

TABLE OF CONTENTS

	<u>PAGE</u>
REVIEW OF LITERATURE	1
Introduction	7
Network analysis	8
The unique values of network analysis	12
Networks and Plant Pathology	14
Using Network analysis to understand plant disease spread	15
Using Network analysis to understand molecular plant pathology	18
A variety of Network Construction Methods	21
Summary	24

REVIEW OF LITERATURE

The pressure of increasing world population brings to bear more on the people of developing countries. In order to deal with this pressure farmers, researchers and extensionists are forced to join “Green Revolution”. The aim of this Revolution is to provide food for the increasing population by enhancing the harvest per unit and the intensification of the number of cultivation seasons. The Green Revolution was introduced during the early part of the 6th decade of the 20th century accelerated food production of the world.

The available statistic record proves that there are some difference of rice production system in different countries. The result of fields survey of Duwayri et al. (2000) showed that rice production system in tropical Asia is driven by high external input usage including the play-off between investment in labor and labor-saving technologies, new seed and inputs, or in other ways of raising fertility and productivity. As a series of consequence of global warming, increased population growth, and increased constraints on resources, agriculture has been changing remarkably from the past.

External input usages in rice production systems

High yield varieties High yielding hybrid seeds which were introduced by Green Revolution were new to our environment. The growers had to practice new techniques to get higher production. Due to the fact that hybrid seeds were more sensitive to nutrients the growers were encouraged to use chemical fertilizer in large quantities as external inputs. As the new crops were foreign to the environment they were susceptible to pests and

diseases. Consequently the necessity arose to apply chemicals, which became an additional burden to growers. Application of chemical fertilizer and pesticides increased the cost of production.

Chemicals The hybrid varieties were dwarf in nature and could not compete with weeds. Application of herbicide or manual weeding was essential to mitigate the competition between hybrid varieties and weeds. From land preparation to harvesting all cropping practices, related to hybrid varieties were more labor intensive.

Mechanization Mechanization was an integral component of Green Revolution. To increase the working efficiency of the production system machinery like tractors have been introduced. These machines required mineral fuel.

Irrigation The water consumption of new crops was also higher. Therefore it was necessary to improve irrigation facilities. The external resources were used extensively in this agricultural system many of the resources used in high external inputs.

The importance of Rice Pest

Deepening into each type of rice pest, estimation of global yield losses caused by weeds, disease, animal pests are at 10.2 %, 12.2 % and 15.5 % of the attainable yields, respectively. Whiteheads caused by stem borers) is the most destructive insect pest causing approximately 1.9 to 5.8 % yield loss, and for disease, sheath blight is the highest damage the yield loss at 5 to 10 % Willocquet et al. (2004).

Rice pests are the organism causing the economic loss in rice production, such as arthropods (insects and mites), pathogens (bacteria, fungi, and virus), which cause the diseases, nematode, mollusk (snails), and vertebrates (rodents, birds).

There are more than 100 insect feeding on rice, but fortunately most of them do not damage reaching to economically destructive enough to need to control because the rice plants have ability to recover from such injuries GRiSP (2013) if they are in the vegetative stage. The relative importance and widely distributed rice insect pest are stem borers, leafrollers, brown planthoppers (*Nilpavata lugens*, BPH) and gall midge. Particularly, some insects cause serious problems various from country to country, For example China, Vietnam, India and Thailand has recently experienced serious problems with brown planthopper. Stem borer and leaf-feeding insects are also found in most rice growing areas and can sometimes cause major yield losses.

A wide range of rice disease affect rice Ou (1985), among which blast , sheath blight, bacterial blight, brown spot, and several viral disease, including rice tungro, are of primary concern. As with insect pests, rice disease can be categorized as chronic yield reducers (e.g., brown spot and sheath blight), whereas other diseases cause sporadic, often large-scale, and extremely damaging epidemic (e.g., blast and most virus disease). Such uncertainty of risk adds another layer of complicity to assessing priorities and developing sustainable decision making process, from the field to the nation scales.

Rice Disease and Insect Pests Occurrence

Rice Brown Spot Rice brown spot (BS), caused by *Cochliobolus miyabeanus* (Ito and Kuribayashi) Drechs. ex Dastur. (Anamorph: *Bipolaris oryzae* (Breda de Haan) Shoemaker), is the disease chronically affects millions of hectares worldwide every year. It especially occurs in environments where water supply is scarce and it is often combined with imbalances in plant mineral nutrition, especially the lack of nitrogen. These factors are commonly associated with resource poor farmers' fields, which at this beginning of the 21st century, are still. Reported yield losses in relative terms vary widely from 4

to 52 %. However, accurate, systematic estimates are lacking. BS is conventionally perceived as a secondary problem that reflects rice crops that experience physiological stresses, e.g. drought and poor soil fertility, rather than a true infectious disease. Much remains to be understood about the mechanisms leading to epidemics and crop losses Barnwal et al. (2013).

The studies of crop losses for long term strategies of pest management

An ultimate goal of the studies of plant pathology is to suggest recommendations for disease management. Future crop protection strategies must take into account sustainability, production efficiency and environmental friendly protection and providing a rationale for integrated pest management or IPM Mew et al. (2004); Zadoks and Schein (1979); Savary et al. (2012).

