J. Agric. Food. Tech., 7(9)10-21, 2017

ISSN 2090 – 424X
Journal of Agriculture and Food Technology
www.textroad.com

Identification and Validation of Putative *Erwinia mallotivora* Effectors proteins via Quantitative Proteomics and Real Time Analysis

[1]Norliza Abu Bakar, [1]Rafidah Barun, [1]Lina Rozano, [1]Laziana Ahmad, [1]Rohaiza Ahmad Redzuan, [2]Mohd Firdaus Mohd Raih, [1]Amin Asyraf Tarmizi

[1]Biotechnology & Nanotechnology Research Centre, MARDI, G.P.O. Box 12301, 50774 Kuala Lumpur, Malaysia.
[2]School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

ABSTRACT

Papaya dieback disease symptom can be depicted through the manifestation of greasy water-soaked lesions and spots on leaves and crowns as well as defoliation and blemishing of papaya tress and fruits. *Erwinia mallotivora* which was phenotypically, biochemically and genotypically shown as the close relative of *E. papayae*, is accountable for the devastating papaya dieback disease in Malaysia. Virulent factors are known to be secreted out from plant pathogen for them to take effect. Only a small amount of information is available on *E. mallotivora* mechanism of virulence. In an attempt to understand the molecular mechanisms leading to the bacterial pathogenesis; protein profiling of *E. mallotivora* was studied using iTRAQ mass spectrometry analysis. *Erwinia mallotivora* was grown in selected nutrient rich and minimal media that simulate the environment of papaya leaf apoplast. Identification of hypersensitive response and pathogenicity (hrp) and virulence proteins from papaya dieback pathogen, *E. mallotivora,* were conducted through increased fold changes in gene expression of the pathogen in media that stimulate the expression of virulent related proteins after LC MS-MS, bioinformatics and database search analysis. Analysis of the results showed that the expression of proteins grown in the minimal media were different when compared to the proteins expressed in rich media. Putative pathogenicity and virulence related proteins were identified. They include chorismate mutase, lipoprotein, achromobactin, Type III Secretion System (T3SS), proteases and hydrolase proteins. Validation of selected T3SS proteins and chorismate mutase through quantitative real-time polymerase chain raction (qPCR) showed a correlation with the results obtained. The identified proteins may represent important proteins that contribute to the bacterial pathogenicity and virulence. The approach used in this study may be useful for further analyses in order to understand the molecular mechanism underlying papaya plant defense against *E. mallotivora* with the hope to develop novel strategies to control this important plant pathogen. This may provide insights into the bacterial pathogenesis and pave the solutions for disease management.

KEYWORDS: *Erwinia mallotivora*, iTRAQ, virulent factors and hypersensitive response proteins

1. INTRODUCTION

Papaya (*Carica papaya*), belonging to the Caricaceae family, is a highly popular fruit tree especially in the tropical and subtropical areas of the world. The fruit is well known not only for its nutritious quality but also for its medicinal properties. Ranking first in nutritional scores when compared to 38 common fruits, papaya fruits are highly recommended due to the presence of fiber, potassium, folate, niacin, thiamine, riboflavin, iron, calcium and vitamin A and vitamin C. The leaves, roots, fruit and stems of papaya are also a source of an important medicinal substance known as papain; a valuable proteolytic enzyme [1,2]. *Carica papaya* has been widely cultivated in all parts of the world including South-east Asia region resulting in increased global production due to popular demands from consumers. Prior to 2005, Malaysia was once rated as the second most significant world exporter of papaya fruits with a total volume of 58,149 mt accounting for 21% of the global trade [3]. In Malaysia, papaya fruit is of considerable economic importance having an export worth of about RM100–120 million annually. However, the threat of the dieback disease are proven to be eminent and has since become the major causal agent for rapid decline in popular local cultivars for fruit export like 'Eksotika', 'Solo' and 'Sekaki' [4,5]. The disease, known as papaya dieback, is caused by a specific Gram-negative bacteria *Erwinia mallativora* [6].

Erwinia mallotivora is a part of the family of Enterobacteriaceae and is a Gram-negative bacterium which was first isolated in 1976 from a bacteria leaf spot disease in *Mallotus japonicus* plant [7]. Similar to other plant bacterial

*** Corresponding Author:** Norliza Abu Bakar, Biotechnology & Nanotechnology Research Centre, MARDI. G.P.O. Box 12301, 50774 Kuala Lumpur, Malaysia. Tel: 603-89437664 Email: lizaab@mardi.gov.my

pathogens, *E. mallotivora* enters its papaya host through stomata and wound openings. Subsequently, the entire parts of the papaya plant including shoot, leaf, frond, and bar and also the fruit will be infected that leads to early slimy, greasy, water-soaked patches as well as spot signs and symptoms on the foliage and petioles. Eventually, these result in necrosis, blemished or premature fruit drop and followed by death of the papaya plant [6,8].

Proteomics technology encompasses a broad range of technologies aimed at systematic analysis of proteins for protein identification, structure, functions, molecular characterization and identification of interacting partners [9]. Quantitative MS based proteomics is an emerging technique in the field of plant proteomics. When combined with genomics, transcriptomics and metabolomics platforms, proteomics technology has the ability to disclose the actual functionality of genes in a collective way [10]. Isobaric tags for relative and absolute quantitation (iTRAQ) and label-free methods are broadly useful for quantitative MS based proteomics. The iTRAQ method depends upon the differentiation of primary amino groups in intact proteins using isobaric tag for relative and absolute quantitation (iTRAQ) [11]. Development of the iTRAQ reagents and technology has the means to provide quantitative information of four or eight samples for the discovery and elucidation of protein markers and differential protein expression analyses [12]. The isobaric mass design of the iTRAQ reagents enables differentially labeled proteins to show up as single peaks throughout the MS/MS analysis. The iTRAQ reporter ions are released during collision-induced dissociation (CID) and quantifications are performed depending on the peak area of the reporter ions [13,14]. The technology has proven to be effective for quantification of proteins belonging to organisms which included *Trichoderma reesei* [15], *Phanerochaete chrysosporium* [16], *Aspergillus fumigatus* [17], and *Arthrobotrys oligospora* [18]. Pathogenic bacteria exude proteins that are necessary for supply associated with nutrients, cell-cell interaction, detoxing of the environment, and killing of prospective competitors during intrusion of their host plants [19]. The pathogenicity of Gram-negative plant-pathogenic bacteria relies on hypersensitive response as well as the pathogenicity (hrp) genes, which control the pathogen ability to multiply in susceptible hosts and cause disease [20]. A plethora of methods have been applied when elucidating mechanisms of bacterial pathogenicity. Protein elements of the secretion plant pathogen apparatus usually are not constitutively expressed and therefore are only expressed at minimal levels when bacteria are grown in nutrient-rich medium. Naturally, these genes expression are induced and highly expressed provided that the bacteria enter in the intercellular space (apoplast) of plant tissues [21]. To compensate the difficulty of obtaining quantitative expression of phytopathogens upon infection with their host, formulated media that mimic the plant apoplast have been widely used to increase expression of virulent proteins in an *in vitro* condition [22].

