bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Analyzing Key Nodes and Epidemic Risk in Seed Networks: Sweetpotato in Northern Uganda

Kelsey F. Andersen[1,2,3], Chris E. Buddenhagen[1,2,3], Paul Rachkara[4], Richard Gibson[5], Stephen Kalule[4], David Phillips[5], Karen A. Garrett[1,2,3]

[1]Plant Pathology Department, University of Florida, Gainesville, FL 32611-0680, United States

[2]Institute for Sustainable Food Systems, University of Florida, Gainesville, FL 32611-0680, United States

[3]Emerging Pathogens Institute, University of Florida, Gainesville, FL 32611-0680, United States

[4]Department of Rural Development and Agribusiness, Gulu University, Gulu, Uganda

[5]Natural Resource Institute, University of Greenwich, United Kingdom

Corresponding authors: KA Garrett (karengarrett@ufl.edu) and KF Andersen (andersenk@ufl.edu)

Mailing address: Plant Pathology Department, University of Florida, Gainesville, FL 32611-0680, United States

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

The structure of seed system networks provides important information about epidemic risk within the network. We evaluated the structure of a sweetpotato seed system in Northern Uganda in terms of its utility for distributing improved varieties and its vulnerability to the spread of potential seed-borne pathogens. Sweetpotato sellers were surveyed in the Gulu Region of Northern Uganda. Weekly vine sales transactions were tracked through the growing season (April- October) creating a robust dataset of planting material sales over time, including price, village sold to, volume, and information about the buyer and seller. From this dataset of known transactions and the distance between villages, a network of vine movement was constructed. *In silico* simulations of the introduction of a novel virus into the systems indicated the potential for rapid spread. Through simulation of multiple epidemic starting points, nodes of particular importance to disease sampling and mitigation were identified. This method can serve as an example, with potential to be used across a wide variety of seed systems.

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Seed systems, both formal and informal, are a critical component of global food security. Yet efforts to implement seed systems that work better for smallholder farmers in low-income countries have too often been unsuccessful (Gibson and Kreuze 2015; Thomas-Sharma et al. 2016). Improving seed security – defined as timely access to quality planting material by all, at a fair price (Sperling 2008, Gibson 2011) – is vital for improved livelihoods, particularly for smallholder farmers. Seed systems are complex, highly nuanced networks with a suite of actors that move both material and information. Plant disease epidemiologists have given little attention to the study of seed system networks, although seed systems play a fundamental role in the spread of plant disease epidemics. Network analysis can be applied to study layered biophysical and information networks (Garrett 2012b, 2017; Pautasso and Jeger 2008). Agricultural cropping systems are inherently multi-layer networks of information and biological material. Network analysis can be applied to define key nodes and actors in a system, provide diagnostic metrics, and forecast the risk of network fragmentation and pathogen introduction (Garrett 2012b; Harwood et al. 2009; Moslonka-Lefebvre et al. 2011; Pautasso and Jeger 2008; Sanatkar et al. 2015). Understanding the structure of seed systems can inform policy and intervention strategies, where intervention may be particularly important in times of acute insecurity due to weather or biotic stressors. Decisions made in times of crisis can have long-lasting impacts on communities (Sperling 2008). Healthy local seed systems have basic characteristics: access to multiple, well adapted varieties that meet the local product requirements, at a fair price (Sperling 2008). Having good models of seed systems supports scenario analyses to determine how likely a seed system is to remain healthy under different types of stressors.

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

This study examines sweetpotato vine transactions in Northern Uganda, illustrating the potential for seed system network analysis to define a system and assess how it compares to other seed systems. Sweetpotato is a major staple food crop in many African countries, and Uganda is the second largest producer in Africa, fourth globally (FAOSTAT 2013). Sweetpotato is generally grown by women in Uganda in small plots of land, close to the household, and is important for household food security (Behrman 2011; Johnson and Gurr 2016). In the last decade, sweetpotato has increased in importance due to the introduction of a betacarotene-rich biofortified crop known as Orange-fleshed Sweetpotato (OFSP) by HarvestPlus, part of the CGIAR Research Program on Agriculture for Nutrition and Health. OFSP varieties were introduced with the goal of addressing vitamin A deficiency in women and children in this region (Behrman 2011).

Network analyses of seed systems have focused on understanding the effects of social ties on seed system structure, and how well networks may conserve variety diversity (Pautasso et al. 2013). Abay et. al (2011) applied network analysis in a study designed to characterize informal barley seed flows in the Tigray, Ethiopia region, with the goal of informing breeding or intervention strategies. The authors used network properties, such as betweenness and degree centrality, to characterize key nodes and their role in connecting the seed system. Each of these metrics measures the importance of a node in the network in terms of the number of connections it has, or the number of paths across the network of which it is a part. Network analysis can also be used to evaluate nodes important for sampling and mitigation of the movement of pathogens or other contaminants through networks (Sutrave et al. 2012). In a study of wheat grain movement in the United States and Australia, network analysis was used to identify key locations that could be strategically targeted for sampling and management of mycotoxins

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Hernandez Nopsa et al. 2015). Network modeling is increasingly being used to evaluate likely disease spread, and can provide insights into the utility of forecasting and backcasting (Sanatkar et al. 2015). To our knowledge, the current study is the first to model the epidemiology of a novel epidemic introduction in a seed system.

Viruses and other seed-transmitted diseases are important risks to yield and quality degeneration within seed systems, particularly informal seed systems and without healthy seed certification protocols. Newly introduced viruses can be particularly severe, as methods for detection may be limited or unavailable, and resistance may be unavailable for several years. It is imperative, therefore, that the risk of novel pathogen introduction into a seed system be understood for the swift recommendation of intervention (such as sampling, quarantine, variety deployment, and education). This problem was illustrated in 2011 when maize lethal necrosis (MLN) was first reported in Kenya (Wangai et al. 2012) and soon was found in several Sub-Saharan African countries. MLN symptom expression results from coinfection with *Maize chlorotic mottle virus* (MCMV) and a potyvirus (Mahuku et al. 2015). Since its introduction, MLN has been detected in several East African countries including Ethiopia, Uganda, South Sudan, Tanzania, DRC, and Rwanda, and up to a 22% yearly yield loss has been reported (De Groote et al. 2016; Mahuku et al. 2015). Characterization of seed systems using network analysis may inform strategic intervention, such as quarantine and sampling, in these scenarios. Similarly, in 2015 *Cassava mosaic virus* (CMV) was first detected in Southeast Asia and efforts are still ongoing to mitigate spread and deploy resistance (Graziosi et al. 2016; Wang et al. 2016).