Realistic approach to crop protection should consider the various pests, including disease, that affect to a crop. Cropping practices represent major interactions within the disease tetrahedron Zadoks and Schein (1979).

In order to improve effective pest management in the targeted production situation, the empirical approaches have been proposed and used, which is based on spontaneous or manipulated injuries due to disease, insects, and weeds. This approach have been used to characterize rice production situation and injuries profile in the tropical Asia, and quantify their interaction Savary et al. (2000).

Analysis of the production situation x injuries profile doublet may help identify opportunities for improved pest management strategies and determine the underlying decisions, either tactical or strategies, that might be alter for better and sustainable dis-

ease management Mew et al. (2004).

Characterizations at the field level observed sites of South and Southeast Asia led to the conclusion that the observed injuries profiles (i.e., the combination of disease and pest injuries that may occur in a given farmers' field) are strongly dependent on the production situation. Production situation was defined as the natural resources need to produce the amount of rice under different production environment and vary socio-economic system. These studies were conducted by survey Savary et al. (2000) which generated a considerable amount of data set used to characterize associations between crop health and PSs in 456 lowland rice farmers' fields in tropical and subtropical Asia (India, the Philippines, Vietnam, and China). These data were collected at the individual farmer's field level, from field measurements of crop status and injuries (diseases, pests, and weeds) at four development stages, along with information on PSs and direct yield estimates. The data were analyzed using non parametric multivariate technique: cluster analyses with chi-square distance and correspondence analysis. The result showed PSs differed markedly with respect to fallow period duration, crop establishment method, water shortage, and mineral fertilizer inputs. Syndromes of injuries strongly differed, in particular with respect to six rice diseases: sheath blight, sheath rot, brown spot, leaf blast, neck blast, and bacterial leaf blight. Injury syndromes are so strongly associated with the injuries caused by these six diseases, the Ss involving these six diseases only can be used as markers of the overall injury syndromes.

From the attribution of production situations, provides the strategic decisions for pest management Mew et al. (2004). Because patterns of cropping practices and injury profiles depend on the considered production situation, management suggests handling not only of a single pest but several pests at a time (a pest injury profile) which

is meaningful to the individual farmer, and linked to production situations, which in turn are linked to the socioeconomic environment.

Even analysis of survey data has shed light on the disease, pest and weeds from various production scenarios. However, the knowledge gap about their organization and the interrelationships of the variables still exists.

Network: Introduction

Farmers lose an estimated average of 37% of their rice crop to pests and diseases every year. In addition to good crop management, timely and accurate diagnosis can significantly reduce losses.

The applications of network analysis have increased exponentially over the past two decades in various disciplines. Even though documented applications of network analysis in plant pathology are still relatively sparse, network applications in the social sciences, systems biology and ecology have been increasingly found. Here I review the empirical works that exist and argue that network analysis is a promising approach for exploring questions in the context of plant pathology.

This chapter contains four sections to thoroughly review of network analysis and its applications. In the first sections, I introduce a brief overview of the concepts and methods of network analysis, and I then discuss the unique values of network analysis that are not found in other approaches. In the third section I focus on the application of network analysis to current applications of plant pathology research, in particular, plant disease epidemiology and molecular plant pathology, for which network analysis has been broadly applied, and increasingly documented. In last section I will show a variety of applications of network analysis to study biological networks. Differential and dynamic network analysis are promising tools to understand biological responses to the different conditions or times.

Network Analysis

Introduction to network analysis

Network analysis is used for determining relationships between elements of interest. It offers toolkits for visualizing data in a network model and measuring its properties, and network thinkings (Proulx et al., 2005). It has been widely used by various branches of science, such as social science, ecology, biology, computer science, and many others to study the interactions between elements, e.g., the relationships of students in school (Moody, 2001), species in food webs (Krause et al., 2003), interactions of genes or proteins in cells (Guimera and Amaral, 2005), or the connections of computer in the network (Pastor-Satorras and Vespignani, 2001; Newman, 2006).

Newman (2003) loosely categorized four types of networks based on different complex data. The first category is social network, representing sets or groups of people forming some patterns of contacts or interactions between them such as the patterns of friendship or business relationships. Analyzing the structure of whole social entities gives us the perspectives from a social network, which enables us to explain the patterns observed. Moody (2001) analyzed the social behaviors in high school students using social network approaches. (Kasari et al., 2011) applied network analysis to compare the social relationships and friendships between children with and without autism spectrum disorder (ADS). The second type of network is an information network or knowledge network. The classic example of this network is the network of citations between academic papers (Newman, 2003). The articles cited other papers, which have related topics. They formed a citation network that has vertices as articles and direct links as citations. The citation network visualizes the structure and the movement of the

information. The third category, technological network, is object connected network, or man-made network which represents a physical connection between objects. This network is mostly applied for illustrating physical structures and systems such as the electrical power grid, the connections of rivers, transport systems, etc. The fourth category of network is a biological network. It represents the biological systems such as genes to genes, genes to protein, protein to protein interactions, which enable biologists understand the connections and interactions between individual constituents including genes, proteins, and metabolites at the level of the cell, tissue and organ to ultimately describe the entire organism system. Biologists use biological networks in various branches of biology at different levels (from a single molecule to an entire organism). For example, Yang et al. (2014); Barabási and Oltvai (2004) studied in the patterns of gene expression in different conditions and different types of cells (normal cells and cancer cells) in order to characterize the genes that change and do not change following the particular conditions; Freilich et al. (2010) applied a molecular ecological network analysis to study the communities of soil microorganisms. Networks revealed the complex relationships between microbial species in soils and their communities. Moreover, network analysis enables ecologists to understand ecological properties and predict the ecological roles of species in a soil ecosystem. Although the application of each type of network approach varies, all four categories of networks share a common empirical focus on relational structure and a similar set of mathematical analysis.

