DOI: 10.14456/gag.2019.6 Genomics and Genetics 2019, 12(2&3): 47–53

Difference in defense mechanism of two cassava cultivars to bacterial blight disease inferred by analysis of interspecies protein-protein interaction networks

Ratana Thanasomboon[1,2], Saowalak Kalapanulak[2,3], Supatcharee Netrphan[4] and Treenut Saithong[2,3]*

[1]*Biological Engineering Department, Engineering Faculty, King Mongkut's University of Technology Thonburi, Bangkok, 10140*
[2]*Systems biology and bioinformatics laboratory, King Mongkut's University of Technology Thonburi, Bangkok, 10150*
[3]*Bioinformatics and Systems Biology Program, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok, 10150*
[4]*National Center for Genetic Engineering and Biotechnology, Pathum Thani, 12120*

**Corresponding author: treenut.sai@kmutt.ac.th*

ABSTRACT

Cassava (*Manihot esculenta* Crantz) is an important economic crop worldwide as well as Thailand. One of the major problems resulting in its yield loss is cassava bacterial blight (CBB) disease caused by *Xanthomonas axonopodis* (*XAM*). However, the basic knowledge regarding cultivar-specific defense mechanism is limited. Therefore, in this work protein-protein interaction (PPI) networks of two cassava cultivars; KU50 and AM560, and *XAM* were reconstructed using domain-domain interaction method (DDI) to study interactions between *XAM* and cassava. The $PPI_{XAM\text{-}KU50}$ consisted of 24,509 interactions from 4,754 proteins while $PPI_{XAM\text{-}AM560}$ contained 28,962 interactions from 6,153 proteins. The results demonstrated that *XAM* used different mechanisms to invade specific cassava cultivar. This study would provide knowledge regarding the protein-protein relationship between bacterial pathogen and cassava.

Keywords: cassava; *Manihot esculenta* Crantz; protein-protein interaction; *Xanthomonas axonopodis*

INTRODUCTION

Protein-protein interaction (PPI) plays an important role for studying the molecular interactions between plants and pathogens. Many plant proteins interact directly with the pathogen proteins, and some of them can initiate plant defense responses to the infection. Understanding the PPI between plant proteins and pathogen proteins is a critical step for proposing precise targets against the pathogens and leading to super-crop breeding program. By using yeast two-hybrid system (Y2H), the first PPI involved in plant-pathogen was discovered in tomato. The interaction between the R protein of tomato, a serine/threonine protein kinase, and the AVR-Pto protein of *Pseudomonas syringae pv. tomato* could determine resistance to bacterial speck of tomato (Scofield *et al*., 1996). Moreover, Jia *et al.* (2000) also found the interaction between the *Magnaporthe grisea* AVR-Pita protein and the Pi-ta protein functioning in resistance to rice blast disease.

Nowadays, the advanced high-throughput technologies provide extensive biological information about plants and pathogens in the genomics, transcriptomics, proteomics levels which allows us to investigate the PPIs for each organism, and also the interactions between species at the whole genome level. For example, many genes potentially associated with pathogenicity of *Ralstonia solanacearum* (Salanoubat *et al*., 2002) and plant defense in *Arabidopsis thaliana* (Bishop *et al*., 2000) were reported. The PPI networks between plants and pathogens were predicted by interolog or domain-based methods for studying inter-species interactions, such as *R. solanacearum and* Arabidopsis (Li *et al*., 2012) and *Ustilaginoidea virens* and rice (Zhang *et al*., 2017). These predictions provided the potential PPIs leading to finding of important proteins and vital biological processes during the infection.

