# Computational identification of protein-protein interactions Between Rice and Xanthomonas Oryza plant proteomes

A dissertation submitted in the partial fulfilment for the award of Degree of Master of Science in Genomic Science

## MASTER OF SCIENCE IN GENOMIC SCIENCE

By

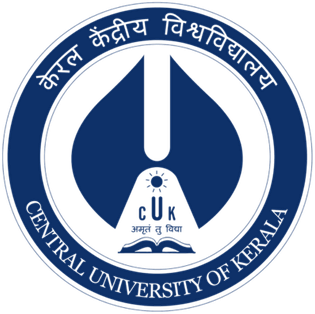
**UTTKARSH VERMA**

**Reg. No. BGS052128**

Under the guidance of

**Dr. Alagu Manickavelu**

professor

Department of Genomic Science

Department of Genomic Science School of Biological Science Central University of Kerala Kasaragod

August 2023

# CERTIFICATE

This is to certify that the dissertation entitled “ **Computational identification of protein-protein interactions Between Rice and Xanthomonas Oryza plant proteomes”** submitted in partial fulfilment of the requirement for the award of the degree of **Master of Science in Genomic Science**, to Central University of Kerala, done by Ms. **UTTKARSH VERMA, (Reg. No. BGS052128)** is an authentic work carried out by her at **Central University of Kerala, Periye, Kasaragod, Kerala,** under the guidance and supervision of **DR. Alagu Manickavelu** sir during the period of **May 2023-August 2023**. It is further certified that this work or part of this work has not been submitted elsewhere for any other degree.

**Signature of the supervisor Signature of H.O.D**

**Signature of the examiner**

# DECLARATION

I UTTKARSH VERMA, hereby declare that the dissertation entitled “**Computational identification of protein-protein interactions Between Rice and Xanthomonas Oryza plant proteomes**” a Master’s degree project under the guidance of Dr. Alagu Manickavelu, Professor Department of Genomic Science, Central University of Kerala. This is an original record of my Master’s Degree project work which is submitted for the partial fulfilment of the requirement of Master’s Degree.

Place: Central University of Kerala

Date: (Signature)

**UTTKARSH VERMA**

Reg No. BGS052001

MSc Genomic Science

Central University of Kerala

# ACKNOWLEDGEMENT

I express my deepest gratitude to our guide Alagu Manickavelu, Professor Department of Genomic Science for his steadfast guidance, supervision and the support extended by him during my entire work. I heartily thank Miss. Bhavana K and Dr.joshana mam, for their attentive support and earnest help throughout. I owe my sincere gratitude to Dr. Sameer sir (Head, Dept of Genomic Science), Dr. A Manickavelu (Professor, Dept. of Genomic Science) Dr. Ranjith Kumavath (Assistant Professor, Dept. of Genomic Science) Dr. Tony Grace (Assistant Professor, Dept. of Genomic Science), for their warm affection and persistent support throughout the course. I feel privileged to be a part of this workgroup and I am thankful to all the lab members for their support and cooperation. I am thankful to all my friends for their constant support and encouragement. With immense pleasure, I would like to express my deep sense of gratitude to my family for their love, constant support, and encouragement throughout my studies.

(Signature)

**Uttkarsh Verma**

# INDEX

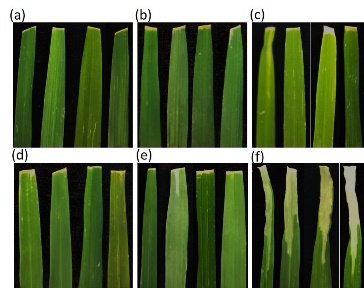
|  |  |  |
| --- | --- | --- |
| **S. No** | **Content** | **Page no.** |
| **1** | **Abstract** |  |
| **2** | **Introduction** |  |
| **3** | **Objectives** |  |
| **4** | **Materials and methods** |  |
| **5** | **Results** |  |
| **6** |  |  |
| **7** |  |  |
| **8** | **Reference** |  |

**Abstrat**

Rice is the mostly staple food in the india basiclly North and north west parts of India(Punjab, Haryana & Delhi), Gujrat, west Uttar Pradesh, south Karnataka, Tamil Nadu, parts of Telangana , leeward side of Maharashtra, parts of Madhya Pradesh and Chhattisgarh, are the areas of low rainfall in India.Desert areas of Rajasthan, Kutch in Gujrat and Leh & Ladakh regions of Jammu & Kashmir are the driest regions of India. These have had no rains at all in certain years. As Xanthomonas Oryza pv. oryzae (Xoo) is the bacteria that basically cause disease in rice (Oryza) that is basically host specific also it plays a role in for fundamental understanding of pathogen biology such as the rice production worldwide. Due to the economic importance, extensive genetic and genomic studies have been conducted to elucidate the molecular mechanism of rice response to Xoo and Xoc in the last two decades. As recent study uses to find the interaction between rice and Xoo and R gene or their product and effectors. These diseases lead to 16% of global crop yield losses, Plants have evolved sophisticated innate ability of each cell to fend off the attack, as we all know every disease is caused by some microorganism whether in plant or in animals as the infection of Rice (oryza) by Xoo is causing bacterial blight (BB)