The pathogenic mechanism of *E. mallotivora* has not been fully explored yet. Little is understood about the active molecular mechanisms in the pathogenicity and host specificity of *E. mallotivora* which can be crucial in determining its pathogenicity mechanism. For better understanding of the disease, knowledge on the underlying mechanism of the papaya dieback disease is crucial especially in management and future studies of the disease. In this study, *E. mallotivora* were grown in selected nutrient rich medium and minimal medium that simulate the environment of the leaf apoplast. Identification of potential proteins associated with hypersensitive response and pathogenicity (hrp) from *E. mallotivora* was carried out via proteomics technology by comparing the protein profiles of the bacteria grown in these media. Real-time analysis (RTA) was carried out to further validate the expression of these putative virulent factors and hypersensitive response proteins.

2.MATERIALS AND METHODS

2.1Bacterial strain and culture media

Erwinia mallotivora strain was obtained from Horticulture Research Centre, Malaysian Agriculture Research and Development Institute (MARDI). Inductions of differential proteins were carried out using minimal medium represented by M9 minimal medium and rich medium represented by Luria Bertani (LB) medium in triplicates. The M9 minimal medium was consisted of 10 mM $(NH_4)_2SO_4$, 5 mM $MgSO_4$, 1 mM $CaCl_2$, 0.16 mM KH_2PO_4, 0.32 mM K_2HPO_4 and 0.01% (w/v) yeast extract. NaCl (20 mM), sucrose (10 mM) and methionine (2 ug/uL) were added prior use as recommended by Schulte and Bonas [23]. *Erwinia mallotivora* was grown at 28ºC in nutrient broth to OD600 of 1.0. The cells were pelleted down and washed twice with 1x Phosphate Buffer Saline (PBS) before being resuspended in M9 and LB media according to their intended subculturing media. The cells were grown for 24 hours until the OD reach A600 = ~1.8. The cells were pelleted by centrifugation at 4000 x g for 15 minutes for protein extraction step.

2.2 Protein Extraction and Quantification

To identify *E. mallotivora* secreted effector proteins under inducing conditions by quantitative proteomic, 50 ml of culture suspension was collected by centrifugation. The cells were lysed using Bugbuster reagent

(Novagen) and the culture supernatants released by the lysed bacteria were collected and passed through a 0.22 µm membrane filter. Precipitation of proteins were carried out using Trichloroacetic acid (TCA) at 4°C overnight before centrifugation at 16,000 *g* for 10 min at 4°C. The pellet was washed 3 times with ice cold 80% acetone, air dried at room temperature and resuspended in resolubilisation buffer (7 M urea, 2 M thiourea and 4% 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS)] for 30 minutes with vigorous vortexing and sonications. The quantity of solubilised protein was measured at 595 nm absorbance using spectrophotometer. 20ug of proteins were loaded onto 12% SDS PAGE for visualisation. The quantified samples were later freeze-dried using VirTis freeze dryer.

2.3 Quantitative Mass Spectrometry Analysis using iTRAQ reagent

For iTRAQ analysis, proteins were dissolved in iTRAQ dissolution buffer and precipitated using ITSI Bio's ToPREP kit to remove iTRAQ interfering chemicals in the sample. Quantification of proteins was carried out using Bradford protein assay techniques. 100 ug of peptides from each samples were labeled with 2 plex iTRAQ reagents according to the manufacture's instruction (Applied Biosystems Inc., Foster City, CA). Approximately 35 µg of protein was used for all samples for the iTRAQ analysis. All of the labelled proteins were reduced, alkylated and digested using trypsin as the enzyme. After digestion, the samples were labeled with iTRAQ 3-plex reagents as followed:

i. *Erwinia mallotivora* grown in rich medium;LB: iTRAQ reagent 116
ii. *Erwinia mallotivora* grown in minimal medium:M9: iTRAQ reagent 118.

Labeled peptides were combined and cleaned by strong cation exchange (SCX) chromatography. The iTRAQ-labeled peptides were eluted through the SCX column using 225 mM and 450 mM Ammonium acetate. Eluted peptides were dried and re-dissolved in 20 µl 5% acetonitrile in 0.1% formic acid. Samples were loaded onto a PicoFrit C18 nanospray column (New Objective) utilising a Thermo Scientific Surveyor Autosampler operated in the no waste injection mode. Peptides were eluted through the column utilising a linear acetonitrile gradient from 5 to 45% acetonitrile over 230 minutes. This was followed by organic washes for an additional twenty minutes. A data-dependent was utilised where a full MS scan from m/z 400-1500 as well as MS/MS scans on the three most abundant ions were carried out. Each ion was subjected to CID for peptide identification followed by PQD for iTRAQ quantitation.

Raw data files obtained were searched using Proteome Discoverer 1.3 (Thermo Scientific) and the SEQUEST algorithm. Trypsin was the selected enzyme allowing for up to two missed cleavages per peptide. Methylthio Cysteine, N-terminal iTRAQ 8-plex, and Lysine iTRAQ 8-plex were used as static modifications and oxidation of methionine as variable modification. Proteins were identified when two or more unique peptides had X-correlation scores greater than 1.5, 2.0, and 2.5 for respective charge states of +1, +2, and +3. ProteinPilot™ 2.0 software (Applied Biosystems) was used for peptide matching, protein identification, and relative protein quantitation. The relative abundance of each and every peptide was established by ProteinPilot™ utilising the peak areas of signature ions from the iTRAQ-labeled peptides. The mass spectra data obtained were searched for similarity against GenBank, EMBL, Swiss Prot databases using BLAST. The iTRAQ Ratio of Treated / Control of more than 1.5 were considered to be up regulated up regulated, whilst ratio of less than 0.67 were considered down regulated. Ratio from 1.5 - 0.67 were considered to be moderate or no changes.

2.4 Bioinformatic Analysis to Determine Virulent Factors

Sequences for known plant bacteria virulence factors were downloaded from Pfam database (http://pfam.xfam.org/). Hmmer software which can be used on the website http://hmmer.org were downloaded and used to identify virulence genes/proteins from the unregulated proteins sequences obtained from the iTRAQ analysis.

2.5 RNA isolation and Quantitative Real Time Analysis

Two starter cultures of *E. mallotivora* were prepared by cultivating a single colony of bacteria in 5 ml of LB liquid broth overnight at 28°C. Next day, the 5 ml cultures were spun down at 10,000 rpm in 5 minute and washed twice with PBS to eliminate the remaining of LB liquid broth substances. Pellets obtained were resuspended in 50 ml of M9 minimal media. For rich media, another 5 ml culture was straightly inoculated into 50 ml LB liquid broth. Both cultures in the different media were then grown overnight in incubator shaker at 200 rpm, 28°C.

Bacterial RNA extractions were performed by using RiboPure™- Bacteria kit (Ambion, USA). 1 ml culture of bacteria from each flask containing different media were collected and used to extract the RNA by following the kits protocol. Integrity and concentration of RNAs were then checked using agarose gel and NanoDrop

Spectrophotometer ND-1000 (Thermoscientific, USA). Complimentary DNA (cDNA) synthesis was then carried out by using Quantitect Reverse Transcription kit (Qiagene, USA). Starting material which has been used for the cDNA synthesis was 2 ug of *E. malotivora* RNA from both samples.

