobo, T., Asada, M., Kowyama, Y., & Hattori, T. (1999). ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. *The Plant Journal*, 19(6), 679–689. <https://doi.org/10.1046/j.1365-313x.1999.00565.x>
- Nakashima K, Tran LS, Van Nguyen D, Fujita M, Maruyama K, Todaka D, Ito Y, Hayashi N, Shinozaki K, Yamaguchi-Shinozaki K. Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. *Plant J*. 2009;51(4):617–630. <https://doi.org/10.1111/j.1365-313X.2007.03168.x>
- Solano R, Stepanova A, Chao Q, Ecker JR. Nuclear events in ethylene signalling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes Dev*. 1998;12(23):3703–3714. <https://doi.org/10.1101/gad.12.23.3703>
- Xu, L., Liu, F., Lechner, E., Genschik, P., Crosby, W. L., Ma, H., Peng, W., Huang, D., & Xie, D. (2002). The SCFCO1 Ubiquitin-Ligase complexes are required for jasmonate response in Arabidopsis. *The Plant Cell*, 14(8), 1919–1935. <https://doi.org/10.1105/tpc.003368>
- Ohno, S. (1970). Evolution by gene duplication. In *Springer eBooks*. <https://doi.org/10.1007/978-3-642-86659-3>
- Guo, X., Mandáková, T., Trachtová, K., Özüdoğru, B., Liu, J., & Lysak, M. A. (2020). Linked by ancestral bonds: multiple Whole-Genome duplications and reticulate evolution in a Brassicaceae tribe. *Molecular Biology and Evolution*, 38(5), 1695–1714. <https://doi.org/10.1093/molbev/msaa327>

Wang, T., Van Dijk, A. D., Cai, X., Wu, J., Bonnema, G., & Wang, X. (2025). Brassica diversity through the lens of polyploidy: genomic evolution, introgression, and homoeologous exchange. *Horticultural Plant Journal*. <https://doi.org/10.1016/j.hpj.2025.08.002>

Yang, Z. (1998). Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular Biology and Evolution*, 15(5), 568–573. <https://doi.org/10.1093/oxfordjournals.molbev.a025957>

Zou, Z., Zhang, L., & Zhao, Y. (2024). Integrative Analysis of Oleosin Genes Provides Insights into Lineage-Specific Family Evolution in Brassicales. *Plants*, 13(2), 280. <https://doi.org/10.3390/plants13020280>

De Souza Cândido, E., Cardoso, M. H. E. S., Sousa, D. A., Viana, J. C., De Oliveira-Júnior, N. G., Miranda, V., & Franco, O. L. (2014). The use of versatile plant antimicrobial peptides in agribusiness and human health. *Peptides*, 55, 65–78. <https://doi.org/10.1016/j.peptides.2014.02.003>

Sanchez-Reinoso, Z., García-Vela, S., Clément, J., & Bazinet, L. (2024). Combining statistical, machine learning and experimental approaches for screening of novel antimicrobial peptides of calf cruor hydrolysates. *Food Bioscience*, 65, 105734. <https://doi.org/10.1016/j.fbio.2024.105734>

Souza, P. F. (2020). The forgotten 2S albumin proteins: Importance, structure, and biotechnological application in agriculture and human health. *International Journal of Biological Macromolecules*, 164, 4638–4649. <https://doi.org/10.1016/j.ijbiomac.2020.09.049>

Guan, Y., Ramalingam, S., Nagegowda, D., Taylor, P. W. J., & Chye, M. (2008). *Brassica juncea* chitinase BjCHI1 inhibits growth of fungal phytopathogens and agglutinates Gram-negative bacteria. *Journal of Experimental Botany*, 59(12), 3475–3484. <https://doi.org/10.1093/jxb/ern197>

Nawrot, R., Barylski, J., Nowicki, G., Broniarczyk, J., Buchwald, W., & Goździcka-Józefiak, A. (2013). Plant antimicrobial peptides. *Folia Microbiologica*, 59(3), 181–196. <https://doi.org/10.1007/s12223-013-0280-4>

Ishaq, N., Bilal, M., & Iqbal, H. M. (2019). Medicinal Potentialities of Plant Defensins: A Review with Applied Perspectives. *Medicines*, 6(1), 29. <https://doi.org/10.3390/medicines6010029>