Cassava (*Manihot esculenta* Crantz) is an important staple crop since its edible starchy tuberous roots feed at least 800 million people worldwide (FAO, 2013). It can tolerate poor quality soil and the valuable starch extracted from cassava roots is used for many industrial purposes, such as food and beverages, cosmetics, medicine and printing. Cassava is also the economic crop for Thai farmers as 80 percent of its productions were exported making Thailand the largest

Received 28 June 2019; Accepted 14 August 2019

cassava exporter in the world (FAO, 2015). Although many cultivation systems have been developed to increase the productivity, the biotic stresses from various pathogenic diseases have caused in yield losses in many countries (Bock and Woods, 1983, Dutt *et al*., 2005). The cassava bacterial blight (CBB) disease is considered to be one of the important diseases in many cassava-growing locations. The symptoms are characterized by angular leaf spotting and blight, wilting, dieback, and vascular necrosis leading to cassava death. A previous report showed that *Xanthomonas axonopodis (XAM)*, gram negative bacteria, was the cause of this disease (Lozano, 1986). *XAM* invades and destroys the spongy mesophyll, then enters the vascular tissues where it propagates and spreads throughout the plant. In general, the plant shows the symptoms around 11-13 days after infection (Lozano, 1986), thus, plant probably performs some actions against the pathogen during this period before its death.

In this work, the PPI networks between *Xanthomonas axonopodis (XAM)* and two cassava cultivars; AM560 (experimental trait) and KU50 (high yield commercial trait) were reconstructed using bioinformatics approach based on protein domain method. This method relied on information on domain-domain interaction (DDI) which was based upon the observation that the proteins usually interact via specific domains. The resulting PPI networks provided the different interactions between each cassava cultivar and pathogen that would help us to gain a better understanding for cultivar-specific defense mechanism between plant–bacteria interactions.

MATERIALS AND METHODS

Reconstruction of species-specific PPI networks using domain-based method

PPIs of two cassava cultivars; KU50, AM560 and *XAM* pathogen were reconstructed using DDI information. This was based on the hypothesis that PPIs occur when binding domains of both proteins interact. The information of cassava AM560 and KU50 proteins were retrieved fromPhytozome V.9 (Goodstein *et al*., 2012) and Wang et al (2014), respectively while information regarding *XAM* proteins were collected from Bart et al (2012). All proteins from each species were predicted for their domain based on Pfam (Finn *et al*., 2013) with E-value and aligned sequence length coverage cutoffs at 0.001 and 90%, respectively. The DDIs retrieved from iPfam database (Finn *et al*., 2014) were used to predict the PPIs for each organism. In this work, the PPI was predicted when at least one DDI occur between the protein pair.

Reconstruction of cassava–pathogen PPI networks using domain-based method

The interspecies PPI networks were reconstructed based on *XAM* secreted proteins and whole genome sequence of cassava proteins. To identify the invaded protein from pathogen, all *XAM* proteins in its genome were identified as effector proteins by using Effective database (Jehl *et al*., 2010) which was based on the identification of eukaryotic-like protein domains and the recognition of signal peptides in amino acid sequences. The *XAM* effector proteins and all proteins of the two cassava cultivars were used to predict interspecies interactions based on their domains and DDI information. The criteria of interspecies PPI prediction were the same as individual species PPI prediction.

Functional analysis

Biological functions of the proteins based on ontology of the constituent proteins in all constructed networks were examined through AgriGO (Du *et al*., 2010) and visualized by REVIGO (Supek *et al*., 2011).

RESULTS AND DISCUSSION

Predicted PPI network of the two cassava cultivars, AM560 and KU50, and bacterial pathogen, *XAM*

The domain-based prediction approach provided PPI networks of specific cassava cultivars and the pathogen. Based on at least one DDI occurs between each protein pair, the PPI networks of KU50, AM560 and *XAM* were predicted. The KU50 PPI network contained 2,247,030 interactions from 16,049 proteins (~41% of all 38,845 proteins in the genome) while the AM560 PPI network consisted of 3,625,335 interactions from 19,161 proteins (~56% of total 34,152 proteins in the genome). For *XAM* PPI network, it consisted of 21,102 interactions from 2,227 proteins (~83% of total 4,216 proteins in the genome). The number of proteins and interactions of each organism were shown in Table 1.