Disease all the disease is cause by some protein that play a role in sepression of growth of rice as well as the protein also increase toxiticity in the leaf of rice as by this its really challanging part to understand molecular level the infection of leaf is done by effector protein which have the spefic binding domine of receptor protein of oryza, as its very challinging part to undertand the working and finding the exect molecular intraction is now a days done by ppi(protein protein intraction) or by Ai ML(machine learning ) tools as this part of biology basically deals with computatinal part of biology with great understanding of both molecular part as well as compunational part of science, for this process we have chossen the AI(artificial intillengence) ML(machine lerning) .

Bacterial blight disease incidence across the state of India (red circles indicate the disease incidence in various locations across the Indian states).



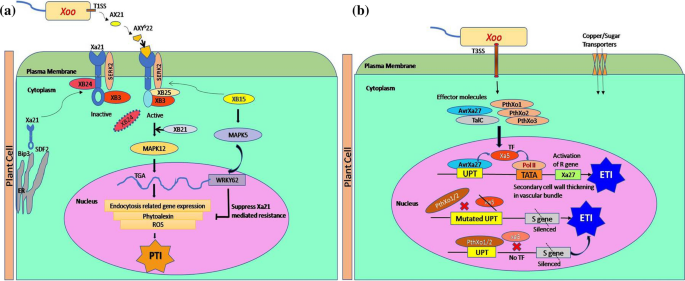
These abstract reviews the state-of-the-art bioinformatics tools designed for predicting and characterizing protein-protein interactions. We discuss different computational approaches employed by these tools, such as machine learning, network analysis, and structural modelling.

Furthermore, we explore the challenges and limitations associated with PPI prediction and highlight the potential future developments in this field. We highlight the importance of integrating multiple tools and data sources to enhance the accuracy and reliability of PPI predictions. By harnessing the power of bioinformatics tools, researchers can decipher the complex landscape of PPIs, uncover novel interaction networks, and gain insights into cellular processes and disease mechanisms.

**Introduction**

Protein-protein interactions (PPIs) are fundamental to the functioning of biological systems, regulating various cellular processes and mediating complex molecular mechanisms. Understanding the landscape of PPIs is crucial for unravelling the intricate relationships between proteins and their roles in biological pathways. Traditional experimental methods for studying PPIs have limitations in terms of coverage and scalability. However, recent advances in bioinformatics tools and computational approaches have revolutionized our ability to predict, analyze, and understand PPIs on a large scale.

In 2020, a groundbreaking development in protein structure prediction and PPIs occurred with the introduction of AlphaFold2. AlphaFold2 is an advanced deep learning system that uses artificial intelligence (AI) and deep neural networks to predict protein structures accurately and infer potential protein interactions. The significance of AlphaFold2 lies in its ability to provide accurate and reliable predictions of protein structures, which are crucial for understanding their functions and interactions. Several computational methods have been developed to predict PPIs based on different principles and algorithms. Machine learning algorithms, such as support vector machines (SVMs) and random forests, have been applied to predict PPIs from protein sequence data (Jones et al., 2012; Pazos & Valencia, 2002). Network-based approaches, including graph theory and network propagation methods, analyze the topology of PPI networks to identify functional modules and uncover hidden interactions (Bader & Hogue, 2003; Vella et al., 2017). Structural modelling techniques, such as protein docking and molecular dynamics simulations, provide insights into protein complexes' three-dimensional structure and binding interfaces (Yan et al., 2020; Schneidman-Duhovny et al., 2005). The original Alpha Fold was introduced by Senior et al. in 2019, demonstrating impressive progress in protein structure prediction. However, AlphaFold2, presented by Jumper et al. in 2020, marked a significant advancement with its remarkable performance in the Critical Assessment of Structure Prediction (CASP) competition. AlphaFold2 outperformed other computational methods and even rivalled experimental techniques in accurately predicting protein structures. Its success has since attracted widespread attention and sparked enthusiasm within the scientific community.



Rice immune response to *Xanthomonas oryzae* pv. *oryzae* infection. (**a**) Pathogen Triggered Immunity: RLK protein XA21 synthesized and processed at endoplasmic reticulum was transported to plasma membrane and association of an ATPase, XA21 binding protein 24 (XB24) to the juxta membrane domain auto phosphorylates XA21 to remain inactive.