Shao, Q., Liu, X., Su, T., Ma, C., & Wang, P. (2019). New insights into the role of seed oil body proteins in metabolism and plant development. *Frontiers in Plant Science*, 10. <https://doi.org/10.3389/fpls.2019.01568>

Rahman, M., Browne, J. J., Van Crugten, J., Hasan, M. F., Liu, L., & Barkla, B. J. (2020). In silico, molecular docking and in vitro antimicrobial activity of the major rapeseed seed storage proteins. *Frontiers in Pharmacology*, 11. <https://doi.org/10.3389/fphar.2020.01340>

Minkiewicz, P., Iwaniak, A., & Darewicz, M. (2019). BIOPEP-UWM Database of Bioactive Peptides: Current opportunities. *International Journal of Molecular Sciences*, 20(23), 5978. <https://doi.org/10.3390/ijms20235978>

Huan, Y., Kong, Q., Mou, H., & Yi, H. (2020). Antimicrobial peptides: Classification, design, application and research progress in multiple fields. *Frontiers in Microbiology*, 11. <https://doi.org/10.3389/fmicb.2020.582779>

Zasloff, M. (2007). Antimicrobial peptides, innate immunity, and the normally sterile urinary tract. *Journal of the American Society of Nephrology*, 18(11), 2810–2816. <https://doi.org/10.1681/asn.2007050611>

Ma, X., Wang, Q., Ren, K., Xu, T., Zhang, Z., Xu, M., Rao, Z., & Zhang, X. (2024). A review of antimicrobial peptides: structure, mechanism of action, and molecular optimization Strategies. *Fermentation*, 10(11), 540. <https://doi.org/10.3390/fermentation10110540>

Kumar, P., Kizhakkedathu, J., & Straus, S. (2018). Antimicrobial peptides: diversity, mechanism of action and strategies to improve the activity and biocompatibility in vivo. *Biomolecules*, 8(1), 4. <https://doi.org/10.3390/biom8010004>

Kang, S., Kim, D., Mishig-Ochir, T., & Lee, B. (2012). Antimicrobial peptides: Their physicochemical properties and therapeutic application. *Archives of Pharmacal Research*, 35(3), 409–413. <https://doi.org/10.1007/s12272-012-0302-9>

Su, Z., Yu, H., Lv, T., Chen, Q., Luo, H., & Zhang, H. (2025). Progress in the classification, optimization, activity, and application of antimicrobial peptides. *Frontiers in Microbiology*, 16. <https://doi.org/10.3389/fmicb.2025.1582863>

Jenssen, H., Hamill, P., & Hancock, R. E. W. (2006). Peptide antimicrobial agents. *Clinical Microbiology Reviews*, 19(3), 491–511. <https://doi.org/10.1128/cmr.00056-05>

Huang, H. W. (2006). Molecular mechanism of antimicrobial peptides: The origin of cooperativity. *Biochimica Et Biophysica Acta (BBA) - Biomembranes*, 1758(9), 1292–1302. <https://doi.org/10.1016/j.bbamem.2006.02.001>

Huang, Y., Huang, J., & Chen, Y. (2010). Alpha-helical cationic antimicrobial peptides: relationships of structure and function. *Protein & Cell*, 1(2), 143–152. <https://doi.org/10.1007/s13238-010-0004-3>

Benincasa, M., Runti, G., Mardirossian, M., Gennaro, R., & Scocchi, M. (2016). Methods for elucidating the mechanism of action of Proline-Rich and other non-lytic antimicrobial peptides. *Methods in Molecular Biology*, 283–295. https://doi.org/10.1007/978-1-4939-6737-7_20

Sharma, N., Patiyal, S., Dhall, A., Pande, A., Arora, C., & Raghava, G. P. S. (2020). AlgPred 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes. *Briefings in Bioinformatics*, 22(4). <https://doi.org/10.1093/bib/bbaa294>

Brown, K. L., & Hancock, R. E. (2005). Cationic host defense (antimicrobial) peptides. *Current Opinion in Immunology*, 18(1), 24–30. <https://doi.org/10.1016/j.coi.2005.11.004>