Although the three PPI networks were successfully reconstructed, this method was confined by the amount of domain and domain-domain interaction information. Around 38, 19 and 17 percent of all proteins in KU50, AM560 and *XAM* genome, respectively, could not be studied because their domain information are unavailable. Thus, if there are more available domain and DDI information in the future, the predicted networks will be nearly completed to cover the actual genome-scale study.

Table 1 Protein information and characteristics of the reconstructed PPI networks of cassava cultivars, KU50 and AM560, and *XAM.*

Description	cassava		pathogen
	KU50	**AM560**	***XAM***
Numbers of proteins in the genome	38,845	34,152	4,216
Numbers of proteins with domain information	23,917	27,732	3,494
Numbers of proteins in the reconstructed PPI network	16,049	19,161	2,227
Numbers of PPIs based on DDI method	2,247,030	3,625,335	21,102

The interspecies PPI networks providing the possible protein interactions between cassava and XAM during the infection

The PPI network between cassava and *XAM* was predicted by using DDI information of cassava and *XAM* proteins. The $PPI_{XAM\text{-}KU50}$ interactive protein network contained 24,509 interactions from 368 *XAM* and 4,386 KU50 protein (Figure 1a) while the $PPI_{XAM\text{-}AM560}$ interactive protein network consisted of 28,962 interactions from 422 *XAM* and 5,731 AM560 proteins (Figure 1b). The result showed that *XAM* used almost the similar proteins to infect KU50 and AM560 cultivars (Figure 2a). Around 364 effector proteins interacted with cassava proteins involving in signaling pathways, response to stimuli, biological regulations and transmembrane transport in the pathogen. This result was corresponded to a previous report that showed important functions of XAM proteins during infection (Li and Wang, 2011). On the other hand, our prediction demonstrated that some *XAM* proteins interacted with only one cultivar leading to specific interactions based on each unique species (Figure 2a). Only four *XAM* proteins infected with KU50 proteins functioning as carboxylesterases via eight PPIs (Figure 2b). This result corresponded with a previous study demonstrating that carboxylesterases involved in Arabidopsis-pathogen interactions (Marshall et al., 2003). Furthermore, 58 unique *XAM* proteins infecting AM560 were found to interact with 17 AM560 proteins leading to specific 69 interactions between *XAM* and AM560 (Figure 2c). Most of the *XAM*-targeted KU50 proteins were TonB-dependent receptor and ligand-gate channel proteins which have roles in signaling and transport processes (Chimento *et al*., 2003). These specific targets and their interactions in $PPI_{XAM\text{-}KU50}$ and $PPI_{XAM\text{-}AM560}$ interactive protein network probably related to different responses during bacterial infection in specific cassava cultivars, KU50 and AM560.

The *XAM*-targeted proteins were further analyzed based on GO enrichment analysis. For KU50, most of the target proteins involved in cellular homeostasis, biological regulation and signaling processes while most of the *XAM*-targeted proteins in AM560 related to protein phosphorylation, responses to stimuli and cellular metabolism processes. These results showed that each cassava trait used different processes to defense pathogen, even though *XAM* used the similar proteins to invade KU50 and AM560. Moreover, these target proteins in both traits were further used to find their partners through KU50 and AM560 PPI network in order to understand the plant responses during *XAM* infection. The protein-protein interactions between *XAM*-targeted proteins in KU50 and their partners (1st neighbor proteins) in KU50 PPI network was visualized. It consisted of 1,638,939 interactions from 10,937 proteins. It implied that the *XAM* targeted proteins could interact with the other 68 percent of all proteins in the KU50 PPI network which could affect various biological processes such as glycosyl compound biosynthesis, energy couple proton transport and cell recognition (Figure 3a). While PPI between *XAM*-targeted proteins in AM560 and their partners in the AM560 PPI network, contained 2,682,573 interactions from 12,556 proteins (65 percent of total proteins in AM560 PPI network) which involved in post-translational modification, cell cycle and response to stress processes (Figure 3b). These results suggested that the plant proteins responding to *XAM* infection in KU50 and AM560 are different based on PPI information.