The accuracy and reliability of protein structure prediction achieved by AlphaFold2 have profound implications for the study of PPIs. Accurate predictions of protein structures enable researchers to infer potential interactions between proteins and provide insights into their binding interfaces. This information is crucial for understanding the molecular mechanisms underlying biological processes, drug discovery, and the design of therapeutics targeting specific PPIs.The application of AlphaFold2 in the field of PPIs has opened up new avenues for research. By combining AlphaFold2 predictions with existing experimental data and bioinformatics tools, researchers can gain deeper insights into the networks and dynamics of PPIs. This integrated approach enhances our understanding of complex biological systems and facilitates the discovery of novel interactions and functional modules. In this research paper, we aim to explore the impact of AlphaFold2 in the field of PPIs and its implications for protein structure prediction. We will review the development of AlphaFold2, highlighting its performance in protein structure prediction and its potential for predicting PPIs. We will also discuss integrating AlphaFold2 predictions with other bioinformatics tools and experimental approaches to enhance the accuracy and reliability of PPI predictions.

Furthermore, we will examine recent studies and applications that have utilized AlphaFold2 in understanding specific biological processes, elucidating disease mechanisms, and identifying potential therapeutic targets. By showcasing the diverse applications of AlphaFold2, we aim to

demonstrate its potential to transform the field of PPI research and drive future advancements. Through this research paper, we aim to provide a comprehensive overview of the impact and potential of AlphaFold2 in the study of PPIs. By harnessing the power of AI and deep learning, AlphaFold2 has the potential to revolutionize our understanding of protein structures, interactions, and their role in complex biological systems.

**OBJECTIVE**

MATERIALS AND METHODS

SAMPLE COLLECTION

The data was taken by NCBI website on the bases of various articles majorly which are having role in defense and metabolic pathway with proper literature and function of each protein was known: the data was taken by based on three pathways related directly or indirectly in both the organism Xoo and RICE.

TABLE 1. LIST OF BLOOD AND TISSUE SAMPLE COLLECTED.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Effector protein** | | | |  | **Receptor protein** | | |
|  | **PROTEIN NAME** | **NCBI Reference Sequence:** | **ORGANISM** |  | **Protein name** | **NCBI Reference Sequence:** | **Organism** |
| 1 | YopJ family type III secretion system effector XopJ [Xanthomonas] | WP\_109292224.1 | xanthomonas oryzae pv. oryzae | 1 | receptor like kinase, partial [Oryza sativa Indica Group] | ABR25930.1 | Oryza sativa Indica Group |
| 2 | xopx |  | xanthomonas oryzae pv. oryzae | 2 | ethylene receptor 3 isoform 2 precursor [Oryza sativa Japonica Group] | NP\_001388942.1 | Oryza sativa Indica Group |
| 3 | xopq |  | xanthomonas oryzae pv. oryzae | 3 | wall-associated receptor kinase-like 3 [Oryza sativa Japonica Group] | XP\_025880762.1 | Oryza sativa Indica Group |
| 4 | xopl |  | xanthomonas oryzae pv. oryzae | 4 | putative receptor protein kinase ZmPK1 [Oryza sativa Japonica Group] | XP\_025877832.1 | Oryza sativa Indica Group |
| 5 | avrbs1 |  | xanthomonas oryzae pv. oryzae | 5 | TPA\_inf: WRKY transcription factor 30 [Oryza sativa Japonica Group] | DAA05095.1 | Oryza sativa Indica Group |
| 6 | ABC transporter ATP-binding protein [Xanthomonas oryzae] | UNE64853.1 | xanthomonas oryzae pv. oryzae | 6 | RecName: Full=Calcium-binding protein CBP; Short=OsCBP | Q2QY10.1 | Oryza sativa Indica Group |
| 7 | nucleoside hydrolase [Xanthomonas oryzae] | UNE64851.1 | xanthomonas oryzae pv. oryzae | 7 | subtilisin-like protease Pr1B, partial [Metarhizium anisopliae] | AAC49831.1 | Oryza sativa Indica Group |
| 8 | IS3 family transposase [Xanthomonas oryzae] | UNE64850.1 | xanthomonas oryzae pv. oryzae | 8 | serine/threonine-protein kinase SAPK9 [Oryza sativa Japonica Group] | NP\_001391667.1 | Oryza sativa Indica Group |
| 9 | MFS transporter [Xanthomonas oryzae] | UNE64796.1 | xanthomonas oryzae pv. oryzae | 9 | type-1 proteins geranylgeranyltransferase subunit beta [Pyricularia oryzae Y34 | ELQ34401.1 | Oryza sativa Indica Group |
| 10 | AdeC/AdeK/OprM family multidrug efflux complex outer membrane factor [Xanthomonas oryzae] | UNE64758.1 | xanthomonas oryzae pv. oryzae | 10 | bidirectional sugar transporter SWEET2b | NP\_001395931.1 | Oryza sativa Indica Group |
| **DATA WAS COLLECTED BY NCBI:** [National Center for Biotechnology Information (nih.gov)](https://www.ncbi.nlm.nih.gov/) | | | | | | | |

**PROTEIN- PROTEIN PREDICTION/ DOCKING**

The protein protein intraction is the tecniques use for the

Pymol: splicing the binding region of protein

Qdock: find the rate of intraction

Matlab: to find the stability of protein

GROMAX: MD simulations of protein in different behavior

Result:

Reference