Viral diseases are major biotic limiters to sweetpotato production in Uganda and throughout Sub-Saharan Africa, with the most yield-limiting being sweet potato virus disease

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(SPVD), which occurs when a plant is co-infected with *Sweet potato feathery mottle virus* (SPFMV) and *Sweet potato chlorotic stunt virus* (SPCSV) (Karyeija et al. 1998). Seed degeneration is defined as the successive loss in yield over generations of vegetatively-propagated seed material due to the accumulation of viruses and other seed borne pathogens (Thomas-Sharma et al. 2016). Degeneration is highly problematic in informal seed systems where farmers tend to save seed season-to-season and where certified seed sources are rare or non-existent. Both SPFMV and SPCSV can be transmitted through vegetatively-propagated material, with evidence of yield degeneration over five generations in high pressure fields in Uganda (Adikini et al. 2015). SPVD has not yet been reported in Northern Uganda, likely because the extended dry season in this region is unfavorable for the whitefly vector (Richard Gibson, *personal observation*). Changes in climate patterns or vector range, however, could potentially extend the range of this disease into this region. The potential for novel pathogen introduction makes it important to model epidemic scenarios to inform intervention strategies.

Understanding key properties that make a seed system successful can help inform seed system development and strategy. Seed system networks are comprised of a suite of actors, or nodes, including farmers, buyers, multipliers, NGOs, breeding organizations, traders, villages, and buyers. The connections between these nodes, generated through formal and informal interactions (links), are complex and require analyses that address this complexity. Network analysis allows us to simulate multiple scenarios in known systems, including the impact of node loss or epidemic spread. In this study we propose a framework for analyzing such networks that can be transferred to a broad range of seed systems, including informal systems. In this analysis we aim to; i) characterize key network and node properties within an informal seed system important to regional food security; ii) evaluate the variety dissemination within the network; iii)

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

evaluate scenarios for the introduction of a potential seed-borne pathogen into the system, and determine the optimal nodes for sampling and intervention.

Methods

Study System

In Northern Uganda, sweetpotato seed material is sold in small bundles of vine cuttings. In this region, sweetpotato vine distribution is largely informal, consisting of smallholder farmers who have access to fields with adequate moisture to produce roots and vines through the extended dry season, which typically lasts from December to April (Gibson 2013). These counter-season multipliers generally produce local landraces which tend to be well-adapted white-fleshed cultivars. (Gibson 2013). Vine cuttings are not easily stored, and because of a single, extended dry season in Northern Uganda, vines need to be obtained by farmers at the beginning of each season (Gibson et al. 2011). There are also several formal institutions involved in sweetpotato breeding and distribution in Uganda, which include the National Sweetpotato Program (NSP), private sector enterprises, and NGOs (Gibson 2013).

Survey Methods

A survey of vine multipliers and sellers was conducted in 2013-2014 in the Gulu Region of Northern Uganda. Survey methods are fully described by Rachkara et al. 2017, *in press.* In the first year of the study (2013), the transactions of a small number of local multipliers were tracked throughout the season. A more complete cohort of multipliers and sellers were surveyed in 2014. All seller names have been anonymized to protect the identity of individuals who participated in this survey. Each seller was visited weekly from the start of the growing season (April) and surveyed twice per week to record all transactions that occurred in the period since they were last

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

visited until the end of the season (August). Volume of transaction (number of bundles), price, variety, origin of buyer, and buyer type (farmer or seller) were recorded. In this study, a small bundle refers to 50 vines cut to 40 cm in length. Large bundles are equal to 20 small bundles. For consistency, large bundles will be used to describe volume in this paper. Because of the high volume of transactions, names of individual buyers were not collected and therefore sales transactions were summarized by the village from which the buyer originated.

Seed Network Data Analysis

Nodes in this analysis include sellers and villages, with one set of directed links representing vine sales from an individual single seller to an individual village. Villages in this region of Uganda can have between 40-60 households. The categorization of the network into two distinct types (village and seller), produces a single, bipartite graph. Although several transactions were recorded on a weekly basis, transactions were generally aggregated so that links represent the existence of at least one transaction over the course of the season. This part of the network is based entirely on the data from (Rachkara et al. 2017).

Key network properties, such as number of nodes, network density, and modularity, were calculated for each year of the survey. Key node properties, such as coreness, closeness, and in- and out-degree, were measured for both villages and sellers. These metrics provide insights, beyond those that can be gained from simple summary statistics, into the role of a node for both variety distribution and access, as well as vulnerability to pathogen introduction. All analyses were conducted in the R programing environment (R Core Team 2016) using several software packages, including igraph (Csárdi and Nepusz 2006).

Modeling Disease Risk in Simulation Experiments

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The introduction of novel pathogens can pose a high risk to seed systems, especially informal systems. It is important to understand the potential for epidemic spread within a known network structure. In this analysis, we use both actual 2014 transaction history and spatial nearness to simulate the introduction of a pathogen into villages.

In order to better understand epidemic risk in the seed system, a second set of undirected links were added in addition to the known transaction network. These represent the potential for movement of virus vectors, as well as the potential for informal exchange of potentially infected planting materials, between nearby villages. The existence of a link between two villages was evaluated as a function of the distance between them. The likelihood of transmission between villages was modeled as a function of the Euclidian distance between each pair of villages in the network. The transmission probability from one village to another follows a power-law distribution with greater risk for pairs nearer each other. A link is intended to represent both the potential for seed movement and vector movement, and it is assumed that there is a higher chance of exchanging seed with nearer neighbors than with villages that are far away. The power law equation used is $Y=AX^{-\beta}$, where X = Euclidean distance between two villages and Y is proportional to the probability of movement between the villages. When A=1 and β =1.5, this results in 27% of village pairs being linked.

