Network biology discovers pathogen contact points in host protein-protein interactomes

Hadia Ahmed1, T.C. Howton2, Yali Sun2, Natascha Weinberger3, Youssef Belkhadir3 & M. Shahid Mukhtar2,4

In all organisms, major biological processes are controlled by complex protein–protein interactions networks (interactomes), yet their structural complexity presents major analytical challenges. Here, we integrate a compendium of over 4300 phenotypes with Arabidopsis interactome (AI-1MAIN). We show that nodes with high connectivity and betweenness are enriched and depleted in conditional and essential phenotypes, respectively. Such nodes are located in the innermost layers of AI-1MAIN and are preferential targets of pathogen effectors. We extend these network-centric analyses to Cell Surface Interactome (CSI-LRR) and predict its 35 most influential nodes. To determine their biological relevance, we show that these proteins physically interact with pathogen effectors and modulate plant immunity. Overall, our findings contrast with centrality-lethality rule, discover fast information spreading nodes, and highlight the structural properties of pathogen targets in two different interactomes. Finally, this theoretical framework could possibly be applicable to other inter-species interactomes to reveal pathogen contact points.

1 Department of Computer Science, University of Alabama at Birmingham, 115A Campbell Hall, 1300 University Boulevard, Birmingham, AL 35294, USA. 2 Department of Biology, University of Alabama at Birmingham, 464 Campbell Hall, 1300 University Boulevard, Birmingham, AL 35294, USA. 3 Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna Biocenter (VBC), Dr Bohr Gasse 3, 1030 Vienna, Austria. 4 Nutrition Obesity Research Center, University of Alabama at Birmingham, 1675 University Blvd, WEBB 568, Birmingham, AL 35294, USA. These authors contributed equally: Hadia Ahmed, T.C. Howton, Yali Sun. Correspondence and requests for materials should be addressed to M.S.M. (email: smukhtar@uab.edu)
networks consist of systems’ components, referred to as nodes and interactions between them, termed ‘edges’. Network representation of a typical biological system constitutes the direct and indirect interactions among diverse molecular components. These molecular players, proteins in particular, participate in a wide range of biological processes, cellular pathways, and signaling cascades. To achieve these cellular functions, proteins operate in conjunction with other partners, typically through direct physical protein–protein interactions (PPIs). The overall proteome-scale of these cellular interactions constitutes an “interactome”. Thus, elucidating the physical characteristics and functional interaction properties of an interactome could potentially reveal novel relationships between host proteins, new community structures as well as unique nodes with signaling cascades. Such structural and functional topological features provide a range of information on individual nodes and edges, distinct modules, and the entire network as a whole. Considering that diverse networks share similar organizational landscapes, and the rate of information flowing through a network is dependent on the connectivity of its components, several parameters of centrality measurements may act as indicators of important nodes in an interactome. For instance, network architectural properties can determine the connectivity and the critical distribution of a particular node within a network. These include degree, the number of connections of a node; betweenness, the fraction of the shortest paths that pass through a node; and eigenvector, a measure of the influence of a node in a network (Fig. 1a). Scale-free topology of a network follows a power law degree (a heavy-tailed) distribution exhibiting a few nodes with increased connectivity. Recently, k-shell decomposition was shown to identify influential spreaders of information in social platforms and scientific publishing society. Thus, deciphering the network architecture and understanding these topological properties could lead to the discovery of novel components in a complex system, which then provide biological insights as well as testable hypotheses.