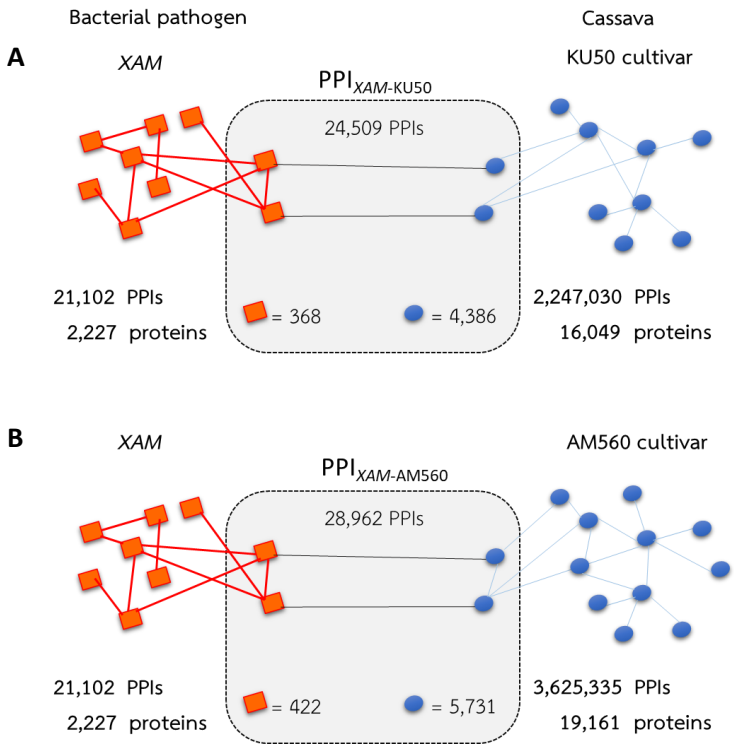

Figure 1 Scheme of PPI networks between cassava and *XAM*. The interspecies PPI networks were represented in a) $PPI_{XAM\text{-}KU50}$ and b) $PPI_{XAM\text{-}AM560}$ interactive protein network, respectively. The nodes represented proteins; cassava (blue circle) and *XAM* (orange rectangular) while edges represented interactions between *XAM* proteins (orange), cassava protein (blue) and *XAM*-cassava proteins (black).