The first simulation experiment evaluated the potential spread through the network of a pathogen associated with planting material of a single variety. At the start of the simulation (Time 1), each seller known to sell this variety had a 5% probability of transmitting the virus to villages to which transaction connections had been made in the 2014 season. Villages that became infected after a single round of simulations were assigned an infected status in the subsequent time step (Time 2). It is not only important to understand what transmission might

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

happen within the network, but also to understand the chance that neighbors may become infected. Transmission in the next time step was a function of nearness of infected villages to neighboring villages. That is, the closer the neighboring village is, the higher the potential to become infected in the subsequent time step (Time 3).

In the second simulation experiment, we assess the value of each village as a monitoring or sampling location. Again, links from sellers to villages are based on known transactions in 2014 and the probability of a link existing between villages is calculated based on a power law function of distance, with a higher probability of a link between villages that are geographically close. In this scenario, all nodes (sellers and villages) are assigned an equal likelihood of being the starting point for an introduced pathogen. For each possible combination of an epidemic starting node and a sampling node, we determine the number of nodes that could become infected by the time the pathogen is detected at the sampling node. Summarizing over all the potential starting nodes allows a comparison of the importance of each node as a potential sampling node. Simulations were implemented using custom R code.

Results

Network Properties

In 2013 and 2014, 5 and 27 sellers were tracked, respectively. A total of 878 transactions that occurred in this season were recorded, to 99 distinct villages (Table 1). Using an adjacency matrix constructed of aggregated transactions from sellers to villages, a network graph was constructed (Figure 1). It is important to note here that although transactions were collected over time, the presence of a directed link here represents at least one transaction in the 2014 season. This graph has a density of 0.013 and a modularity of 0.56 (Table 1). Because density is a

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

function of the proportion of possible connections within a network, it makes sense that this would be low for the graph, as there are no connections between village or sellers, therefore many potential connections are not realized. Such a high modularity is likely a function of the high intrinsic community structure in this network.

Node in- and out-degree are the number of links to and from a node, indicating how well connected individual agents are in a network. The degree distributions of this network are highly right skewed (Figure 2, *b*, *d*), meaning that the majority of villages and sellers have a small number of connections, while a small number of sellers have many connections. The network graphs in Figure 2 have node size that is proportional to node in- (a) and out-degree (b). These metrics are particularly important for characterizing the risk of disease spread, as those sellers and villages with high node degree have a higher potential to transmit or contract disease.

Variety Dissemination

A total of 15 cultivars were sold during the 2014 season, being a mix of landraces and varieties introduced by the national breeding program (Figure 3). Six of these cultivars were considered biofortified-OFSP cultivars. Interestingly, for the top cultivars, there appears to be a disproportionate number of transactions for the volume that was sold (Figure 3). For example, the white-fleshed land race, Ladwe Aryo, is sold in hundreds of transactions throughout the season, but in less volume than the OFSP cultivar, Ejumula. This suggests that a large volume of certain varieties (particularly those that are orange-fleshed) are sold to a small number of farmers. When the network is examined in terms of variety, disaggregation becomes apparent (Figure 4). Evaluating graphs of the distribution of the top eight varieties (Figure 4) indicates that only a small number of sellers and villages are exchanging orange-fleshed varieties. It

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

appears that most villages only buy a single variety, even when they have access to multiple sellers.

Disease Risk in Simulation Experiments

The potential emergence of a novel virus, or an increase in the incidence and severity of a known virus, pose a threat to seed systems, especially where material is propagated vegetatively. In this scenario analysis, virus spread was modeled by introducing infected material of a single variety into the known transmission network. In Time 1 (Figure 5a) 21 sellers have a 5% potential of transmitting infected material to known villages. In the set of simulations shown here, six villages obtain infected vines (Figure 5b). It is common practice in this informal seed system for vines to be shared between friends, family, and neighbors. Because of this known network property, it is assumed that villages that are nearest to infected villages will have the highest chance of receiving planting material that has been infected, such that they also become infected. In the third time-step presented here, 16 new villages acquire the pathogen, based on proximity. By Time 3, 22% of villages now have infected planting material and the capacity to spread infected planting material (Figure 5c). In this analysis of the specific structure of this sweetpotato seed network, we identify the nodes that are particularly important for mitigating disease spread. We identify the 20 nodes that would be most favorable for sampling for virus introduction as the fewest number of villages will be infected by the time the virus reaches these nodes (Figure 6).

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Discussion

Seed system assessment should be done to determine the normal status of a system, thus enabling strategic intervention in times of need. Understanding key system properties prior to emergencies is essential for quality development aid or plans for recovery from a new epidemic. The analytic framework proposed in this study can be used as a template to characterize and model other seed systems, particularly those of vegetatively-propagated crops with high potential for seed-borne pathogen introduction. It is clear from this study that sweetpotato vines in Northern Uganda are sold in a complex and highly connected informal seed system. There exists a trade-off of centrality within this seed network. That is, it is favorable to be a node with high in-degree and centrality because of increased access to a diversity of vines and sellers. However, these metrics may make a village or individual more susceptible to pathogen introduction. In one scenario analysis presented here, virus transmission within an informal seed system can be rapid, with 22% of villages becoming infected within a season. Disease transmission in this system is a function of human movement of planting material and vector transmission of viruses.

This study also used scenario analysis to characterize villages for their utility as sampling hubs based on a simulation where all nodes had an equal chance of being the point of epidemic start. Based on this analysis, a subset of 20 nodes were identified as potential monitoring targets because there is a high likelihood that the virus will be detected in these locations before the rest of the network becomes colonized. This method can be used to identify sentinel nodes to prescribe sampling efforts, and can be particularly useful in informal systems where production and distribution are not centralized, and therefore are constructed due to a number of social and economic factors. This method can serve as a complement to new diagnostic technologies, such as loop-mediated amplification (LAMP) assays, which are becoming increasingly available to

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

practitioners in the field and have the potential for rapid on-site detection of viruses and other pathogens (Johnson and Gurr 2016; Sasaya 2015).