Network analysis can be a powerful tool to study plant disease. Moslonka-Lefebvre et al. (2011); Jeger et al. (2007) applied network models and concepts to study disease spreads in regional networks of plant nurseries and garden centers of *Phytophthora ramorum*, the oomycete causing Sudden Oak Death in the West Coast of the USA

and leaf blight and dieback in many ornamental shrubs both in America and Europe. The result could be applied to design measures to control plant disease epidemic from movement of infected material among plant nurseries. Shaw and Pautasso (2014) recently reviewed critically about network analysis to plant disease management (*i.e.*, to design ways to reduce the flow of disease in traded plants, to find the best sites to monitor as warning sites for annually reinvading diseases, to understand the fundamentals of how plant pathogen spreads in different structures of simulated trade network models). Windram et al. (2014) reviewed the applications of network analysis in molecular plant pathology. Network models were applied to reveal plant defense mechanism during plant-pathogen interactions.

Concepts, principles, and methods of network analysis

A network represents relationships between elements of interest, which are defined by links (edges) among nodes (vertices). Nodes can be units of interests or studies, and links represent interactions between nodes. Network analysis aims to identify the patterns of associations among nodes, not only features or attributes of particular nodes.

Network analysis follows three principles. Nodes and their behaviors are mutually dependent, not autonomous; links between nodes can be channels for transmission of both material (*e.g.*, money, disease) and non-material (*e.g.*, information, knowledge, relationship, interaction) and; persistent pattern of association among nodes create structure that can define, enable, or restrict the behavior of a node.

Network models have two different organizational structures depending on goals of the representation and analysis (Borgatti et al., 2013). Flow models, commonly named directed graphs, represent the network as a system of pathways along which

something move such as transportation networks (*e.g.*, of highways, railways and air-lines) or communication networks. Since flow networks show the directions of movement between nodes, thus these networks are interesting to study their behaviors. Analysis of flow networks can identify which nodes are more active or which ones are more important connectors. Jeger et al. (2007); Shaw and Pautasso (2014) applied such network models to study plant disease spread. Architectural models, or indirected graphs, are mainly used to determine the structure of the network, seeking to discern whether specific structures lead to similar outcomes or whether nodes in similar network positions behave in similar ways. Ecological applications related to the ecology and spatial structure of “community” tend to be organized and analyzed as architectural models. For example, Faust et al. (2012) studied networks of soil microbial interactions. Network models can describe how microbial populations change over time, which will require the use of dynamic models of microbial communities. Beyond these basic principles, network analysis enables the calculation of structural properties of nodes, groups, or the entire network.

Measuring network properties

A network is made up of nodes and links from relational data. It is constructed from an adjacency matrix, which is obtained from analysis using metric algebra techniques. The row and column headings for an adjacency matrix are identical, listing the names of the components involved in the network. In the simplest case, the cells of the matrix are coded with “1” if a link exists between the node or “0” if no edge exists. However, a link can be valued. Value indicates a characteristic of the relationship that the research has quantified. The values may be binary, such as whether two friends recognize each other, or variable strength (*e.g.*, the number of mutual friends between two

friends). A network link need not imply positive or cooperative interaction; they can also be a negative or competitive interaction between two individuals.

The distribution of links in a network suggests two important structural characteristics: centrality (importance) of nodes in the network and division of the network into subgroups. Variants of centrality in a network include degree, closeness, and betweenness. Degree centrality of a node is the sum of the value of the links between that node and every other node in the network. This measure tells us how well-connected a particular node is to the other nodes. Closeness centrality is calculated using the length of the path between a node and every other node. This measure could estimate the time required for information or resources to propagate to a given node in a network. Betweenness centrality corresponds to the number of paths in the network that pass through a particular node, and therefore measures the dependence of a network on a particular node for maintaining connectedness (Toubiana et al., 2013). Deng et al. (2012); Newman (2003); Toubiana et al. (2013) are recommended references for descriptions of the theory and uses, as well as the formal calculation of these measures.

The unique values of network analysis

There are four key points that will help to understand network analysis 1) how it differs from traditional approaches of scientific research; 2) how it relates to those traditional approaches; 3) how networks are constructed, manipulated and measured; and 4) what value network analysis offers beyond traditional approaches.