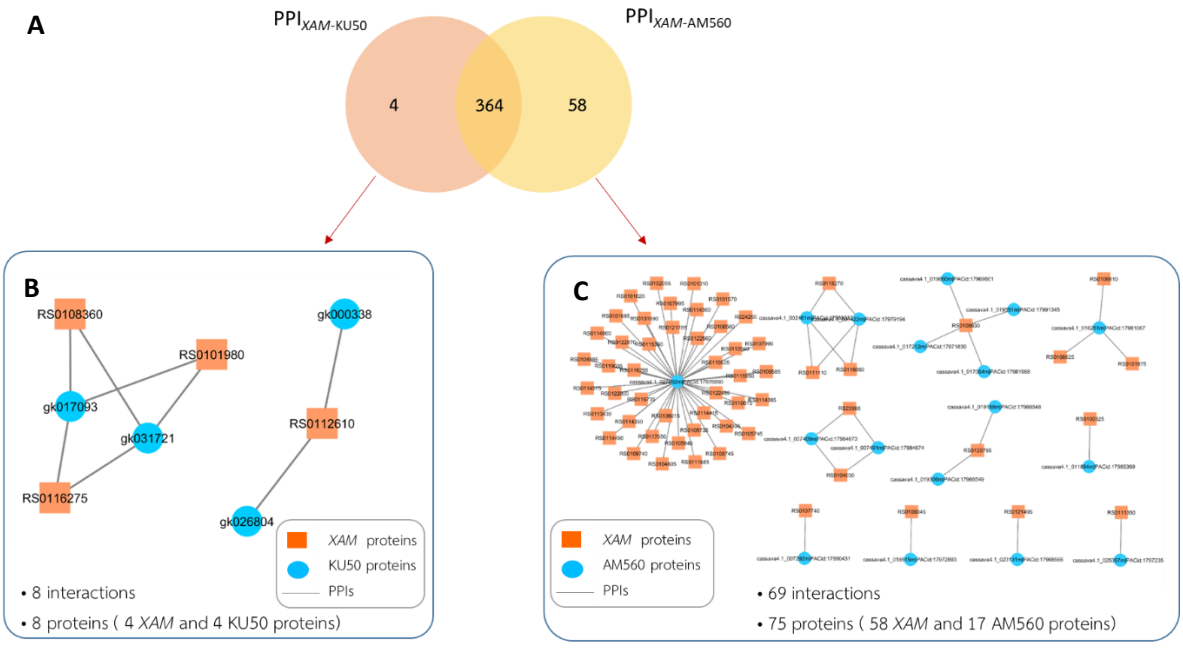

Figure 2 *XAM* effector proteins in interspecies PPI network of each cassava cultivar, KU50 and AM560. a) Venn diagram demonstrating *XAM* effector protein comparison when invaded to specific cassava cultivar, KU50 (368) and AM560 (422). PPIs between unique *XAM* proteins with KU50 proteins (b) and with AM560 protein (c) were showed. The nodes represented proteins; cassava (blue circle) and *XAM* (orange rectangular) while edges represented interactions between *XAM* and cassava protein.

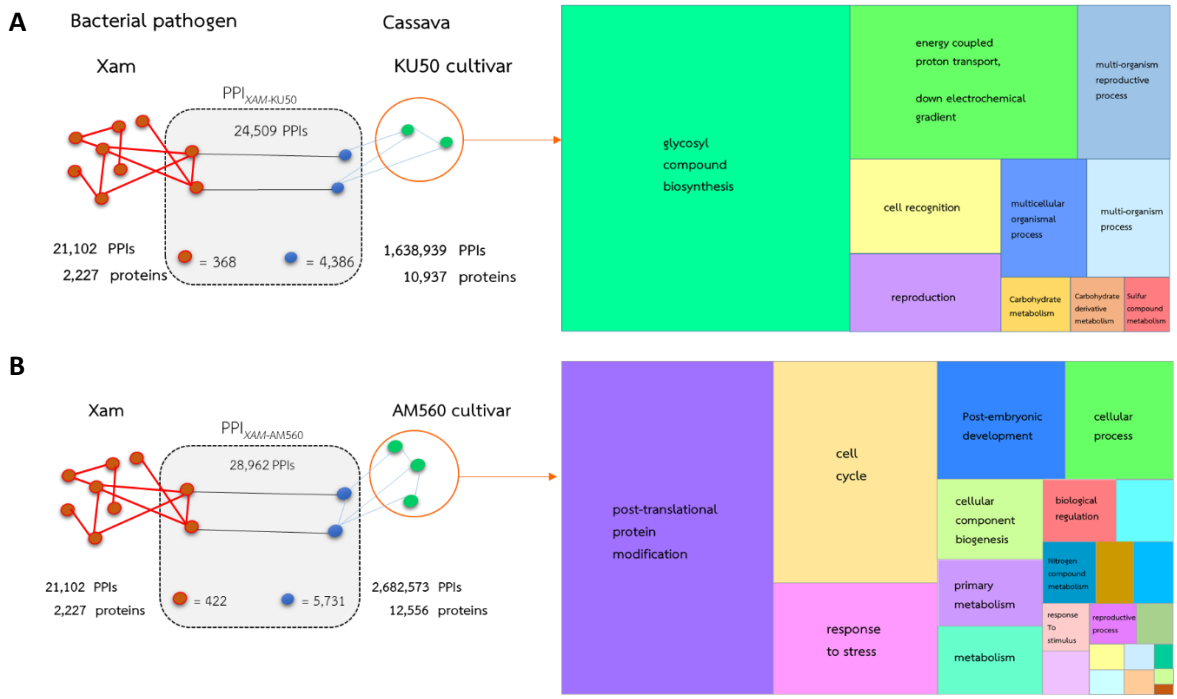

Figure 3 Functional content of cassava proteins which are the first neighbor of *XAM*-targeted proteins in KU50 (a) and AM560 (b). Their functional contents were represented as tree map based on GO enrichment analysis (*p*-value <0.05) in biological process.