Realtime qPCR experiment was carried out to validate several virulence/effector genes which play important role in *E. mallotivora* pathogenicity. The selected genes from previous iTRAQ analysis are protease 2, hydrolase, luxR transcriptional regulator, chorismate mutase, hrpN, hrpS, hrpL and hopF2. This experiment was carried out by using StepOne Plus from Applied Biosystem, (CA, USA). In each well of 0.1 ml in microAmp Fast 96 well reaction plate (Applied Biosystem, USA), a RT-qPCR reaction in 10 ul of total volume was set up as follows: 5 ul of SensiFAST Probe Hi-Rox mix (2x) (BIOLINE, UK), 1 ul of diluted purified cDNA template, 0.5 ul of PrimeTime qPCR assay (1x concentration of 500 nm primers and 250 nm probe) (IDT, USA) and 3.5 ul of steril distilled water. The thermal cycling program at ABI StepOne Plus (Applied Biosystem, Foster City, CA) was: 95°C for 3 min, 95°C for 10 s and 60°C for 30 s in 40 cycles. 96 wells plate containing of qPCR reactions was sealed with clear microAmp Optical Adhesive film (Applied Biosystem, Foster City, CA). The plate then was briefly vortex and spun down at 3000 rpm for 2 min before put in the qPCR machine. Prime Time ® Standard qPCR assay from IDT DNA (IDT, USA) were synthesised by the corresponding service provider and used as probes in genes expression analysis. Two genes which are *gltA* and *glk* were selected as the housekeeping genes. The results were analysed by comparative Ct (ΔΔct) method using StepOne Plus software to validate and determine the change in expression of target in a sample relative to same target (GOI) in a reference sample.

3. Results and Discussion

Isobaric tags for relative and absolute quantification (iTRAQ) coupled with multidimensional liquid chromatography (LC) and tandem MS analysis are powerful tool for quantitative proteomic analysis which has been widely used in many researches. In this particular} study, the iTRAQ approach was applied to determine differential protein expression profiles of *E. mallotivora* pathogen cultured in 2 different media; under effector inducing condition and normal condition followed by selected genes validation performed by using quantitative Real-time experiments.

3.1 *Erwinia mallotivora* growth and protein extraction

Virulence factors and pathogenicity(hrp) proteins or effector proteins are important factors during invasion of any plant pathogen. *Erwinia mallotivora,* like other phytopathogenic bacteria, delivers a number of virulent factors and hrp proteins that are responsible for its pathogenesis. For the experiment, *E. multivora* strain was obtained from Horticulture Research Centre, Malaysian Agriculture Research and Development Institute (MARDI). To overcome problems associated with plant protein contamination and low level expression of effector proteins, a whole range of media to stimulate the environment of the plant tissues that can promote optimal production of effector proteins in *E. mallotivora* were exploited for this study. A method for efficient concentration and fractionation of proteins secreted by *E. mallotivora* under effector inducing conditions was employed to facilitate visualisation and subsequent identification of *E. mallotivora* effector proteins. Nutrient rich medium was represented by LB broth. Minimal media which mimic the environment of plant apoplast were formulated based on past research. The minimal media used in the experiment were as recommended by Schulte and Bonas [23]. However, based on the analysis of *E. mallotivora* genome sequence, sucrose instead of glucose was used as the carbon source. Modifications were made to the secreted protein preparation protocols to aid the removal of polysaccharides and contaminating compounds to obtain better resolution and visualisation of the extracellular proteins. Based on 1D SDS PAGE, the proteins extracted were visible in Coomassie-Blue-stained SDS PAGE gel. Overall, acceptable protein recoveries were obtained for the proteins of *E. mallotivora* grown in both media used (Figure 1). Expression of effector proteins in *in vitro* condition was successfully increased using the formulated media which allowed the discovery of novel effector proteins. Based on the SDS PAGE analysis, there was a differential expression of proteins from *E. mallotivora* that were cultured in effector inducing media compared to the nutrient rich media. Further investigation was carried out using proteomic based approach to identify the proteins that up regulated in the effectors inducing media. This is as expected as SDS PAGE does not have the resolution and ability to separate, identify or quantified differently expressed proteins. The main advantage of this approach is that there was no contamination of proteins produced by the plant as normally occurs if this study was to be carried out i*n vivo* using infected plant samples.

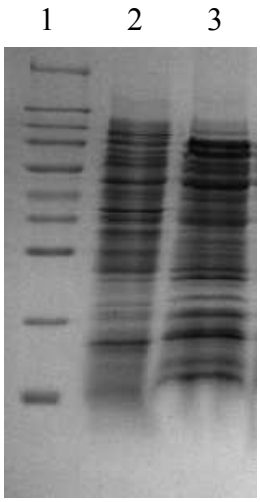


Figure 1. Approximately 20ug of protein from each sample was run on a 10% SDS PAGE before being sent for iTRAQ analysis. Significant differences were observed between the proteins of *E. mallotivora* grown in minimal medium (lane 2) with samples grown in nutrient rich medium (lane 3) based on the SDS PAGE.

3.2 The iTRAQ analysis

Proteins associated with *E. mallotivora* were obtained (Figure 1) and compared using iTRAQ proteomic analysis for the identification of potential virulent factors *in vivo*. The iTRAQ protein mass spectrometry and analysis were conducted for the intracellular protein samples. The schematic presentation of the iTRAQ mass spectrometry analysis is shown in Figure 2. For iTRAQ analysis, proteins were precipitated using ITSI Bio's ToPREP kit to remove iTRAQ interfering chemicals in the sample. Quantification of proteins was carried out using Bradford protein assay techniques. The concentrations of proteins obtained after the precipitation were as followed: *E. mallotivora* grown in plant extract media - 63 µg, *E. mallotivora* grown in rich media - 38.7 µg and *E. mallotivora* grown in minimal media - 114.2 µg. 35 µg of protein was used for all samples for the iTRAQ analysis. All of the labelled proteins were reduced, alkylated and digested using trypsin as the enzyme.

Precipitated proteins were dissolved in iTRAQ, digested using the trypsin enzyme and the control sample was labelled with iTRAQ 8 plex reagents 113 while the treated sample was labelled with iTRAQ 8 plex reagent 115. Labelleled samples were cleaned and fractionated by cation exchange chromatography and loaded into the mass spectrometer to obtain peptides peaks for identification and quantification.

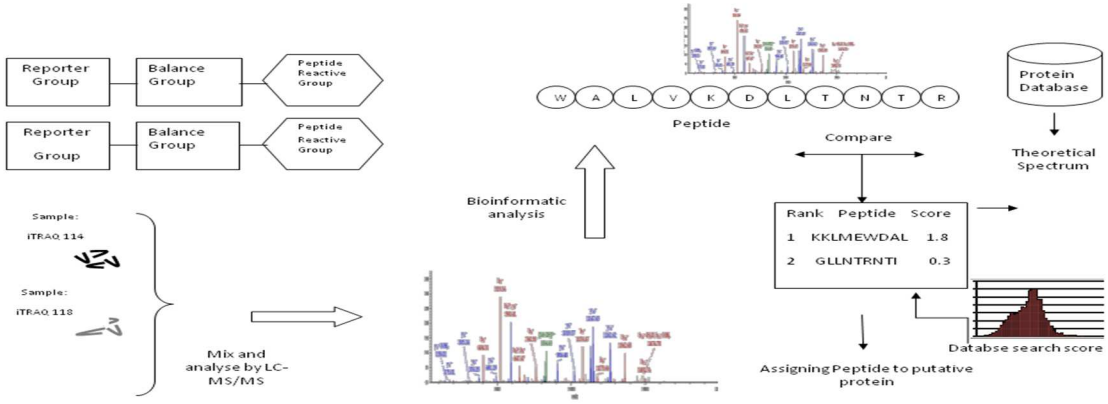


Figure 2. Schematic presentation of iTRAQ analysis for *E. mallotivora* grown in LB rich medium and *E. mallotivora* grown in M9 minimal medium.