The first point of network analysis is that there are two types of data represented in the network graphs; technical and rational data. The first is data characterizing the actors or variables being studied referring to attributes. Attributes describe characteristics of individual actors or variables, for example their race, income or physical location,

and are the primary variables considered in traditional approaches. The second type of data is relational data, that is, data about the relationships between individual nodes. For example, Lazega and Pattison (1999) represented the network model of cooperation among lawyers in three law firms, through the exchange of various type of resources among them. This model consisted of over 70 lawyers in three different offices in three different cities. Relational data reflected to resources exchange and additional attribute information were included type of practice, genders, and seniority of each lawyers.

Relationships are also referred to as edges (links) in network analysis. Edges cannot be attributed to any single actor. Rather, edges only exist between nodes. This leads to the second point about network analysis that it requires a different conceptual approach. Because edges only exist between nodes, it is useful to think of edges existing in a separate dimension from nodes, who are connected in physical space. This dimension is sometimes referred to as relational space. To visualize the difference, think of someone far away with whom you correspond regularly, say using a phone, email, or Facebook. Even though the two of you are not physically close, you have a strong relationship. The two of you are distant in physical space but close in relational space. This notion of relational space is in part what means when he refers to the space of flows as something distinct from the space of places (Castells, 2001).

The third point that distinguishes network analysis from other approaches is it involves different methods of analysis. Because traditional research methods consider variable attributes in a wide variety of statistical analyses such as measures of center (*e.g.*, mean, median, etc.) and dispersion (*e.g.*, standard deviation, range, etc.), these methods are sometimes referred to as variable analysis, whereas, network analysis models relational data and to measure various characteristics of network structure.

For example, for lawyers data (Lazega and Pattison, 1999), it is natural to ask to what extent two lawyers that both work with third lawyer are likely to with each other as well. This notion corresponds to the social network concept of transitivity and can be captured numerically through an enumeration of proportion of vertex triples that form triangles, so-called cluster coefficient.

The idea that network structure may be correlated with variable attributes and behaviors is the fourth point to consider in comparing network analysis to other approaches. In network analysis, the arrangement of the network in relational space is basically correlated with the behavior and attributes of those variables. For example, in the network created by Lazega and Pattison (1999) lawyers of the same firm may share similar attributes such as office location or department, and lawyers in similar roles within that network may share similar behaviors. Basically, conventional approaches measure various attributes of variable (nodes in a network) and attempts to discern something about the relationships between actors (edges in a network) based on those attributes. When the network structure is simple and the differences in node attributes are clear, the conventional analytic approach is sufficient. However when relationships are complex or node attributes are more nuanced, clear answers using conventional analysis may prove elusive. As a result, network analysis offers a tool to help researchers visualize the large network and disentangle some of the relational complexities within the network, just as cluster analysis and multivariate analysis for help research disentangle the complex data.

Networks and Plant Pathology

Recently, a broad expansion of applications of network analysis has occurred across many disciplines. It has been evaluated as a promising tool to study a complex

system. Plant pathologists have applied network analysis for their research. Moslonka-Lefebvre et al. (2011); Jeger et al. (2007); Windram et al. (2014) supported that network analysis can be fruitful models in many applications relevant to plant pathology because of its generality and flexibility. For example, the network of main fresh cut flowers movements among European countries was determined the likelihood of introduction of new pathogens and other organisms associated with plants (Eagling, 2007), and plant-pathogen interaction network models were applied to present plant defense mechanisms (Dietz et al., 2010).

The development of network analysis challenges conventional approaches to uncover rational complexities of plant pathology studies. Two fields of research relevant to plant pathology presented particularly strong growth and proved that network analysis has significant potential to augment traditional analysis methods. The first is plant disease epidemiology, which investigates questions related to plant disease spread. The second is plant molecular biology, which investigates questions related to biological networks.

Using Network analysis to understand plant disease spread

Network analysis challenges conventional approaches of studies in plant disease epidemiology, especially underling the spatiotemporal flow of plants and their pathogens (Pautass et al., 2010). When plant pathological studies were restricted to a single geographical location, there was a limit to thinking about the connections or relationships between plants or fields in different locations. However, network analysis can enlarge the view of studies and can consider whether or not plant pathogens are moving from one field to others in the regions of interest. For example, networks of plant disease spread in trade networks presented the flows of plant disease from in-

fectured units (infected plants or epidemic areas) to susceptible units (susceptible plants or areas).

Network models of epidemic development

The idea of plant epidemics is that the probability of infection embedded in the connection or the contact patterns between susceptible/infected plants, and it forms as the networks. Pautasso et al. (2010) showed a network model of epidemic development (susceptible-infected-susceptible model) in a directed network. In the network model, vertices were represented plant, and their attributes were presented the infectious status (healthy or infected). The epidemic is started at a single node, then nodes with a connection from the starting infected node will be infected at the next time step with a certain probability of transmission. In turn, already infected nodes will be infected at the next time step depending on their infection status and on a certain probability of persistence. The probability of infection transmission is the same for all connections between infected nodes and susceptible nodes over times. Similarly, the probability of infection persistence is the same for infected nodes in a certain network replicate. For each network structure, the two probabilities of persistence and transmission define an epidemic threshold, which is independent of the starting node of the epidemic. This epidemiological model does not result in either susceptible or infected nodes, as nodes will have a infection status along a continuum. Key quantities for epidemiological dynamics in networks were reviewed in Moslonka-Lefebvre et al. (2011).