In addition, the important proteins of cassava during *XAM* infection could be identified through node degree defined as the number of interactions that a node (protein) obtains. For $PPI_{XAM\text{-}KU50}$ network, two key proteins involved in *XAM* infection were identified through node degree information. Firstly, the peptidylprolyl isomerase (RS0111635) of *XAM* showed the highest (907) number of connections with KU50 proteins functioning as calcium dependent protein kinase (CDPK), calmodulin, receptor-like protein kinase (RLK), peptidyl-prolyl cis-trans isomerase (PPIases) and mitogen-activated protein (MAP) kinase proteins which involved in many biological processes, including intracellular signaling, transcription, inflammation, and apoptosis (Hoffmann and Schiene-Fischer, 2014). In pathologic process, the peptidylprolyl isomerase was reported for modifying outer membrane proteins to avoid or suppress host cell immune response and use host peptidylprolyl isomerase for modification of effector proteins necessary for pathogenesis (Kromina *et al*., 2008). The second key protein was methyltransferase (gk000101) of KU50 protein which obtains the highest number of interactions with pathogen proteins including histidine kinase, methyltransferase ATPase and chemotaxis proteins. This result corresponded to Berr *et al.* (2010). They reported that methyltransferase played a crucial role in Arabidopsis defense system. For $PPI_{XAM\text{-}AM560}$ network, the peptidylprolyl isomerase (RS0117785) showed the highest (1,187) number of connections with AM560 proteins. This protein could interact with many types of AM560 proteins involving various biological processes as mentioned above. Moreover, the interspecies network also provided interesting AM560 proteins which could be the targets of many *XAM* proteins. For example, signal transduction histidine kinase (cassava4.1_026514m) that interacted with 40 *XAM* proteins. This protein was reported to play a major role in cell signaling process (Bhate *et al*., 2015).

The relationship between cassava and bacterial pathogen was investigated through protein-protein interaction network. The KU50, AM560 and *XAM* PPI as well as interspecies PPI network were constructed. The $PPI_{XAM\text{-}KU50}$ and $PPI_{XAM\text{-}AM560}$ network provided possible active *XAM* and cassava proteins during infection. The list of key protein targets of *XAM* invading KU50 and AM560 was proposed. However, the PPIs between *XAM* and specific cassava cultivars were proposed based on annotated protein domain information, that are around 62%, 81% and 83% of all proteins in the genome of KU50, AM560 and *XAM*, respectively. Therefore, more information of protein domains would enable us to more comprehend plant-bacterial pathogen relationship and shed light on resistant trait screening in markers-assisted breeding programs.

ACKNOWLEDGEMENTS

The authors would like to thank The National Center for Genetic Engineering and Biotechnology (BIOTEC, NSTDA) for R.T post-graduate scholarship. This work was supported by National Research Council of Thailand (NRCT) and National Science and Technology Development Agency (NSTDA) under Thailand Research Organizations Network (research grant: P-13-50437).

REFERENCES

Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, Krasileva K, Dahlbeck D, Medina C, Alicai T (2012) High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance, Proceedings of the National Academy of Sciences, 109 (28): E1972–E1979.

Berr A, McCallum EJ, Alioua A, Heintz D, Heitz T, Shen WH (2010) Arabidopsis histone methyltransferase set domain group8 mediates induction of the jasmonate/ethylene pathway genes in plant defense response to necrotrophic fungi, Plant Physiology, 154(3): 1403–1414.

Bhate MP, Molnar KS, Goulian M, DeGrado WF (2015) Signal transduction in histidine kinases: insights from new structures, Structure, 23(6): 981–994.

Bishop J, Dean AM, Mitchell-Olds T (2000) Rapid evolution in plant chitinases: molecular targets of selection in plant-pathogen coevolution, Proceedings of the National Academy of Sciences, 97(10): 5322–5327.