Bioinformatics analysis was carried out for all the mass spectrometry data has been obtained for all the protein analysis. Peptide mass and all protein sequences data generated by mass spectrometry were searched against the non-redundant protein at National Centre for Biotechnology Information (NCBI) database and other relevant databases database for bacteria. Identification of proteins was based on the nearest sequenced homologue alignment results and the protein identities were only considered statistically significant if $p<0.05$. More than 380 proteins were identified out of which more than 270 proteins had quantitative iTRAQ ratios. Based on the quantification analysis carried out, treated samples with ration more than 1.5 were considered up regulated while protein with ratio less than 0.67 were considered down regulated. Gene ontology classifications were assigned to the proteins and the proteins were categorized as having functions in defence response, signal transduction and signalling, metabolism, protein destination and storage, and protein synthesis.

For virulence factor identification of quantitatively upregulated proteins, sequences for known plant bacterial virulence factors were downloaded from Pfam database (http://pfam.xfam.org/). Hmmer software (http://hmmer.org) was used to find homologous sequences in the sequential database to be used in multiple sequence alignments based on hidden Markov models. The two softwares were used to identify virulence genes/proteins from the iTRAQ up regulated proteins data. Filtering processes were carried out by referring to the e-value of each output to further eliminate duplicate output and output with high e-value. Table 1 highlight the proteins with significant differences in protein expression and were considered to be putative virulence factors proteins from *E. mallotivora* based on the proteomic and bioinformatics analysis carried out. Phytopathogen virulence factor expression has been known to be induced by culture condition that mimics the plant apoplast. Interestingly, a Chorismate mutase-P and prephenate dehydratase was found in this study to be of the highest ratio with 28.430 fold in expression. Chorismate mutases are generally crucial enzymes associated with the shikimate pathway that plays a major role in the conversion of chorismate to the precursor for tyrosine and phenylalanine synthesis [7]. Similar study by Ferreira et al. [24] demonstrated that this protein seemed to be highly expressed and it was among the potential virulence factor candidates expressed by *Xanthomonas citri.* subsp. *citri* when grown inside a medium mimicking plant growth condition. This protein is postulated to be engaged in *Erwinia mallotivora* pathogenicity through the suppression of salycilic acidity defence-related hormone in plant. Plant-connected microbes are recognized to secrete chorismate mutases which have been postulated to operate as common tools for host manipulation. This protein was proven to become involved throughout the *Xanthomonas* pathogenecity [25]. Chorismate mutase Cmu1 created by Ustilago maydis continues to be proven to become a virulence factor for that organism [26]. Nematode secreted chorismate mutase seemed to be proven to help plant development. Inside a related study, use of exogenous salicylic acidity was proven to improve papaya plant tolerance towards *E. mallotivora* infection. This really is synchronized using the report by Djamei et al. 2011 [27] that confirmed Cmu1 chorismate mutase activity which influenced *U. maydis* virulence and salicylic acidity levels through metabolic priming strategy. Much like chorismate mutase, Isochorismatase hydrolase; a protein that's involved with SA/JA defense signaling's networking was discovered to be upregulated with 5.84 fold elevated in expression in our study. Zhu et al. 2017 [28] revealed that *Verticillium dahliaes* Isochorismatase hydrolase interfered with the salicylate and jasmonate defense signaling in potato. In that study, they measured salicylic acidity (SA) and jasmonic acidity (JA) accumulation in potato plants which were inoculated with Isochorismatase hydrolase mutant virus and compared all of them with wild type *Verticillium dahliaes*.

Table 1. Putative *E. mallotivora* virulence proteins identified by liquid chromatography mass spectrometry/mass spectrometry and bioinformatics analysis.

Accesion number	Protein name	Protein score	Fold change in expression	Protein MW	Protein PI
387872378	Chorismate mutase-P and prephenate dehydratase	2.12	28.430	43.0	6.68
188534885	Lipoprotein	6.66	17.341	7.8	5.06
300715857	Hydrolase	2.57	6.135	28.5	6.62
387871475	Spheroplast protein Y	9.27	7.287	18.0	9.70
188532625	Achromobactin transport system permease	2.37	5.093	36.5	11.78
300718096	Isochorismatase hydrolase	2.63	5.084	24.5	5.81
300714654	Lipopolysaccharide core biosynthesis glycosyl transferase	2.51	4.374	29.3	9.16

387871386	Protein araA	3.34	4.275	55.2	6.71
300715701	Quorum-sensing transcriptional regulator	2.07	3.650	28.9	7.61
300714900	Plasmid transfer protein TraF	8.53	3.205	44.1	5.72
188534867	Protease 3 (Pitrilysin)	4.76	2.518	108.4	7.14
312171706	Lipoyltransferase	3.75	2.421	20.1	8.19
300716507	Outer membrane lipoprotein SlyB	4.98	2.340	15.4	9.25
478727272	Flagellar hook-associated protein 1	2.68	2.299	35.6	4.55
188534622	Multidrug transporter membrane\ATP-binding components	5.29	2.188	65.8	9.16
300716447	Osmotically-inducible lipoprotein E	8.45	1.823	12.1	7.84
312171154	Putative lipoprotein	3.06	1.823	14.7	6.79
300716486	Major outer membrane lipoprotein	9.81	1.774	8.4	8.87
334737961	Type III secretion system effector Eop1	2.53	1.753	44.4	9.79
300714933	Hydrolase	2.78	1.702	25.2	5.83
385787155	Hrp/hrc Type III secretion system-Hrp/hrc secretion/translocation pathway-hrpX protein	7.19	1.654	56.7	6.61
385785771	Cytotoxic necrotizing factor type 2	5.77	1.905	122.8	5.94
194350779	HrcV, partial	2.69	1.854	51.4	5.88
312172585	Exonuclease III	5.89	1.853	34.4	6.80
188535032	LysR family transcriptional regulator	2.69	1.701	33.6	7.14
410502733	Protein hipA	2.63	1.572	47.8	7.03
300716778	Lipoprotein	2.63	1.567	43.6	8.95
300719022	Endonuclease	5.09	1.509	36.5	8.16
188534043	Type III secretion apparatus	2.61	1.471	42.5	7.09
478701781	Invasin ipaB	3.41	2.27	57.1	6.71
292900073	Type III effector protein	5.26	2.47	76.5	7.64
478706604	Putative avirulence protein DspE (DspA)	5.17	2.42	196.3	9.14
312173037	Zinc metalloprotease zmpB precursor	4.86	2.51	25.9	8.27
312173961	Holin (Lysis protein 13)	2.82	3.34	11.5	8.51
300718169	Multidrug resistance protein	2.56	2.70	110.9	7.02
300715750	NUDIX hydrolase	2.55	3.69	22.7	4.88
188533645	Protease II	2.55	1.73	78.3	5.15
300714725	Antibiotic biosynthesis protein	2.25	2.43	30.1	4.79
300715366	Multidrug ABC transporter ATPase	2.23	3.19	34.5	8.73
478723201	Putative type VI secretion system, effector protein vgrG	2.01	1.86	8.2	10.95

Several peptides identified to be part of lipoprotein or lipoprotein machinery was shown to be highly expressed. Lipoprotprotein is a set of membrane proteins that served as a key element of the outer membrane of bacteria in Enterobacteriaceae family. Lipoproteins are connected with numerous functions varying from cellular physiology to cell division, antigenicity, colonisation, antibiotic resistance mechanism and virulency in phytopathogenic bacteria. Rigano et al. 2007 [29] demonstrated the elevated expression of a homolog protein from the genus *Xanthomonas* lipoprotein when *Xylella fastidiosa* hen the bacteria was grown in biofilm inducing medium. The bacteria are the causal agent for citrus variegated chlorosis in citrus. In another example that demonstrated lipoprotein importance, Otto et al. 2001 [30] revealed the significance of lipoproteins for the production of biofilm that aid genus *Xanthomonas* induction of virulence.