The accumulation of virus is a major cause of yield and quality loss in sweetpotato in Uganda and throughout Sub-Saharan Africa (Clark et al. 2012; Gibson and Kreuze 2015) . There may be additional viruses that serve as hidden yield-robbers, however, as viral symptoms can be cryptic or easily confused with other biotic or abiotic stressors (Mukasa et al. 2003). To better understand the distribution of sweet potato viruses in this region, projects have recently been underway to sequence the pan-African sweetpotato virome (Kreuze 2014). This large-scale sampling and sequencing effort gives insight into the total number and abundance of viruses, both known and novel, across Sub-Saharan Africa. The study of the virome is an exciting new frontier at the intersection of plant disease epidemiology and seed system assessment and has the potential to give insight into the major drivers of yield loss in this region (Johnson and Gurr 2016). Sweet potato virus epidemiology, and specifically the transmission of SPFMV, is not only influenced by crop density and vector transmission potential, but also the abundance of alternative hosts (Tugume et al. 2016; Tugume et al. 2008). The influence of alternative host species on disease loss modeling deserves further attention.

The value of variety adoption is an idea that can spread through the system. Variety adoption in this system appears to follow a “rich get richer” phenomena, meaning that most nodes have a small number of connections and a small number of nodes have a high number of connections. This type of social network pattern can be fit by the power-law distribution, a distribution commonly used to describe pathogen dispersal over a landscape (Barabasi and Albert 1999; Mundt et al. 2009). When examining vine distribution by variety in this network, it is clear that some varieties are not well disseminated throughout the network. Based on the observed data, we

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

cannot be sure if this is because of preference or availability of planting material, or a combination of the two. The reason behind this deserves further attention because the adoption of OFSP is closely tied with human nutritional benefits. Similar methods to those described here may be utilized to target villages for development projects that aim to disseminate varieties to key hubs and maximize their distribution.

Understanding the network structure of seed systems provides a unique lens for understanding variety distribution and pathogen risk. Results from these studies can be strategically utilized to deploy sampling efforts and to disseminate new varieties in informal seed systems. With additional data available, there is the potential to analyze these types of seed transaction networks as components of layered information and biophysical networks (Garrett 2012a, 2017). Future surveys that include questions about social ties and the movement of information among farmers would support better models of variety adoption and distribution in this system. Next research steps will include more finely parameterizing transmission patterns, including the impact of variety resistance, and modeling the gain and loss of key nodes resulting from the introduction of intervention strategies. There is the potential to include data about known yield degeneration rates and known environmental conditions to predict regional yield loss in the case of pathogen introduction. Understanding these system components supports better strategies for seed system development.

Acknowledgements

This work received institutional review board approval through the University of Florida IRB-02 #IRB201700024. We confirm farmer participation was voluntary and personal and demographic information was protected. This research was undertaken as part of, and funded by, the CGIAR

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Research Program on Roots, Tubers and Bananas (RTB) and supported by CGIAR Fund Donors http://www.cgiar.org/about-us/governing-2010-june-2016/cgiar-fund/fund-donors-2/, The Bill and Melinda Gates Foundation (OPP1080975), US NSF Grant EF-0525712 as part of the joint NSF-NIH Ecology of Infectious Disease program, US NSF Grant DEB-0516046, and the University of Florida. The authors confirm they have no conflicts of interest.

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

Abay, F., de Boef, W., and Bjørnstad, Å. 2011. Network analysis of barley seed flows in Tigray, Ethiopia: supporting the design of strategies that contribute to on-farm management of plant genetic resources. Plant Genetic Resources 9:495-505.

Adikini, S., Mukasa, S. B., Mwanga, R. O. M., and Gibson, R. W. 2015. Sweet potato cultivar degeneration rate under high and low sweet potato virus disease pressure zones in Uganda. Canadian Journal of Plant Pathology 37:136-147.

Barabasi, A.-L., and Albert, R. 1999. Emergence of Scaling in Random Networks. Science 286:509-512.

Behrman, J. 2011. HarvestPlus Reaching End Users (REU) Orange-Fleshed Sweet Potato (OFSP) Project: Report of Qualitive Findings from Uganda. International Food Policy Research Institute 1-29.

Clark, C. A., Davis, J. A., Abad, J. A., Cuellar, W., Fuentes, S., Kreuze, J., Gibson, R. W., Mukasa, S. B., Tugume, A. K., Tairo, F. D., and Valkonen, J. P. T. 2012. Sweetpotato viruses: 15 years of progress on understanding and managing complex diseases. Plant Disease 96:168-185.

Csárdi, G., and Nepusz, T. 2006. The igraph software package for complex network research. InterJournal, Complex Systems 1695:1-9.

De Groote, H., Oloo, F., Tongruksawattana, S., and Das, B. 2016. Community-survey based assessment of the geographic distribution and impact of maize lethal necrosis (MLN) disease in Kenya. Crop Protection 82:30-35.

Garrett, K. A. 2012a. Information networks for plant disease: Commonalities in human management networks and within-plant signaling networks. Eur. J. Plant Pathol. 133:75-88.

Garrett, K. A. 2012b. Information networks for disease: commonalities in human management networks and within-host signalling networks. European Journal of Plant Pathology 133:75-88.

Garrett, K. A. 2017. Impact Network Analysis: a framework for evaluating the effects of information and other technologies through linked socioeconomic and biophysical networks. bioRxiv.

Gibson, R. 2013. How sweet potato varieties are distributed in Uganda: actors, constraints and opportunities. Food Security 5:781-791.

Gibson, R. W., and Kreuze, J. F. 2015. Degeneration in sweetpotato due to viruses, virus-cleaned planting material and reversion: a review. Plant Pathology 64:1-15.

Gibson, R. W., Namanda, S., and Sindi, K. 2011. Sweetpotato seed systems in East Africa. African Crop Science Conference Proceedings 10:449-451.