Several proteome-scale interactomes have been generated in both prokaryotes and eukaryotes including human, and the reference plant Arabidopsis thaliana (hereafter Arabidopsis). These not only mapped the network and module organization of protein interactions onto the overall cellular organization and function but also allowed understanding of genotype-to-phenotype relationships as well as evolution of biological networks and ancestral gene function. As such, several studies in yeast interactomes suggest that high degree (hubs) and high betweenness (bottlenecks) are likely to be encoded by essential genes, a phenomenon termed as centrality-lethality rule. In addition, PPI networks can also be exploited to decipher the complex interplay between hosts and their pathogens during the process of infection. Analyses of interspecies interactomes demonstrated that proteins corresponding to hubs and bottlenecks are targets of pathogen attack. Thus, a conceptual challenge posed by the centrality-lethality rule in analyzing inter-species interactome dataset stems from diverse lifestyles of pathogens on their hosts. Of particular interests are the pathogens that must keep their hosts alive (e.g. obligate biotrophs) throughout their life cycle. Therefore, association of hubs and/or bottlenecks (potential pathogens’ targets) with essentiality/lethality would principally undermine the pathogens’ infectious process. Thus, the phenotypic characteristics of nodes defined as hubs and/or other network centrality measures are a requisite layer of information to biologically understand inter-species interactome datasets.

Previously, we generated an Arabidopsis binary PPI map using ~8000 open reading frames representing ~30% of its protein-coding genes. Known as Arabidopsis Interactome version 1 “main screen” (AI-MAIN), this network encompasses 5664 binary interactions between 2661 proteins. We showed that AI-MAIN displays properties of a scale-free network that exhibits only 15 nodes with more than 50 interactions, i.e., ≥50 edges. These high-degree nodes are referred as hubs. In addition, we also constructed two inter-species Plant–Pathogen Interaction Networks (PPIN-1 and PPIN-2) by systematically interrogating interactions between Arabidopsis proteins and pathogen proteins that are translocated inside the plant cells during infection (also termed pathogen effectors). Specifically, these effectors were derived from three distantly related pathogens. Unexpectedly, however, we determined that these independently evolved effectors interact with a limited repertoire of 201 Arabidopsis proteins (hereafter host or effector targets). Subsequently, we demonstrated that these effectors can modulate host targets to establish effector-triggered susceptibility (ETS). We also showed that these targets participate in various layers of plant immunity including microbial-associated molecular patterns (MAMPs) and Effector-Triggered Immunity (MTI and ETI, respectively). While most nodes corresponding to effector targets in AI-MAIN are highly connected (average degree), less than 6.5% of these nodes were defined experimentally as proteins belonging to the hub class. Thus, the predictive power of computational methods relying solely on centrality measures, particularly hubs, to determine if a given node in an interactome is more inclined to be targeted by pathogen effectors is limited.

Here, we devise a method to predict effector targets in two unrelated experimental interactomes. To fully understand the functional interaction properties of the central nodes within a network, we curate a comprehensive dataset of ~4350 unique phenotypes in Arabidopsis. Unexpectedly, however, we demonstrate that hubs and bottlenecks are enriched in conditional phenotypes and depleted in essential phenotypes contrasting the centrality-lethality rule. We also discover that the nodes located in close proximity of the AI-MAIN core are targeted by effectors. We next apply this network topology framework to the cellular LRR-based Cell Surface Interactome (CST(RK)), an unrelated experimental network that includes >500 interactions between membrane-localized leucine-rich repeat receptor kinases (LRR-RKs). Following centrality measure analyses, we predict a set of 35 LRR-RKs that are located near the core of CST(RK) as the most influential nodes. Using two independent methods, we demonstrate that a subset of these predicted LRR-RKs can physically interact with bacterial effectors. Finally, we provide genetic evidence for the requirement of these newly discovered LRR-RKs modulating in plant immune system activities.

Results
Phenotypic properties of Arabidopsis hubs and bottlenecks. To examine the system-level relationship between genotype-to-phenotype in AI-MAIN, we curated a comprehensive dataset of phenotypes corresponding to loss-of-function mutations in 4344 unique genes in Arabidopsis. We then categorized these genes into five functional groups: essential (ESN), morphological (MRP), cellular-biochemical (CLB), conditional (CND), and no phenotypes (NPH) as described by Lloyd and Meinke (Fig. 1b). We also performed enrichment assays for degree, betweenness, and eigenvector (Supplementary Fig. 1a, 1b). The definition of a high degree node (hub) in an interactome is arbitrary and perhaps depends upon the size and the density of a given network. For instance, we defined hubs with a degree greater than or equal to 50 (hub) in the largest Arabidopsis interactome AI-MAIN as well as PPIN-1 and PPIN-2. The de