Another protein discovered to be upregulated within the inducing media and iTRAQ analysis which may be associated with virulency was achromobactin transport system permease that is an iron siderophore-like protein. In line with the carboxylate and hydroxy donor groups, achromobactin belongs to a different type of siderophore and often is recognized as another siderophore [31]. The proteins are dependent ATP-Binding Cassette (ABC) transporters active in the uptake of heme, b12, or even the divalent cations Mg2+ and Zn2+. Achromobactin is recognized as essential for microbial pathogenesis both in plant and animal hosts [32]. In *D. dadantii*, achromobactin, along with the catechol-type siderophore chrysobactin, are recognized to lead to virulence of the virus in the host plant [33]. *Pseudomonas syringae* is known to be determined by achromobactin because of its pathogenecity with the mediation of iron transport during pathogenesis [34].

A significant and essential aspect in pathogenicity of Gram-negative phytopathogenic bacteria is controlled by T3SS that functions like a conserved secretion system to export and deliver effector proteins inside the cytosol of host plant cells. Peptide sequences that correlate to T3SS were also discovered to be upregulated in *E. mallotivora* in

line with the iTRAQ analysis. The T3SS are present in many Gram-negative bacteria. This sophisticated system works for Gram-negative bacteria to secrete effectors throughout their interaction with host plants where they could manipulate and suppress the host defence and therefore causing illnesses in susceptible host plants [35]. Plant pathogenic bacteria especially in the genus *Pseudomonas, Xanthomonas, Ralstonia, Erwinia* and *Pantoea* possess T3SSs which are mainly conserved and encoded by hrp (sensitive response and pathogenicity) and hrc (sensitive response and conserved) genes [36]. Many Gram-negative phytopathogens secrete several proteins referred to as hairpin proteins that are glycine wealthy as well as heat stable. These hairpin proteins mainly act within the plant extracellular space. Our studies established that hairpin or assistant proteins (hrp) were expressed as much as 7 fold greater within the apoplast mimicking media as also shown by Deng et al. 2010 [37]. Included in T3SS, hairpin proteins are essential especially during effectors delivery as well as their characteristics assist them to communicate with plant cell walls and membranes within the plant extracellular space [38]. According to gene expression, secretion of effectors, effectors translocation and disease symptom development studies, T3SS is certainly among the targets when strategising to manage phytopathogen; as blocking the part of T3SS can result in reduced virulence in bacteria [21]. These effectors or virulence factors have roles in suppressing the host defence mechanisms to guarantee the survival within the host. The host defences that'll be targeted range from the host plant immunity system, cell dying programme, defence hormone path, and proteosome and ubiquitination systems. Interestingly, T3SS effector, Eop1 and putative avirulence protein DspE (DspA) were upregulated by 1.753 and 2.42 fold change within this study. These effector proteins are very well documented to become translocated into plant cells via T3SS. In *Arabidopsis thaliana*, DspA/E was required for *E. amylovora* growth and endurance [35]. An essential plant bacteria affecting the *Rosaceae* subfamily *Spiraeoideae* species; *E. amylovora* also needed the T3SS for the delivery of effector proteins which caused disease in susceptible host plants and eliciting of sensitive response in resistant plants [39,36].

Proteases take part in numerous roles during phytopathogen invasion, mainly through generation of proteins for that phytopathogen sustainability, playing the function to melt the cell wall as well as minimise their host intervention and defense strategy with the degradation from the plant defense proteins. This enables the virus proliferation within the host. Proteases are sorted based on the nature of amino acids of their targets and substrates of which they could be cysteine, serine, aspartic, metallo or even unclassified groups of peptidase [40]. Three different proteases were discovered to be up regulated within our study; namely Protease 3 with 2.518 fold expression, Protease 2 with 1.73 fold expression and Zinc metalloprotease with 2.510 fold expression. Several proteases from phytopathogens were already characterised and proven to become important effectors during host-virus interactions. Included in this are 48 kDa metalloprotease from *E. amylovora* [36,41] Lon, an ATP-dependent protease in *E. amylovora* [39] and also the serine and metallo proteases from *Xylella fastidiosa*, - microbial effectors for the diseases in citrus and grape [42].