Graziosi, I., Minato, N., Alvarez, E., Ngo, D. T., Hoat, T. X., Aye, T. M., Pardo, J. M., Wongtiem, P., and Wyckhuys, K. A. 2016. Emerging pests and diseases of South-east Asian cassava: a comprehensive evaluation of geographic priorities, management options and research needs. Pest Management Science 72:1071-1089.

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 Modeling 220:3353-3361.

Hernandez Nopsa, J. F., Daglish, G. J., Hagstrum, D. W., Leslie, J. F., Phillips, T. W., Scoglio, C., Thomas-Sharma, S., Walter, G. H., and Garrett, K. A. 2015. Ecological networks in

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

stored grain: Key postharvest nodes for emerging pests, pathogens, and mycotoxins. BioScience 65:985-1002.

Johnson, A. C., and Gurr, G. M. 2016. Invertebrate pests and diseases of sweetpotato (Ipomoea batatas): a review and identification of research priorities for smallholder production. Annals of Applied Biology 168:291-320.

Karyeija, R. F., Gibson, R. W., and Valkonen, J. P. T. 1998. The significance of sweet potato feathery mottle virus in subsistence sweet potato production in Africa. Plant Disease 82:4-15.

Kreuze, J. 2014. siRNA deep sequencing and assembly: piecing together viral infections. In Gullino and Bonants, editors. Detection and Diagnostics of Plant Pathogens. Springer.

Mahuku, G., Lockhart, B. E., Wanjala, B., Jones, M. W., Kimunye, J. N., Stewart, L. R., Cassone, B. J., Sevgan, S., Nyasani, J. O., Kusia, E., Kumar, P. L., Niblett, C. L., Kiggundu, A., Asea, G., Pappu, H. R., Wangai, A., Prasanna, B. M., and Redinbaugh, M. G. 2015. Maize Lethal Necrosis (MLN), an emerging threat to maize-based food security in Sub-Saharan Africa. Phytopathology 105:956-965.

Moslonka-Lefebvre, M., Ann Finley, I. D., 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 101:392-403.

Mukasa, S. B., Rubaihayo, P. R., and Valkonen, J. P. T. 2003. Incidence of viruses and viruslike diseases of sweetpotato in Uganda Plant Disease 87:329-335.

Mundt, C. C., Sackett, K. E., Wallace, L. D., Cowger, C., and Dudley, J. P. 2009. Long-distance dispersal and accelerating waves of disease: empirical relationships. Am Nat 173:456-466.

Pautasso, M., and Jeger, M. J. 2008. Epidemic threshold and network structure: The interplay of probability of transmission and of persistence in small-size directed networks. Ecological Complexity 5:1-8.

Pautasso, M., Aistara, G., Barnaud, A., Caillon, S., Clouvel, P., Coomes, O. T., Deletre, M., Demeulenaere, E., De Santis, P., Doring, T., Eloy, L., Emperaire, L., Garine, E., Goldringer, I., Jarvis, D., Joly, H. I., Leclerc, C., Louafi, S., Martin, P., Massol, F., McGuire, S., McKey, D., Padoch, C., Soler, C., Thomas, M., and Tramontini, S. 2013. Seed exchange networks for agrobiodiversity conservation. A review. Agron. Sustain. Dev. 33:151-175.

Rachkara, P., David P Phillips, P., Kalule, S., and Gibson, R. 2017. Innovative and beneficial informal sweetpotato seed private enterprise in northern Uganda. *In press*.

Sanatkar, M. R., Scoglio, C., Natarajan, B., Isard, S. A., and Garrett, K. A. 2015. History, epidemic evolution, and model burn-in for a network of annual invasion: Soybean rust. Phytopathology 105:947-955.

Sasaya, T. 2015. Detection methods for rice viruses by a reverse-transcription loop-mediated isothermal amplification (RT-LAMP). I ed. Springer, New York, NY, US.

Sutrave, S., Scoglio, C., Isard, S. A., Hutchinson, J. M. S., and Garrett, K. A. 2012. Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. PLoS ONE 7:e37793.

Thomas-Sharma, S., Abdurahman, A., Alic, S., Andrade-Piedrad, J. L., Bao, S., Charkowski, A. O., Crook, D., Kadian, M., Kromann, P., Struik, P. C., Torrance, L., Garrett, K. A., and Forbes, G. A. 2016. Seed degeneration in potato: the need for an integrated seed health strategy to mitigate the problem in developing countries. Plant Pathology 65:3-16.

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Tugume, A. K., Mukasa, S. B., and Valkonen, J. P. T. 2008. Natural wild hosts of Sweet potato feathery mottle virus show spatial differences in virus incidence and virus-like diseases in Uganda. Phytopathology 98:640-652.

Tugume, A. K., Mukasa, S. B., and Valkonen, J. P. 2016. Mixed infections of four viruses, the incidence and phylogenetic relationships of sweet potato chlorotic fleck virus (Betaflexiviridae) isolates in wild species and sweetpotatoes in Uganda and evidence of distinct isolates in East Africa. PLoS ONE 11:e0167769.

Wang, H., Cui, X., Wang, X., Liu, S., and Zhou, X. 2016. First Report of *Sri Lankan cassava mosaic virus Infecting Cassava in Cambodia*. Plant Disease 100:1029.

Wangai, A. W., Redinbaugh, M. G., Kinyua, M., Miano, D. W., Leley, P. K., Kasina, M., Mahuku, G., Scheets, K., and Jeffers, D. 2012. First Report of Maize chlorotic mottle virus and Maize Lethal Necrosis in Kenya. Plant Disease 96:1582.