Analysis of plant trade network

Phytophthora ramorum epidemic networks in the horticultural trade are an example of the application of network models in the study of plant disease spread (Harwood et al., 2009; Pautass et al., 2010). Simulations of spread of *P. ramorum* in dif-

ferent network structures (random, small-world and scale-free network) were found that epidemic threshold, the boundary between a no epidemic an epidemic outcome, is significantly lower for scale-free network, a network is dominated by a small number of nodes with many connections, compared to local, random and small-world network structure. Modeling suggested that was possible to control an epidemic by changing the structure of network, without having to decrease the probability of infection persistence at a nursery site and/or of transmission between sites. Simulation showed that correlation coefficient between link in and out nodes increased with connectivity level for all the structures investigated and underlined the importance of targeted control towards node with more connections than others.

Regardless of the network structure and connectivity level, epidemic threshold is negatively correlated to the correlation coefficient between link in and out nodes, (Moslonka-Lefebvre et al., 2009). In presence of high-connected nodes, the most effective way to control disease spread is to move from two-way to a one-way network. That is move from network where overall there is positive correlation among links-in and -out to one where the correlation are negative. In practice this would mean that a nursery network would be dominated by major node, which receives plant materials from many production sites but supply relatively few retail sites, or by major nodes which received plant materials from a few production sites but supply many retail sites. However, the scenario where there are major nodes which both receive plant material from many production sites and supply many retail sites is the most problematic control. The better way to control should target towards the cluster (group of nodes), not only to high-connected nodes.

The last point, the modeling of disease spread in small-size directed networks

showed that increasing the proportion of wholesalers (*i.e.*, traders without a preponderance of incoming or outgoing links) tends to decrease the epidemic threshold in local, random, and small-world network. The opposite result is obtained for the proportions of produces and retails. Scale free networks appear instead to be tolerant to changes in these hierarchical categories as the epidemic threshold in this case is governed by the presence of hub rather than by the features of the majority of nodes in the network (Pautass et al., 2010).

Network models to design strategies of plant disease management

Due to globalization, plant trades among countries are increasing and convenient. They potentially caused risks to plant health when they were not under good control. To control the risks of plant disease epidemic through plant trade, network models were applied to present the flows of trade network of plants and plant products across the world and within countries and to develop strategies of plant disease management (Pautasso, 2015). Networks could be found hubs or highly connected nodes, which represented locations or countries where imported and exported plants or plant parts. Hubs or highly connected nodes were targets for control disease flows in the scenarios that network presented because they were considered to the likelihood of pathogens actually infecting along particular links. Strategies for disease management should be designed by focusing on links to and from hubs, nodes which have high degree of connectivity, so that it increase efficiency to control plant pathogen spreads. The strategies may aim to remove them from the network (Jeger et al., 2007; Moslonka-Lefebvre et al., 2011). Alternatively, strategies may pay attention on them in order to prevent disease spreads. Shaw and Pautasso (2014) suggested placing quarantine efforts on hubs or on connections between major hubs.

Additionally, Pautasso et al. (2008) showed the good examples, which are co-occurrence networks of the *P. ramorum* infected plant genera different environments. The networks may be helpful in identifying host taxa playing an important role in spreading a certain disease in the semi-natural environment, in crop plants, and plants in the trade. Combining genetic network analysis and data on trace forward and trace back on movement of plants nursery trade supported to identified confidentially *P. ramorum* migration. From this approach, it was clear that the pathogen was introduced originally from nurseries, which *P. ramorum* populations in nurseries are genetically ancestral to all Californian forest populations.

Using Network analysis to understand molecular plant pathology

For understanding mechanisms of plant-pathogen interactions, network analysis offers tools to visualize interactions of biological components including genes, proteins, and metabolites, which are related to plant-pathogen interactions. Network concepts enable us to characterize interplay of those components following the properties of network structure. Analyzing structural properties of networks may provide useful clues about the biological functions of individual genes, genes complexes, proteins, protein complexes, pathways they participate in.

Presenting biological data with network model

Networks are used in different contexts as ways to represent relationships between entities, such as interactions between genes, proteins or metabolites. Wu et al. (2007) gave the example of a network model built from gene-for-gene relationships between rice and various avirulence genes of the pathogen *Xanthomonas oryzae* pv. *oryzae* causing bacterial leaf blight of rice. Nodes represented isogenic lines of rice and weighted edges reflected the number of shared genes with high resistance (with respect

to avirulence genes) in the two isogenic lines of rice. For a plant breeder, this graph can help in identifying particularly promising genes for developing a disease resistant variety.

Network analysis to study biological systems

Network analyses have been applied to visualize the myriad information and analyze the complex relationships. To better understand the collective impact of genes on complex traits and determine what governs their organization, biologists are most likely to apply gene co-expression networks (Usadel et al., 2009). Co-expression networks most commonly use the Pearson's correlation coefficient to establish linear pairwise correlations between gene pairs in an adjacency matrix. Another associative metric that can be used is the Spearman correlation coefficient, which captures nonlinear correlations between genes to be uncovered (Usadel et al., 2009; Horvath, 2011). Once a co-expression network has been generated, identifying modules by clustering can help extract biological meaning from the network. Uncharacterized genes within such a module can be candidates for participating in the same process. Similarly, genes directly connected to (or co-expressed with) known central regulators of a developmental process are candidates within this co-functional framework.