Bhardwaj, A. R., Joshi, G., Kukreja, B., Malik, V., Arora, P., Pandey, R., Shukla, R. N., Bankar, K. G., Katiyar-Agarwal, S., Goel, S., Jagannath, A., Kumar, A., & Agarwal, M. (2015). Global insights into high temperature and drought stress regulated genes by RNA-Seq in economically important oilseed crop *Brassica juncea*. *BMC Plant Biology*, 15(1). <https://doi.org/10.1186/s12870-014-0405-1>

Ayadi, J., Debouba, M., Rahmani, R., & Bouajila, J. (2022). Brassica Genus Seeds: A review on phytochemical screening and pharmacological properties. *Molecules*, 27(18), 6008. <https://doi.org/10.3390/molecules27186008>

Kumar, V., Thakur, A. K., Barothia, N. D., & Chatterjee, S. S. (2011). Therapeutic potentials of *Brassica juncea*: an overview. *TANG [HUMANITAS MEDICINE]*, 1(1), 2.1-2.16. <https://doi.org/10.5667/tang.2011.0005>

Kang, N. J., Jin, H., Lee, S., Kim, H. J., Koh, H., & Lee, D. (2019). New approaches towards the discovery and evaluation of bioactive peptides from natural resources. *Critical Reviews in Environmental Science and Technology*, 50(1), 72–103. <https://doi.org/10.1080/10643389.2019.1619376>

Wang, Q., Liu, J., Jiang, M., & Yang, B. (2023). Genome-wide identification, comprehensive characterization of transcription factors, cis-regulatory elements, protein homology, and protein interaction network of the DREB gene family in *Solanum lycopersicum*. *Frontiers in Plant Science*, 13, Article 1031679. <https://doi.org/10.3389/fpls.2022.1031679>

Akbudak, M. A., Filiz, E., & Çetin, D. (2021). Genome-wide identification and analyses of tomato (*Solanum lycopersicum* L.) high-affinity nitrate transporter 2 (NRT2) family genes and their responses to drought and salinity. *bioRxiv (Cold Spring Harbor Laboratory)*. <https://doi.org/10.1101/2021.12.08.471831>

Raturi, A., Shekhar, S., Jha, R. K., Chauhan, D., Pandey, S., Kumari, S., & Singh, A. (2024). Genome-wide comparative analysis of photosynthetic enzymatic genes provides novel insights into foxtail millet and other cereals. *Frontiers in Genetics*, 15. <https://doi.org/10.3389/fgene.2024.1449113>

Zia, K., Sadaqat, M., Ding, B., Fatima, K., Albekairi, N. A., Alshammari, A., & Qamar, M. T. U. (2024). Comparative genomics and bioinformatics approaches revealed the role of CC-NBS-LRR genes under multiple stresses in passion fruit. *Frontiers in Genetics*, 15. <https://doi.org/10.3389/fgene.2024.1358134>

The Brassica juncea Genome. (n.d.). SpringerLink. <https://link.springer.com/book/10.1007/978-3-030-91507-0>

Shahbaz, M., Azeem, F., Rafique, M. U., Siraj, H. M. S., & Rizwan, M. (2023). Heat-Induced Transcriptome and Genome-Wide Analysis of NHL Genes in Maize (*Zea mays* L.) Suggest a Role of ZmNHLs Under Heat Stress. *Journal of Plant Growth Regulation*, 42(11), 6891–6902. <https://doi.org/10.1007/s00344-023-10982-5>

Rahman, M., Guo, Q., Baten, A., Mauleon, R., Khatun, A., Liu, L., & Barkla, B. J. (2021). Shotgun proteomics of *Brassica rapa* seed proteins identifies vicilin as a major seed storage protein in the mature seed. *PLoS ONE*, 16(7), e0253384. <https://doi.org/10.1371/journal.pone.0253384>

Munir, A., Iqbal, S., Khaliq, B., Saeed, Q., Hussain, S., Shah, K. H., Ahmad, F., Mehmood, S., Ali, Z., Munawar, A., Saeed, M. Q., Mehmood, S., Ashraf, M. Y., & Akrem, A. (2025). In Silico Studies and Functional Characterization of a Napin Protein from Seeds of *Brassica juncea*. *International Journal of Agriculture and Biology*, 22(6). <https://doi.org/10.17957/ijab/15.1247>