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

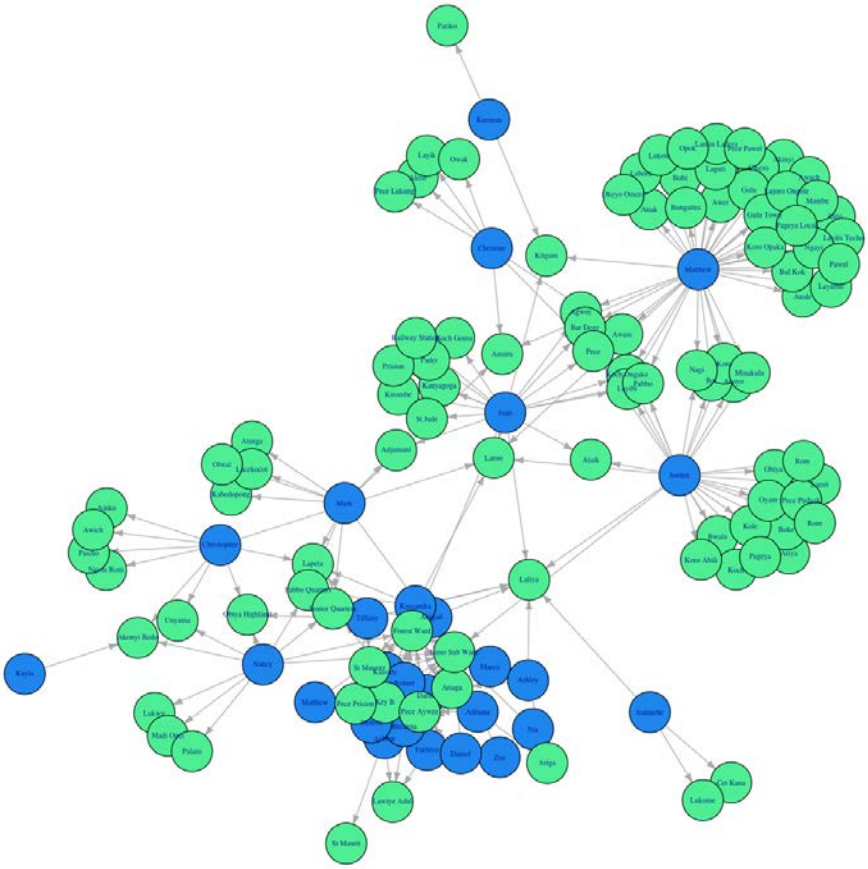
Table 1
Network characteristics of aggregated sweetpotato vine transactions occurring in the 2014 growing season in Northern Uganda.

Network Properties				
Total Nodes	Sellers	Villages	Density	Modularity
126	27	99	0.013	0.556

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1

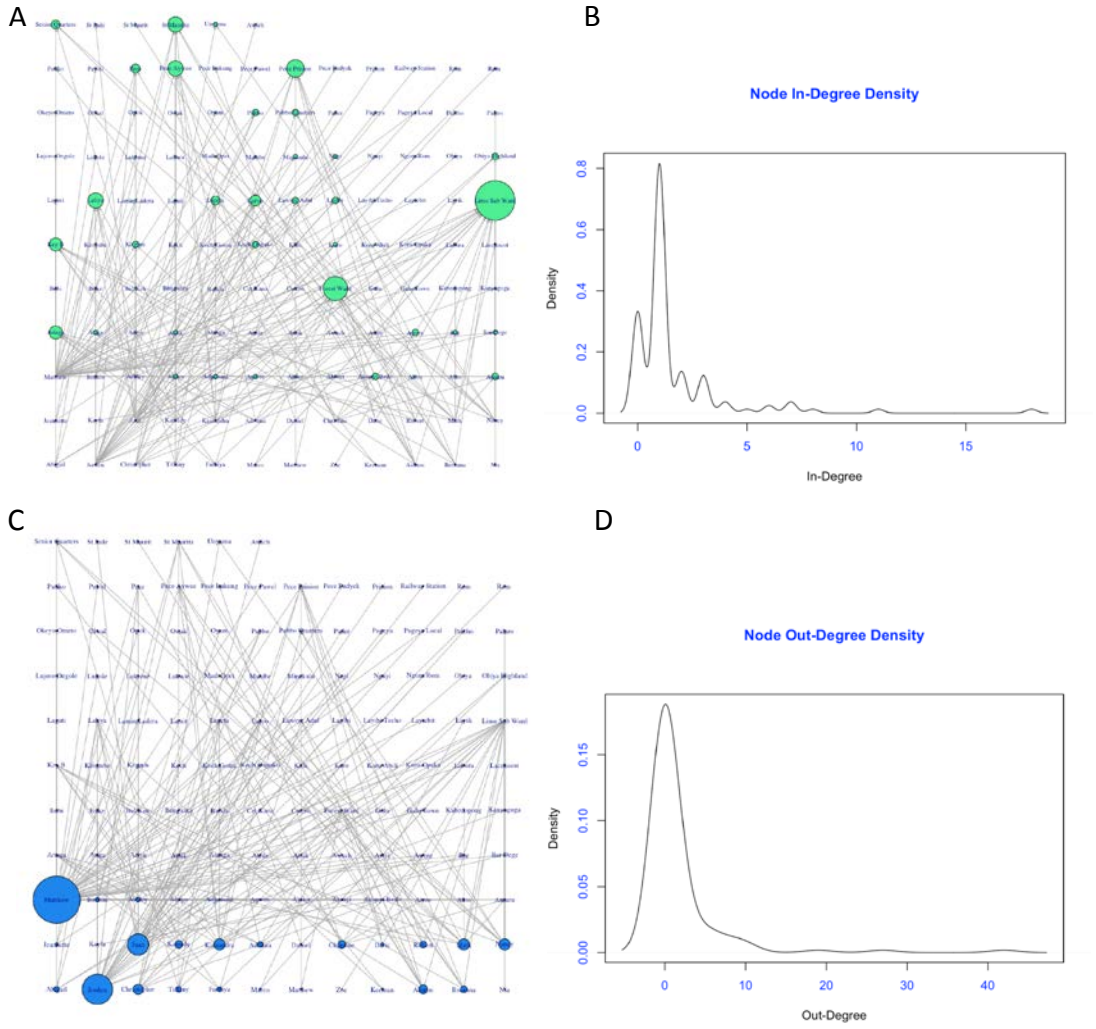
The network structure of sweetpotato vine transactions occurring in the 2014 growing season in Northern Uganda with both sellers (blue nodes) and villages (green nodes). Links represent the occurrence of >1 transaction in the 2014 growing season. Names of villages are true local names of Ugandan villages, while names of sellers have been anonymized.



bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 2

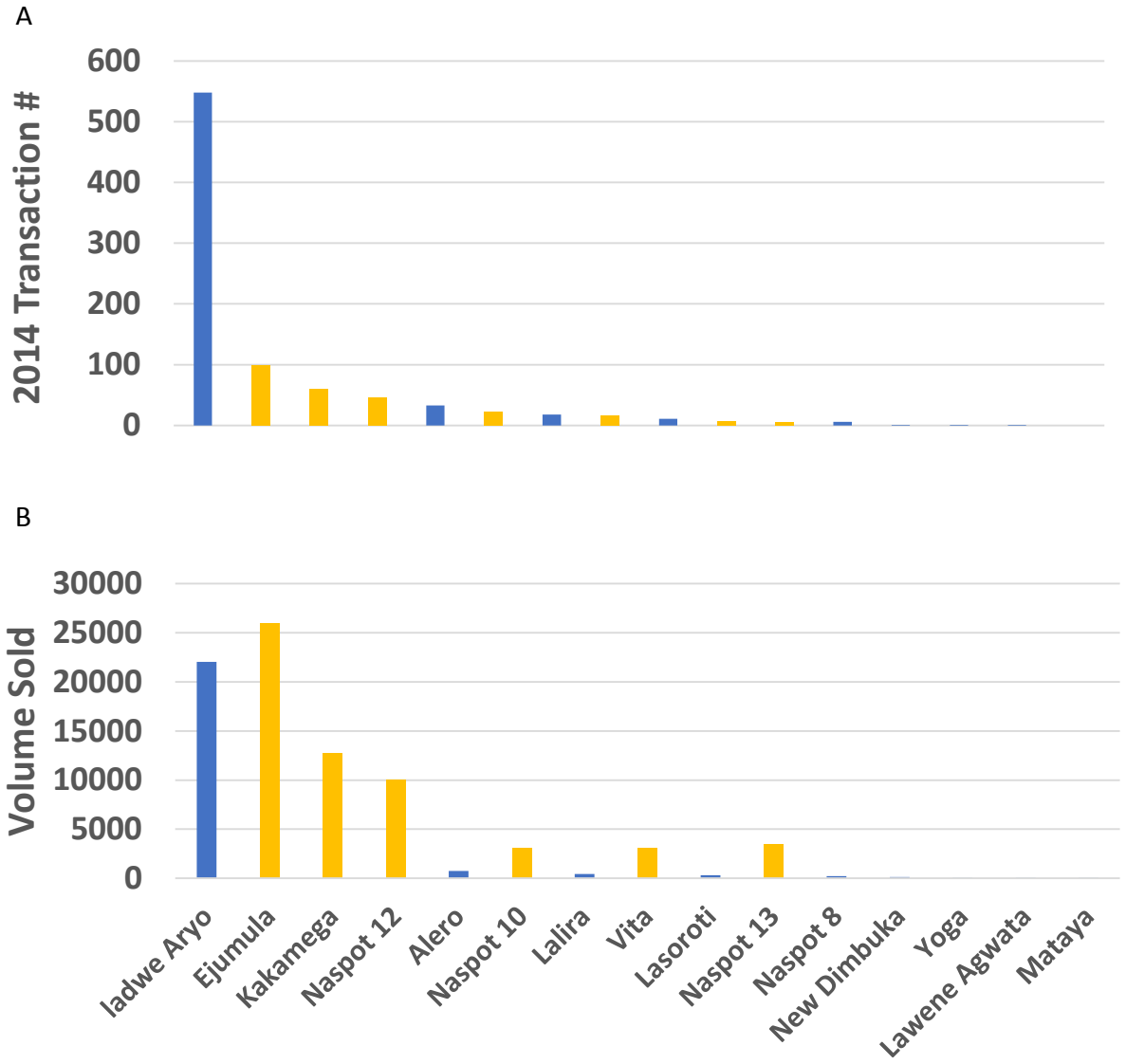
Node degree of villages and sellers. Node size indicates node in-degree for villages (A) and node out-degree for sellers (C). Node degree is the number of connections to or from a given node. Node in- (B) and out- (C) degree density reveal a heavily right-skewed node degree distribution, indicating that a large number of nodes have few connections and a small number of nodes have many connections.



bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3

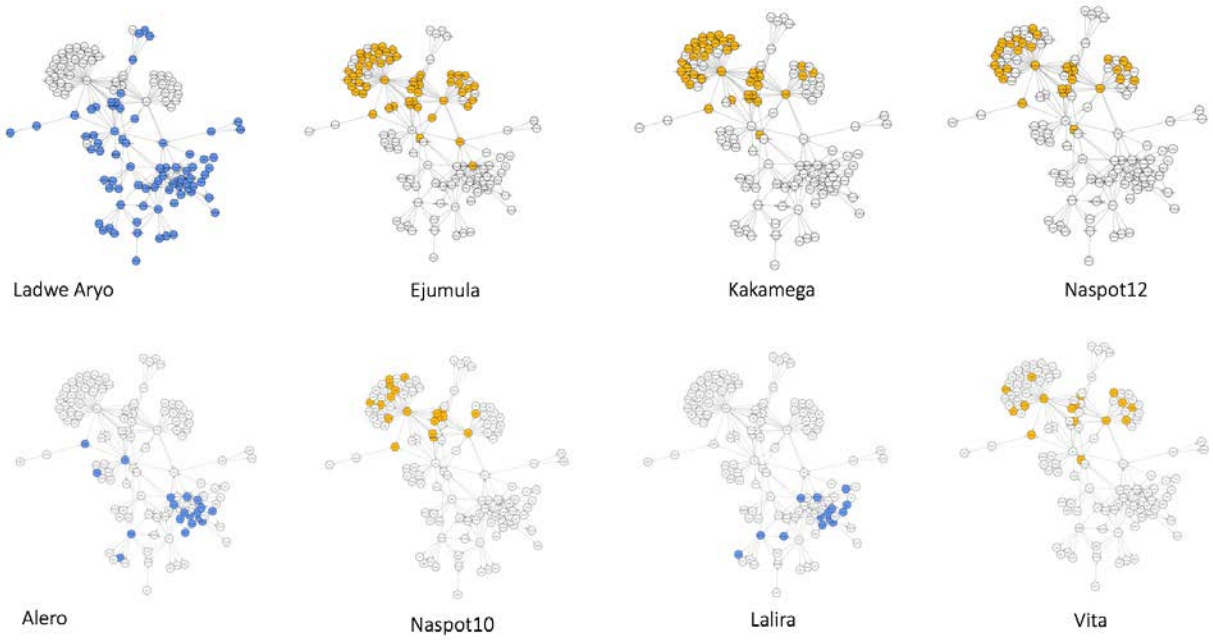
Number of transactions (A) and volume sold (B) for the 16 varieties that were sold over the course of the 2014 growing season. Bar color indicates white-fleshed (blue) and orange-fleshed (orange) cultivars.



bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 4

Networks of dissemination of the top eight cultivars sold during the 2014 growing season. All sellers and villages surveyed are indicated in each network, while colored nodes represent sellers and villages involved in the sale or purchase of the specified cultivar. White villages did not access a given variety in 2014 through this seed network. Node color indicates white-fleshed (blue) and orange-fleshed (orange) cultivars.



bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 5

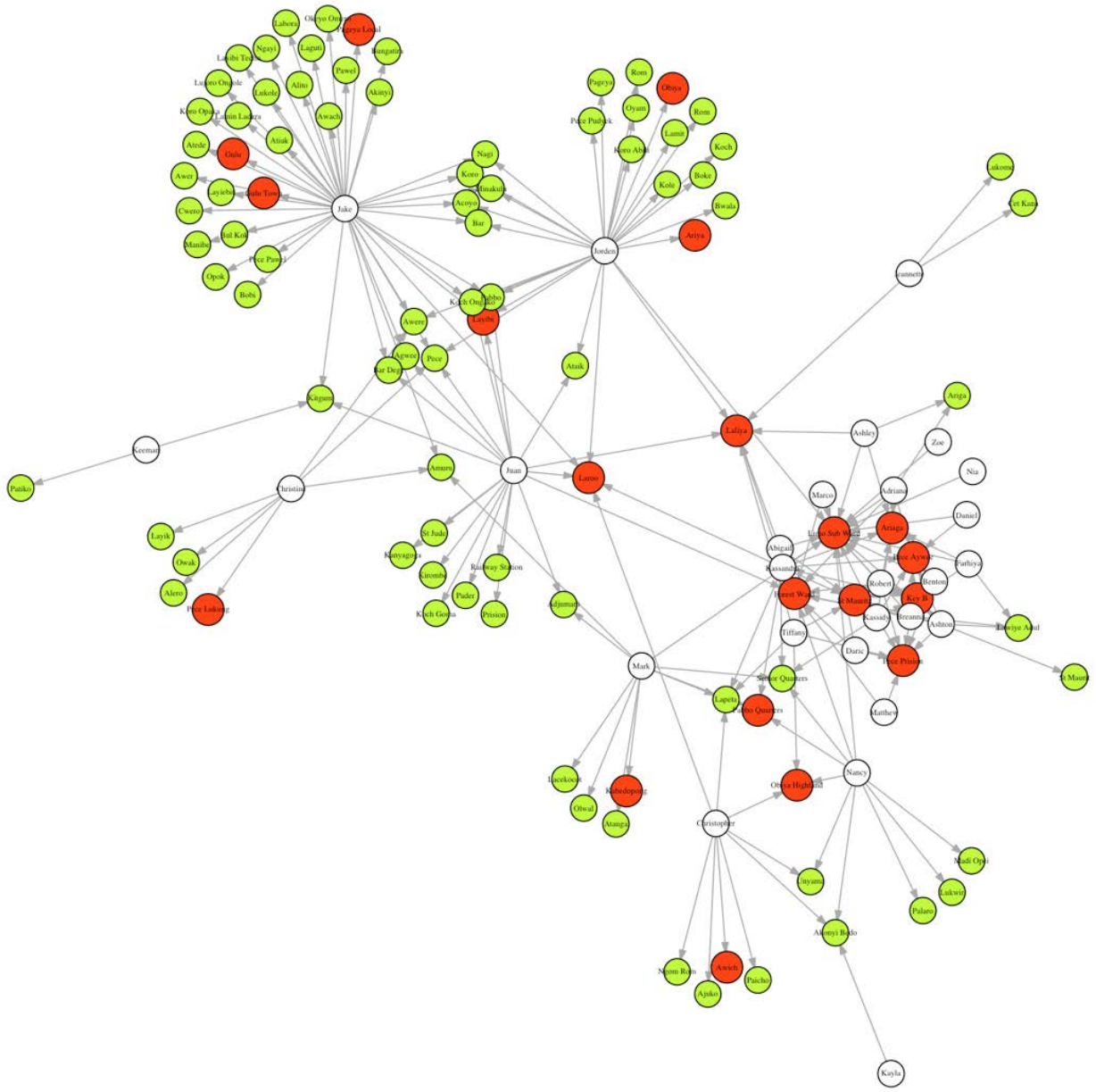
Networks of simulated pathogen introduction into a sweetpotato seed system in Northern Uganda. Twenty-one sellers have the potential to transmit a virus in Time 1 (A) through infected material to villages, based on known transactions, with a 5% chance in Time 2, infecting 6 villages (B). A village in Time 2 has the potential to transmit the virus to neighboring villages, and infected villages are pictured in Time 3 (C).

A. Time 1

B. Time 2

C. Time 3

bioRxiv preprint doi: https://doi.org/10.1101/107359; this version posted February 10, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 6

Network resulting from a simulation where each node had equal potential of being the introduction point of infection. Nodes in red would be favorable for sampling for virus introduction, as the fewest villages will be infected by the time the virus is detected in these nodes. White nodes are sellers and green nodes indicate the remainder of villages.