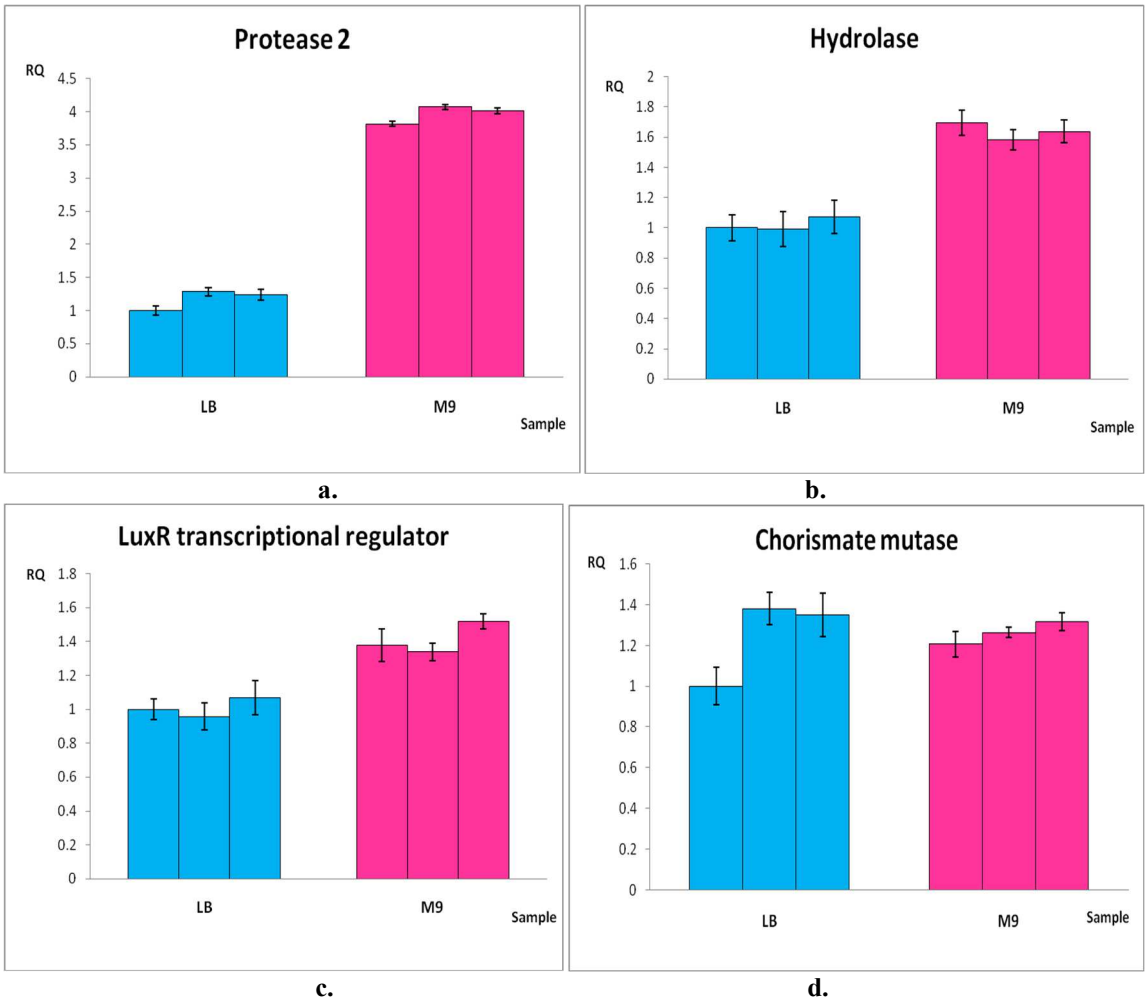
Phytopathogen pathogenesis involves several mechanisms. One of these is the opportunity to detox any substances for instance antimicrobials that may hinder bacteria proliferation via multi drug resistance (MDR) systems [43]. This ability also served as virulence factors or effectors for the plant pathogens where it helps the pathogens to initiate their development in selected host tissue. Being associated with *E mallotivora*, MDR system was proven to be required for *E. chrysanthemi* virulence in the microbial mutagenesis study. In that study, two acridine resistance (Acr)-like systems, two Emr-like systems, and one member of the major superfamily were characterised and the mutants generated generated demonstrated reduced virulence [44]. Several proteins from *E. mallotivora* that constitute the MDR system were interestingly proven to become elevated in expression when grown within the virulent factors inducing medium with fold changes of 2.70, 3.19 and a 2.188 respectively. Another phytopathogen; *Xanthomonas campestris* pv. *phaseoli* was also shown to encode a MDR gene; omlA gene. Similarly, the omlA mutant demonstrated elevated sensitivity towards the antibiotic chloramphenicol, novobiocin, coumermycin, SDS along with a protein synthesis inhibitor indicating its multifunction to improve its survival in the existence of antimicrobials compounds [45].

3.5 Validation of virulence/effector genes of *E. mallotivora*

To confirm the results of the differentially expressed proteins identified by iTRAQ LC-MS/MS analysis, real-time PCR was performed to validate the proteomic results. Recently, *Erwinia mallotivora* complete genome sequence analysis was carried out. Draft genome of *Erwinia mallotivora*, sequenced via RNA transcriptome Next Generation Sequencing (NGS) was obtained and published [46]. Subsequently, complete genome sequence was obtained by the same group using the Pac Bio System (Data not shown). These sequences were used to align with the putative effector proteins obtained through the proteomic study to obtain full length sequences of *E. mallotivora* selected genes for the quantitative real-time PCR study using probe DNA primers. Real-time PCR experiments were carried out to validate several virulence/effector genes which play important role in *E. mallotivora* pathogenicity. The selected genes are protease 2, hydrolase, luxR transcriptional regulator, chorismate mutase, hrpN, hrpL, hrpS

and hopF2. This experiment was carried out by using StepOne Plus from Applied Biosystem. Prime Time ® Std qPCR assay from IDT DNA has been used as probes in Real time experiment and 2 genes which are *gltA* and *glk* were selected as the housekeeping genes. Quantitative Real time study is very important and as it can reveal and validate the expression patterns of selected genes. The results were analysed by comparative Ct (ΔΔct) method using Step One Plus software to validate and determine the change in expression of target in a sample relative to same target gene of interest (GOI) in a reference sample.

From the qRT-PCR analysis, all selected genes expression were in line with previous proteomic data in which the putative virulent factors were found to be more expressed in minimal media as compared to rich media (Figure 3a.-h.). This did not contradict with previous iTRAQ data. Based on the gene expression patterns, it can be concluded that almost all selected potential virulence/effector genes which are protease 2, luxR transcriptional regulator, hydrolase, *hrpL, hrpN, hrpS* and *hopF2* were highly expressed in M9 minimal media compared to LB media (Figure 3a.- h.). The highest fold of gene expression in M9 media is hrpS (Figure 3f.) with average of 70 fold change compared in LB media and followed by *hopF2*, protease 2, *hrpN, hrpL*, hydrolase and *luxR* transcriptional regulator. However for chorismate mutase, the expression was a bit lower in M9 minimal media compared in LB and there was no significant difference in gene expression between the two different media (Figure 3d.). The virulence genes were postulated to have higher gene expression in M9 minimal media as the media condition supposedly to mimic the real plant apoplast where the papaya pathogen *(E. mallotivora*) secreted virulence/effector genes to start its pathogenicity process against host plant. *HrpN, hrpL, hrpS* and *hopF2* genes are involved in type III secretion system of *E. mallotivora* and may play a role as effector/virulence genes to initiate the pathogenicity against the plant host. Results obtained from this finding have been successfully validated and these supported the function and role of the studied virulence genes through the gene expression pattern in inducing media (M9 minimal media).