Zheng and Zhao (2013) used co-expression network inference to investigate plant immunity system of citrus in response to *Candidatus Liberibacter asiaticus*. Gene coexpression networks based on Pearson's correlation coefficients using four transcriptomic data sets studies of citrus infected by *Candidatus Liberibacter asiaticus* bacterium were constructed. The result showed high degree of preservation of gene co-expression patterns across the networks based on different datasets. The network structures revealed hub genes, which have high connectivity. Highly connected nodes (genes) were

closely related to *Arabidopsis* SYP71 encoding a plant syntaxin which functions as a plasma membrane-associated protein transporter. Furthermore, in the study of Mukhtar et al. (2011), plant-pathogen interaction networks revealed the interactions of novel *Arabidopsis* protein-pathogen effectors, provided evidences that pathogen effectors target a limited number of host immune proteins, and demonstrated that effectors from very distantly related pathogens interact with the same host proteins.

The main use of co-expression networks with large collections of static expression data is gene discovery. However, many biologists have attempted to construct differential networks to measure differences in connectivity patterns from datasets with different conditions or targeted experiments (Toubiana et al., 2013). Lu et al. (2013) constructed networks of soil fungal communities. The fungal networks represented two different conditions; yield-invigorating and yield-debilitating soils under prolonged potato monoculture were compared. The authors discussed in differential network concepts that in healthy network three-eighths of fungal groups and soil organic matters were strongly correlated, and in diseases network two of four groups strongly correlated with soil electrical conductivity (EC) and ammonium nitrogen. Differential network analysis showed that average degrees of nodes belonging to *Sordariales* and *Hypocreales* in healthy network substantially were different in disease network. This indicated that they are key species of ecological communities.

A variety of Network Construction Methods

Until this section, I gave the general information of network analysis with network concepts, network modeling and inference. Ultimately, most of the systems studied from a network-based perspective are dynamic in nature. This underscores the importance of developing tools to understand the interplay between network structures and

dynamic processes. I reviewed three network-based analyses: single network, differential network and dynamic network analysis. These techniques from network science have yielded many biological insights (Ideker and Krogan, 2012).

Single Network Analysis

Single-network analysis will be used for analyzing a network based on a single data set. It aims to model patterns of relationships between entities to a network, thereby determining its topological properties such as node degree, degree distribution, average path length, clustering coefficient, modularity index (Newman, 2006) to describe the patterns in the network model. These structural properties offer the potential to explore relationships within complex data sets and identify key factors or clusters.

Correlation based networks are widely used for studying biological networks based on pair-wise correlations between variables. Their edges are obtained using correlation-based measures removing spurious relationships. Correlation based measures include linear based correlation (Pearson's correlation, partial correlation) and rank based correlation (Spearman's correlation, Kendall's correlation). Removal of spurious relationships is particularly important when one attempts to establish causal relationships between entities (Toubiana et al., 2013). A correlation based network allows one to define modules (clusters), intramodular hubs, and network nodes with regard to module membership, to study the relationships between modules, and to compare the network topology of different network.

Differential Network Analysis

Networks can respond differently under various environments or with external signals. They can be simplified by focusing on key components and capture only the essential components differently responding between environments in which they play

a key role in the modeled response (Pe'er and Hacohen, 2011). Networks are examined by adding or deleting some variables. This allows predicting interactions or components that change following the changed structure of networks.

Differential network analysis is applied to identify and describe differences between two networks under different conditions. Differential networks from different data sets with different conditions might display different interactions from the single network based on the whole data set, which are disregarded the conditions. The strongest interactions in differential networks are not necessarily those that are strong in the single network. Conversely, they may be weak or absent in single network. Compared between differential networks under different environment conditions, the differential interactions can be implied that they are a result of response to environmental conditions. Moreover, it can be implied that environments influence on the interaction between pairs of nodes contributing to the differential interactions.

The objective of differential network analysis is to identify the different interactions, that may reveal when condition has changed. For example, the differential network analysis of gene co-expression networks to understand the differences between human and chimpanzee brain was focused on how expression patterns of genes were different in human and chimpanzee brains. These networks are contrasted to find a) non-preserved modules (group of genes with sharing similar expression profiles and presumably similar functions), b) differentially occurred genes, and c) differentially connected genes. The result showed that

Dynamic Network

Biological systems are highly dynamic. They must continuously respond to external signals or the internal state of the system. The responses can be altered slowly

or quickly over time (Pe'er and Hachohen, 2011). Thus, realistically the corresponding biological network models must evolve as well. It seems clear that dynamic networks enable us to see and understand the systems and how their dynamic effects change over time. Some understandings previously have been obtained from studies of dynamics of large networks, for example, gene expression or metabolic fluxes network (Ideker and Krogan, 2012).