Bock K, Woods R (1983) Etiology of African cassava mosaic disease, Plant Disease, 67(9): 994–995

Chimento DP, Mohanty AK, Kadner RJ, WienerM.C (2003) Substrate-induced transmembrane signaling in the cobalamin transporter btub, Nature Structural and Molecular Biology, 10(5): 394.

Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) Agrigo: a GO analysis toolkit for the agricultural community, Nucleic Acids Research, 38: W64–W70.

Dutt N, Briddon R, Dasgupta I (2005) Identification of a second begomovirus, Srilankan cassava mosaic virus, causing cassava mosaic disease in India, Archives of Virology, 150(10): 2101–2108.

FAO (2013) Save and Grow: Cassava Sustainable Production Intensification, Rome, pp.121–128.

FAO (2015) Food Outlook: Biannual Report on Global Food Markets. FAO, Rome, pp. 1–133.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J (2013) Pfam: the protein families database, Nucleic Acids Research, 42(D1): D222–D230.

Finn RD, Miller BL, Clements J, Bateman A (2014) Ipfam: a database of protein family and domain interactions found in the protein data bank, Nucleic Acids Research, 42(D1): D364–D373.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N (2012) Phytozome: a comparative platform for green plant genomics, Nucleic Acids Research, 40(D1): D1178–D1186.

Hoffmann H, Schiene-Fischer C (2014) Functional aspects of extracellular cyclophilins, Biological Chemistry, 395(7): 721–735.

Jehl MA, Arnold R, Rattei T (2010) Effective-a database of predicted secreted bacterial proteins, Nucleic Acids Research, 39: D591–D595.

Jia,Y, McAdams SA, Bryan GT, Hershey HP, Valent B (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance, The EMBO Journal, 19(15): 4004–4014.

Kromina K, Ignatov A, Abdeeva I (2008) Role of peptidyl-prolyl-cis/trans-isomerases in pathologic processes, Biochemistry Supplement Series A: Membrane and Cell Biology, 2(3): 195–202.

Li J, Wang N (2011) Genome-wide mutagenesis of Xanthomonas axonopodis pv. citri reveals novel genetic determinants and regulation mechanisms of biofilm formation, PLoS One, 6(7): e21804.

Li ZG, He F, Zhang Z, Peng YL (2012) Prediction of protein–protein interactions between Ralstonia solanacearum and Arabidopsis thaliana, Amino Acids, 42: 2363–2371.

Lozano JC (1986) Cassava bacterial blight: a manageable disease, Plant Disease, 70(12):1989–1993.

Marshall SD, Putterill JJ, Plummer KM, Newcomb RD (2003) The carboxylesterase gene family from Arabidopsis thaliana, Journal of Molecular Evolution, 57: 487–500.

Salanoubat M, Genin S, Artiguenave F, Gouzy J, Mangenot S, Arlat M, Billault A, Brottier P, Camus JC, Cattolico L, Chandler M (2002) Genome sequence of the plant pathogen Ralstonia solanacearum. Nature, Vol. 415(6871): 497.

Scofield SR, Tobias CM, Rathjen JP, Chang JH, Lavelle DT, Michelmore RW, Staskawicz BJ (1996) Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato, Science, 274(5295): 2063–2065.

Supek F, Bošnjak M, Škunca N, Šmuc T (2011) Revigo summarizes and visualizes long lists of gene ontology terms, PloS One, 6(7): e21800.

Wang W, Feng B, Xiao J, Xia Z, Zhou X, Li P, Zhang W, Wang Y, Moller BL, Zhang P (2014) Cassava genome from a wild ancestor to cultivated varieties, Nature Communications, 5:5110.

Zhang K, Li Y, Li T, Li ZG, Hsiang T, Zhang Z, Sun W (2017) Pathogenicity genes in Ustilaginoidea virens revealed by a predicted protein–protein interaction network, Journal of Proteome Research, 16(3): 1193–1206.