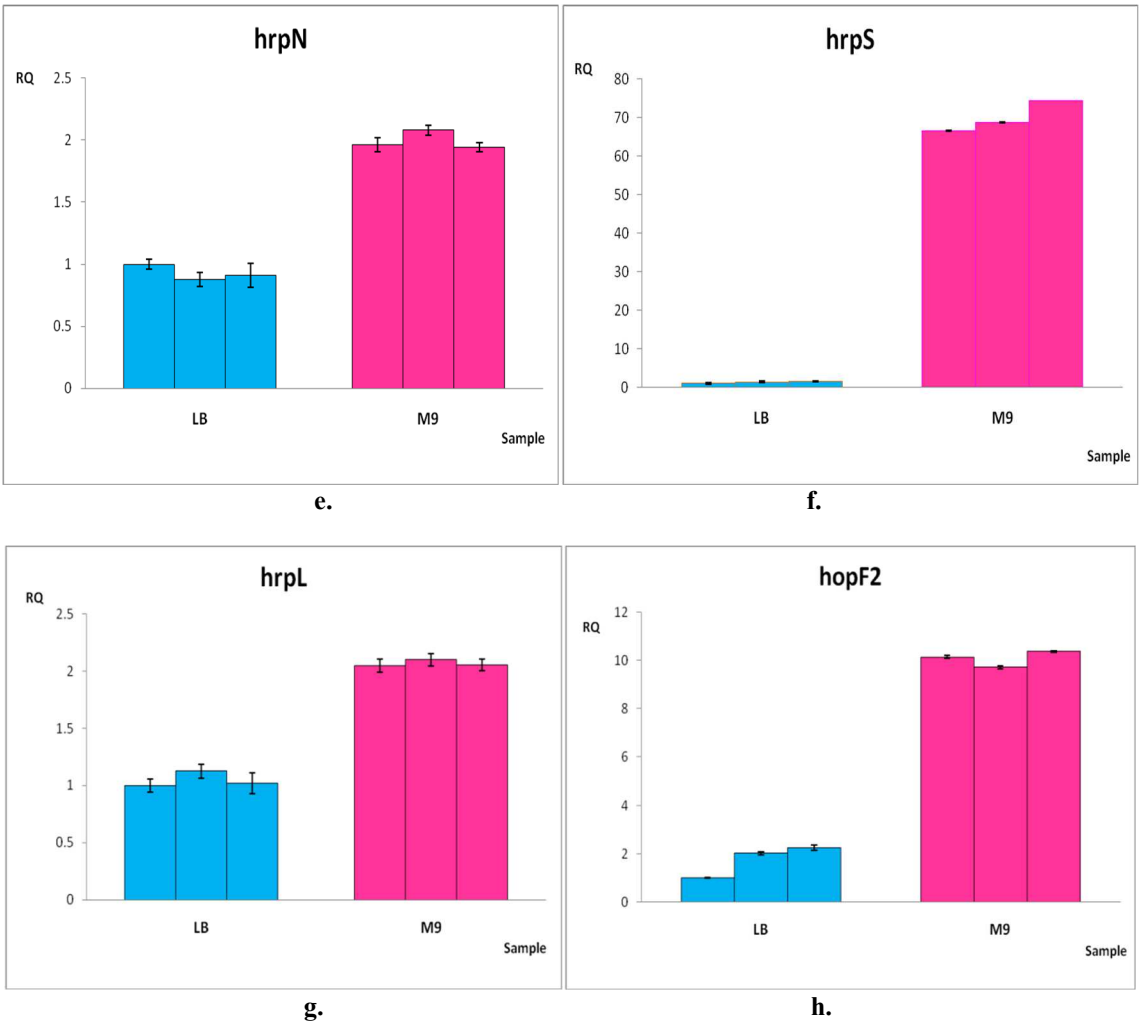


Figure 3: Gene expression pattern of selected virulence/effector genes in LB rich media compared to M9 inducing media condition; a: protease 2; b: hydrolase; c: *luxR* transcriptional regulator; d: chorismate mutase; e: *hrpN*; f: *hrpS*; g: *hrpL*; and h: *hopF2*. Overall, the validation of the qRT-PCR had resulted in a noticeable higher expression of all candidate genes in virulent factors inducing media

The information obtained from this project will facilitate the discovery of potential *hrp* genes from the pathogen and eventually provide a better understanding on the molecular mechanism of papaya dieback pathogen during its pathogenesis. This in return, will allow the discovery of potential markers and targets for identification of the pathogens for detection systems, approaches to monitor the spread of disease, as well as means to devise strategies for the control of plant disease and identification of resistant cultivars.

ACKNOWLEDGEMENT

This study was supported partly by Science Fund Grant, MOSTI Grant 02-03-08-SF0380 and FRGS Fund Grant FRGS/1/2015/ST03/MOA/02/1 from MOHE.

REFERENCES

1. Milind,P., Gurditta, 2011. Basketful Benefits of Papaya. Int Res J Pharma., 2: 6-12.
2. Macalood, J.S.,Vicente,H.J., Boniao, R.D.,Gorospe, J.G., Roa, E.C, 2013. Chemical Analysis of Carica Papaya L Crude Latex. Ameri J Plant Sci., 4: 1941-8

3. Anon, 2006. FAOSTAT. Agricultural production. Crops primary (papaya). Retrieved in 2006 from http://apps.fao.org/faostat/
4. Anon, 2009 Penyakit mati rosot betik. http://www.moa.gov.my/web/guest/ penyakit mati rosot betik.
5. Evans, E. A., and Ballen, F. H, 2012. An overview of global papaya production, trade, and consumption. Topics: Food and Resource Economics, Extension service Institute of Food and Agricultural (IFAS)., 1-7.
6. Noriha, M. A., Hamidun, B. Rohaiza, A. R. and Indu Bala S. J, 2011. *Erwinia mallotivora* sp., a new phatogen of papaya (*Carica papaya*) in Peninsular Malaysia. International Journal of Molecular Sciences 12: 39-45.
7. Kast, P., Grisostomi, C., Chen, I.A., Li, S., Krengel, U., Xue, Y and Hilvert, D, 2000. A strategically positioned cation is crucial for efficient catalysis by chorismate mutase. Journal of Biological Chemistry., 275:36832-36838.
8. Gardan, L., Christen, R., Achouk, W and Prior, P, 2004. Erwinia sp. Nov., a pathogen of papaya (Carica papaya). International Journal of Systematic and Evolutionary Microbiology.,54: 107-113.
9. Ross, P.L., Huang, Y.N and Marchese, J.N, 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Molecular and Cellular Proteomics., 3: 1154–1169.
10. Fukushima, A., Kusano, M., Redestig, H., Arita, M and Saito, K, 2009. Integrated omics approaches in plant systems biology. Curr Opin Chem Biol.,13:532–538.
11. Wiese, S., Reidegeld, K.A., Meyer, H.E and Warscheid, B, 2007. Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research**.** Proteomics., 7**:**340-350.
12. Zieske, L. R, 2006. A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. J. Exp. Bot., 57*,* 1501–1508.
13. Bantscheff, M., Schirle, M., Sweetman, G., Rick,J and Kuster, N, 2007. Quantitative mass spectrometry in proteomics: a critical review. Analytical and Bioanalytical Chemistry., 389(4): 1017–1031.
14. Zhang, H., Lv,L., Liu,H., Cui, L, Chen, G., Bi, P and Zhi Li, Zhang, 2009 .Profiling the potential biomarkers for cell differentiation of pancreatic cancer using iTRAQ and 2-D LC-MS/MS. Proteomics—Clinical Applications., 3(7): 862–871.
15. Adav, S.S., Chao, L.T and Sze, S.K, 2013. Protein abundance in multiplexed samples (PAMUS) for quantitation of Trichoderma reesei secretome. J Proteomics., 3:180-96.
16. Manavalan , A., Adav, S.S and Sze, S.K, 2011. iTRAQ-based quantitative secretome analysis of Phanerochaete chrysosporium. J Proteomics., 21;75(2):642-54.
17. Liu, D., Li, J., Zhao, S., Zhang, R., Wang, M., Miao, Y., Shen, Y and Shen, Q, 2013.Secretome diversity and quantitative analysis of cellulolytic Aspergillus fumigatus Z5 in the presence of different carbon sources. Biotechnol Biofuels.,6(1):149.
18. Liang L, Wu H, Liu Z, Shen R, Gao H, Yang J, Zhang K. 2013. Proteomic and transcriptional analyses of Arthrobotrys oligospora cell wall related proteins reveal complexity of fungal virulence against nematodes. Appl Microbiol Biotechnol. (19):8683-92.
19. Prasannath,K, 2013. Pathogenicity and Virulence Factors of Phytobacteria. Sch. Acad. J. Biosci.,1(1):24-33.
20. Strange, R.N, 2007. Phytotoxins producen by microbial plant pathogens. Nat Prod Rep., 24:127-144.
21. Lindgren, P. B, 1997. The role of hrp genes during plant–bacterial interactions. Annu. Rev. Phytopathol., 35, 129.
22. Wu, H. J., Wang,A.H and Jennings, M.P, 2008. Discovery of virulence factors of pathogenic bacteria. Current Opinion in Chemical Biology., 12:1–9.
23. Schulte, R and Bonas, U, 1992. A Xanthomonas Pathogenicity Locus Is Induced by Sucrose and Sulfur-Containing Amino Acids. Plant Cell.,4(1):79-86.
24. Ferreira, R.M., Moreira, L.M., Ferro, J.A., Soares, M.R.R., Laia, M.L., Varani, A.M., de Oliveira, J.C.F and Ferro, M.I.T, 2016. Unravelling potential virulence factor candidates in Xanthomonas citri. subsp. citri by secretome analysis. PeerJ., 4: e1734.
25. Degrassi, G., Devescovi, G., Bigirimana, J and Venturi, V, 2010. Xanthomonas oryzae pv.oryzae XKK.12 contains an AroQgamma chorismate mutase that is involved in ricevirulence. Phytopathology., 100:262-270.
26. Djamei, A and Kahmann, 2012. Ustilago maydis: dissecting the molecular interface between pathogen and plant. PLoS Pathogens 8:11-814.
27. Djamei, A., Schipper, K., Rabe, F., Ghosh, A., Vincon, V., Kahnt, J., Osorio, S.,Tohge, T., Fernie, A.R., Feussner, I.,Feussner, K., Meinicke, P., Stierhof, Y.D., Schwarz, H., Macek, B., Mann, M and Kahmann, R, 2011. Metabolic priming by a secreted fungal effector. Nature., 478:395–398.
28. Zhu, X., Soliman, A., Islam, M.R., Adam, L.R and Daayf, F, 2017. *Verticillium dahliae's* Isochorismatase hydrolase is a virulence factor that contributes to interference with potato's salicylate and jasmonate defense signaling. Front Plant Sci., 8:399.