The main goal of dynamic network analysis moves away from single network analysis, which characterizes absolute properties of the system. It aims to concentrate on a specific dynamic response. Rather than answer what the key factors in the system, it answers what parts of the system are most affected by perturbation. Most commonly, dynamic networks are applied when the edges among a set of vertices or the sets of vertices itself are changing as a function of time. For example,

Summary

This literature review presented a brief introduction of network analysis and concise concepts and methods. Briefly, a network is usually represented by sets of nodes (or vertices) connected by edges (or links) in various ways. Networks can be categorized to four types, social network, information, technology network, and biological network. Even though four types of networks are described and applied in different context, they share a common empirical focus on relational structure and a similar set of mathematical analyses. Network models are capable of presenting unique values, which traditional approaches cannot present. Network analysis was discussed as applied to two broad areas of study of plant diseases. It firstly was applied to study plant disease epidemics. For example, networks of *P. ramorum* spread through plant nurseries trade. The results enabled us to understand the directions and processes of disease spread.

Additionally, they could predict the movement of disease flows, and improve the implementation of plant disease policy. Secondly, molecular plant pathology showed two applications of network applications. The first use is to apply networks to model large and complex biological datasets. Another use is the consideration network structure to understand biological system. Emergent properties of network structure influences may be identified, measured and analyzed to yield better explanations of the experiments being observed. While numbers of documented plant pathological studies using network analysis are sparse, the literature presented in this review showed a clear and compelling case for plant pathologists to expand the understanding of and use network concepts and methods.

Biological networks are context-specific and dynamic in nature. Under different conditions, the topology of the network changes. Network approaches (differential network, dynamic network analysis) are successful to yield insight on biological system under different conditions or dynamic processes. Network analysis concepts and methods augment existing approaches and provide tools for exploring complex relationships, which have been widely acknowledged as influential but difficult to measure using traditional methods.

LITERATURE CITED

- Barabási, A.-L. and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 5(2):101–113.
- Barnwal, M., Kotasthane, A., Magculia, N., Mukherjee, P., Savary, S., Sharma, A., Singh, H., Singh, U., Sparks, A., Variar, M., et al. (2013). A review on crop losses, epidemiology and disease management of rice brown spot to identify research priorities and knowledge gaps. *European Journal of Plant Pathology*, 136(3):443–457.
- Borgatti, S. P., Everett, M. G., and Johnson, J. C. (2013). Analyzing social networks.
- Castells, M. (2001). *The rise of network society*. Oxford: Blackwell Publishers.
- Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012). Molecular ecological network analyses. *BMC Bioinformatics*, 13:113.
- Dietz, K. J., Jacquot, J. P., and Harris, G. (2010). Hubs and bottlenecks in plant molecular signalling networks. *New Phytologist*, 188(4):919–938.
- Duwayri, M., Van Tran, D., et al. (2000). Reflections on yield gaps in rice production: how to narrow the gaps. *RAP Publication (FAO)*.
- Eagling, D. (2007). Australian trade in agricultural food products – the challenge for plant pathologists. *Australasian Plant Pathology*, 36(6):539–542.
- Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., and Huttenhower, C. (2012). Microbial co-occurrence relationships in the human microbiome. *PLoS Computational Biology*, 8(7):e1002606–e1002606.
- Freilich, S., Kreimer, A., Meilijson, I., Gophna, U., Sharan, R., and Ruppín, E. (2010). The large-scale organization of the bacterial network of ecological co-occurrence interactions. *Nucleic Acids Research*, 38(12):3857–3868.
- GRiSP (2013). *Rice almanac*. International Rice Research Institute (IRRI), 4 edition.
- Guimera, R. and Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature*, 433(7028):895–900.
- Harwood, T. D., Xu, X., Pautasso, M., Jeger, M. J., and Shaw, M. W. (2009). Epidemiological risk assessment using linked network and grid based modelling: *Phytophthora ramorum* and *Phytophthora kernoviae* in the UK. *Ecological Modelling*, 220(23):3353–3361.
- Horvath, S. (2011). Weighted network analysis: applications in genomics and systems biology.