Padovan, L., Scocchi, M., & Tossi, A. (2010). Structural aspects of plant antimicrobial peptides. *Current Protein and Peptide Science*, 11(3), 210–219. <https://doi.org/10.2174/138920310791112093>

Wu, Q., Patočka, J., & Kuča, K. (2018). Insect Antimicrobial Peptides, a mini review. *Toxins*, 10(11), 461. <https://doi.org/10.3390/toxins10110461>

Ngai, P., & Ng, T. (2004). A napin-like polypeptide from dwarf Chinese white cabbage seeds with translation-inhibitory, trypsin-inhibitory, and antibacterial activities. *Peptides*, 25(2), 171–176. <https://doi.org/10.1016/j.peptides.2003.12.012>

Li, K., Ma, B., Shen, J., Zhao, S., Ma, X., Wang, Z., Fan, Y., Tang, Q., & Wei, D. (2021). The evolution of the expansin gene family in *Brassica* species. *Plant Physiology and Biochemistry*, 167, 630–638. <https://doi.org/10.1016/j.plaphy.2021.08.033>

Hammami, R., Hamida, J. B., Vergoten, G., & Fliss, I. (2008). PhytAMP: a database dedicated to antimicrobial plant peptides. *Nucleic Acids Research*, 37(Database), D963–D968. <https://doi.org/10.1093/nar/gkn655>

Yang, H., Bayer, P. E., Tirnaz, S., Edwards, D., & Batley, J. (2020). Genome-Wide Identification and Evolution of Receptor-Like Kinases (RLKs) and Receptor like Proteins (RLPs) in *Brassica juncea*. *Biology*, 10(1), 17. <https://doi.org/10.3390/biology10010017>

Basnet, R. K., Del Carpio, D. P., Xiao, D., Bucher, J., Jin, M., Boyle, K., Fobert, P., Visser, R. G., Maliepaard, C., & Bonnema, G. (2015). A Systems Genetics Approach Identifies Gene Regulatory Networks Associated with Fatty Acid Composition in *Brassica rapa* Seed. *PLANT PHYSIOLOGY*, 170(1), 568–585. <https://doi.org/10.1104/pp.15.00853>

Wei, J., Xu, L., Shi, Y., Cheng, T., Tan, W., Zhao, Y., Li, C., Yang, X., Ouyang, L., Wei, M., Wang, J., & Lu, G. (2023). Transcriptome profile analysis of Indian mustard (*Brassica juncea* L.) during seed germination reveals the drought stress-induced genes associated with energy, hormone, and phenylpropanoid pathways. *Plant Physiology and Biochemistry*, 200, 107750. <https://doi.org/10.1016/j.plaphy.2023.107750>

Jaiswal, M., Singh, A., & Kumar, S. (2022). PTPAMP: prediction tool for plant-derived antimicrobial peptides. *Amino Acids*, 55(1), 1–17. <https://doi.org/10.1007/s00726-022-03190-0>

Chandrashekar, S., Vijayakumar, R., Chelliah, R., Daliri, E. B., Madar, I. H., Sultan, G., Rubab, M., Elahi, F., Yeon, S., & Oh, D. (2021). In Vitro and In Silico Screening and Characterization of Antimicrobial Napin Bioactive Protein in *Brassica juncea* and *Moringa oleifera*. *Molecules*, 26(7), 2080. <https://doi.org/10.3390/molecules26072080>

Montesinos, E. (2007). Antimicrobial peptides and plant disease control. *FEMS Microbiology Letters*, 270(1), 1–11. <https://doi.org/10.1111/j.1574-6968.2007.00683.x>



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis **GENOME-WIDE IDENTIFICATION AND *IN SILICO* CHARACTERIZATION OF PUTATIVE ANTIMICROBIAL PEPTIDES FROM SEED STORAGE AND DEFENSE PROTEINS OF *BRASSICA JUNCEA***

Total Pages: _____

Name of the Scholar: Arpit Raturi

Supervisor

1) Dr. Navneeta Bharadvaja

2) _____

3) _____

Department: Department of Biotechnology, Delhi Technological University, Delhi-110042

This is to report that the below thesis was scanned for similarity detection. The process and outcome are given below:

Software used: Turnitin Similarity Index: _____, Total word Count: _____

Date: _____

Candidate's Signature

Signature of Supervisor