29. Rigano, L.A., Siciliano, F., Enrique, R., Sendín, L., Filippone, P., Torres, P.S., Qüesta, J., Dow, J.M., Castagnaro, A.P., Vojnov, A.A and Marano, M.R, 2007. Biofilm formation, epiphyticfitness, and canker development in Xanthomonas axonopodis pv. citri. Molecular Plant-Microbe Interactions., 20:1222-1230.
30. Otto, K., Norbeck, J., Larsson, T., Karlsson, K.A and Hermansso, M,2001. Adhesion oftype 1-fimbriated escherichia coli to abiotic surfaces leads to altered composition of outer membrane proteins. Journal of Bacteriology., 183:2445-2453.
31. Münzinger, M., Budzikiewicz, H., Expert, D., Enard, C., and Meyer, J.M. 2000.Achromobactin, a new citrate siderophore of *Erwinia chrysanthemi*. Z Naturforsch [C]., 55: 328–332.
32. Dellagi,A., Segond, D., Rigault, M., Fargard. M., Simon, C., Saindrenan, P and Expert, D, 2009. Microbial siderophores exert a subtle role in Arabidopsis during infection by manipulating the immune response and the iron status. Plant Physiol.,150:1687–1696.
33. Franza, T., Mahe, B and Expert, D, 2005. Erwinia chrysanthemi requires a second iron transport route dependent on the siderophore achromobactin for extracellular growth and plant infection. Mol. Microbiol., 55:261–275.
34. Deng, W.L., Rehm, A.H., Charkowski, A.O., Rojas, C.M., and Collmer, A, 2003. *Pseudomonas syringae* exchangeable effector loci: sequence diversity in representative pathovars and virulence function in *P. syringae* pv. *Syringae* B728a. J Bacteriol., 185: 2592–2602
35. Degrave, A., Moreau, M., Launay, A., Barny, M.A., Brisset, M.N., Patrit, O., Taconnat, L., Vedel, R and Fagard M, 2013.The bacterial effector DspA/E is toxic in Arabidopsis thaliana and is required for multiplication and survival of fire blight pathogen. Mol. Plant Pathol.,14:506–517.
36. Piqué, N., Miñana-Galbis, D., Merino, S and Tomás, J.M,2015. Virulence Factors of *Erwinia amylovora*: A Review. Int J Mol Sci. 16(6): 12836–12854.
37. Deng , W., de Hoog, C.L., Yu, H.B., Li, Y., Croxen, M.A., Thomas, N.A., Puente, J.L., Foster, L.J and Finlay, B,B, 2010. Comprehensive proteomic analysis of the type III secretome of *Citrobacter rodentium*. J Biol Chem., 285: 6790-6800.
38. Schneider, D.J. and Collmer, A, 2010. Studying Plant-Pathogen Interactions in the Genomics Era: Beyond Molecular Koch's Postulates to Systems Biology. Annual Review of Phytopathology.,48(1): 457-479
39. Lee, J.H., Ancona, V and Zhao,Y, 2017. Lon protease modulates virulence traits in Erwinia amylovora by directly monitoring major regulators and indirectly through the Rcs and Gac-Csr regulatory systems. Mol Plant Pathol.,doi: 10.1111/mpp.12566.
40. Butler, A.M., Aiton, A.L., Warner, A.H., 2001. Characterization of a novel heterodimeric cathepsin L-like protease and cDNA encoding the catalytic subunit of the protease in embryos of *Artemia franciscana*. Biochem. Cell Biol. 79, 43–56.
41. Zhang, Y.; Bak, D.D.; Heid, H. and Geider, K,1999. Molecular characterization of a protease secreted by *Erwinia amylovora*. J. Mol. Biol.,289:1239–1251.
42. Maria Fedatto, L., Silva-Stenico, M.E., Etchegaray, A., Pacheco, F.T., Rodrigues, J.L and Tsai, S.M, 2006. Detection and characterization of protease secreted by the plant pathogen *Xylella fastidiosa*. Microbiological Research., 161:263—272
43. Burse A., Weingart, H., and Ullrich, M. S, 2004.The phytoalexin-inducible multidrug efflux pump AcrAB contributes to virulence in the fire blight pathogen *Erwinia amylovora*. Mol. Plant-Microbe Interact. 17:43-54.
44. Maggiorani Valecillos, A., Rodríguez Palenzuela, P and López-Solanilla, E, 2006. The Role of Several Multidrug Resistance Systems in *Erwinia chrysanthemi* Pathogenesis. Molecular Plant Microbe Interaction., 19(6): 607–613.
45. Fuangthong, M., Sallabhan, R., Atichartpongkul, S., Rangkadilok, N., Sriprang, R., Satayavivad, J and Mongkolsuk, S, 2008. The omlA gene is involved in multidrug resistance and its expression is inhibited by coumarins in Xanthomonas campestris pv. phaseoli. Arch Microbiol.,189(3):211-218.
46. Rohaiza, A.R., Bakar, N. A., Rozano, L., Badrun, R., Amin, N. M and Raih, M. F, 2014. Draft genome sequence of *Erwinia mallotivora* BT-MARDI, causative agent of papaya dieback disease. Genome Announcements 2(3), e00375-14
47. Vrancken, K., Holtappels, M., Schoofs, H., Deckers, T and Valcke, R, 2013. Pathogenicity and infection strategies of the fire blight pathogen *Erwinia amylovora* in *Rosaceae*: State of the art. Microbiology.,159:823–883.