- Ideker, T. and Krogan, N. J. (2012). Differential network biology. *Molecular Systems Biology*, 8(1):565.
- Jeger, M. J., Pautasso, M., Holdenrieder, O., and Shaw, M. W. (2007). Modelling disease spread and control in networks: implications for plant sciences. *New Phytologist*, 174(2):279–297.
- Kasari, C., Locke, J., Gulsrud, A., and Rotheram-Fuller, E. (2011). Social networks and friendships at school: comparing children with and without ASD. *Journal of Autism and Developmental Disorders*, 41(5):533–544.
- Krause, A. E., Frank, K. A., Mason, D. M., Ulanowicz, R. E., and Taylor, W. W. (2003). Compartments revealed in food-web structure. *Nature*, 426(6964):282–285.
- Lazega, E. and Pattison, P. E. (1999). Multiplexity, generalized exchange and cooperation in organizations: a case study. *Social Networks*, 21(1):67–90.
- Lu, L., Yin, S., Liu, X., Zhang, W., Gu, T., Shen, Q., and Qiu, H. (2013). Fungal networks in yield-invigorating and -debilitating soils induced by prolonged potato monoculture. *Soil Biology and Biochemistry*, 65:186–194.
- Mew, T. W., Leung, H., Savary, S., Vera Cruz, C. M., and Leach, J. E. (2004). Looking ahead in rice disease research and management. *Critical Reviews in Plant Sciences*, 23(2):103–127.
- Moody, J. (2001). Race, school integration, and friendship segregation in America. *American Journal of Sociology*, 107(3):679–716.
- Moslonka-Lefebvre, M., Finley, A., Dorigatti, I., Dehnen-Schmutz, K., Harwood, T., Jeger, M. J., Xu, X., Holdenrieder, O., and Pautasso, M. (2011). Networks in plant epidemiology: from genes to landscapes, countries, and continents. *Phytopathology*, 101(4):392–403.
- Moslonka-Lefebvre, M., Mathieu Pautasso, M., and Jeger, M. J. (2009). Disease spread in small-size directed networks: epidemic threshold, correlation between links to and from nodes, and clustering. *Journal of Theoretical Biology*, 206:402–411.
- Mukhtar, M. S., Carvunis, A.-R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., Tasan, M., Galli, M., Hao, T., Nishimura, M. T., Pevzner, S. J., Donovan, S. E., Ghamsari, L., Santhanam, B., Romero, V., Poulin, M. M., Gebreab, F., Gutierrez, B. J., Tam, S., Monachello, D., Boxem, M., Harbort, C. J., McDonald, N., Gai, L., Chen, H., He, Y., Vandenhaute, J., Roth, F. P., Hill, D. E., Ecker, J. R., Vidal, M., Beynon, J., Braun, P., and Dangl, J. L. (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science*, 333(6042):596–601.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM review*, 45(2):167–256.

- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582.
- Ou, S. H. (1985). *Rice diseases*. International Rice Research Institute (IRRI).
- Pastor-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200.
- Pautass, M., Moslonka-Lefebvre, M., and Jeger, M. J. (2010). The number of links to and from the starting node as a predictor of epidemic size in small-size directed networks. *Ecological Complexity*, 7:424–432.
- Pautasso, M. (2015). Network simulations to study seed exchange for agrobiodiversity conservation. *Agronomy for Sustainable Development*, 35(1):145–150.
- Pautasso, M., Harwood, T., Shaw, M., Xu, X., and Jeger, M. (2008). Epidemiological modeling of *Phytophthora ramorum*: network properties of susceptible plant genera movements in the nursery sector of England and Wales. In *Proceedings of the Sudden Oak Death Third Science Symposium*.
- Pautasso, M., Moslonka-Lefebvre, M., and Jeger, M. J. (2010). The number of links to and from the starting node as a predictor of epidemic size in small-size directed networks. *Ecological Complexity*, 7(4):424–432.
- Pe’er, D. and Hachohen, N. (2011). Principles and strategies for developing network models in cancer. *Cell*, 144(6):864–873.
- Proulx, S., Promislow, D., and Phillips, P. (2005). Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, 20(6):345–353.
- Savary, S., Horgan, F., Willocquet, L., and Heong, K. L. (2012). A review of principles for sustainable pest management in rice. *Crop Protection*, 32:54–63.
- Savary, S., Willocquet, L., Elazegui, F. A., Teng, P. S., Van Du, P., Zhu, D., Tang, Q., Huang, S., Lin, X., and Singh, H. M. (2000). Rice pest constraints in tropical Asia: characterization of injury profiles in relation to production situations. *Plant Disease*, 84(3):341–356.
- Shaw, M. W. and Pautasso, M. (2014). Networks and plant disease management: concepts and applications. *Annual Review of Phytopathology*, 52(1):477–493.
- Toubiana, D., Fernie, A. R., Nikoloski, Z., and Fait, A. (2013). Network analysis: tackling complex data to study plant metabolism. *Trends in Biotechnology*, 31(1):29–36.
- Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F. M., Bassel, G. W., Tanimoto, M., Chow, A., Steinhauser, D., Persson, S., and Provart, N. J. (2009). Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. *Plant, Cell & Environment*, 32(12):1633–1651.

- Willoquet, L., Elazegui, F. A., Castilla, N., Fernandez, L., Fischer, K. S., Peng, S., Teng, P. S., Srivastava, R., Singh, H., Zhu, D., et al. (2004). Research priorities for rice pest management in tropical asia: a simulation analysis of yield losses and management efficiencies. *Phytopathology*, 94(7):672–682.
- Windram, O., Penfold, C. A., and Denby, K. J. (2014). Network modeling to understand plant immunity. *Annual Review of Phytopathology*, 52(1):93–111.
- Wu, X. M., Li, Y. R., Zou, L. F., and Chen, G. Y. (2007). Gene–for–gene relationships between rice and diverse avrbs3/ptha avirulence genes in *Xanthomonas oryzae* pv.*oryzae*. *Plant Pathology*, 56(1):26–34.
- Yang, Y., Han, L., Yuan, Y., Li, J., Hei, N., and Liang, H. (2014). Gene co–expression network analysis reveals common system-level properties of prognostic genes across cancer types. *Nature Communications*, 5.
- Zadoks, J. C. and Schein, R. D. (1979). *Epidemiology and plant disease management*. Oxford University Press, USA.
- Zheng, Z. L. and Zhao, Y. (2013). Transcriptome comparison and gene coexpression network analysis provide a systems view of citrus response to “*Candidatus liberibacter asiaticus*” infection. *BMC Genomics*, 14:27.