ARTICLE NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-04632-8
However, the second largest Arabidopsis interactome, MIND1 (Arabidopsis Membrane-linked Interactome), described hub proteins with degree >70. To demonstrate the robustness of our analysis, we implemented a second cut-off value for nodes displaying greater than or equal to 25 interactions (hub25). Given that high betweenness (bottlenecks) and high eigenvector cut-off values were not defined in either of the two largest Arabidopsis interactomes, AI-1MAIN and MIND1, we also included two cut-off values each for high betweenness (bottleneck 0.025 or bottleneck0.01) and high eigenvector (0.1 or 0.01) (Supplementary Data 2). Our analysis revealed that CND phenotypes are enriched in hub50 and hub25 (hypergeometric $P < 0.05$, Supplementary Data 2) as well as in bottleneck0.01 (hypergeometric $P = 0.055$, Supplementary Data 2). We also discovered that ESN phenotypes were enriched, although not statistically significant, in non-hubs (nodes with less than 25 edges) in AI-1MAIN (hypergeometric $P = 0.11$, Supplementary Data 2). Finally, we did not observe a significant association of high eigenvector nodes in any of the above-mentioned phenotypes (Supplementary Fig. 2 and Supplementary Data 2). To control that the enrichment of CND phenotypes in hubs and bottlenecks is specific, we generated two random networks (“degree-preserving” and “non-degree-preserving”) encompassing nodes and edges similar to AI-1MAIN. Both random networks did not exhibit enrichment in any of the five phenotypes (Supplementary Fig. 3 and Supplementary Data 2). Thus, based on these analyses, we
concluded that high degree (hubs) and high betweenness (bottlenecks) are enriched in CND but not in ESN phenotypes.

Enrichment of CND phenotypes with both hubs and bottlenecks prompted us to test whether high degree and high betweenness share significant fraction of the nodes with each other. Undoubtedly, we observed a strong positive correlation between degree and betweenness (Fig. 1c; \( r^2 = 0.87 \)). An analogous observation has been reported for Compound-Potential Target Network in cardiovascular disease \(^3\) (\( r^2 = 0.77 \)). However, the overlap of nodes corresponding to hubs or bottlenecks with high eigenvector did not yield any significant positive correlation (Supplementary Fig. 4; \( r^2 = 0.55 \)). Taken together, we showed that most central nodes in the network have a high degree and a high betweenness, and that most information perhaps flows through those important nodes. To analyze phenotypic groups’ enrichment assay on nodes that exhibit both hub and bottleneck properties, we categorized the nodes as high degree/high betweenness (HDBH), high degree/low betweenness (HDLB), low degree/high betweenness (LDHB), and low degree/low betweenness (LDLB). While HDHB nodes at two cut-off values were enriched in the CND phenotypic group (hypergeometric \( P < 0.05 \), Fig. 1f and Supplementary Data 2), no significant association of LDBL, HDLB, or LDHB nodes with any phenotypic functional groups was found. Finally, we did not observe enrichment of CND phenotypes with HDHB nodes in the two random networks (Supplementary Data 2). Thus, hubs and bottlenecks are enriched in CND phenotypes in AI-1\(_{\text{MAIN}}\) thereby contrasting the centrality-lethality rule. We also propose that Arabidopsis cells utilize hub and bottleneck proteins to regulate the flow and spread of information to a large number of proteins under diverse physiological conditions.

**Predictability of effector targets in plant interactome.** Previous studies have shown that specialized pathogens have evolved sophisticated mechanisms to manipulate the key components of their hosts’ intracellular networks to their advantage \(^4,29\). Thus, we hypothesized that pathogens use effectors to target the most influential nodes in their host network. To test this concept, we determined if nodes corresponding to hubs, bottlenecks or high eigenvectors were more prone to be effector targets. Our results showed that high degree and high betweenness proteins (HDHB) are likely to be direct physical contact points of pathogen effectors, yet they only account for a small fraction of the range of effector targets determined experimentally in PPIN-1 and PPIN-2 (i.e. 6.45% and 18.71% for two cut-off values applied in our analyses, respectively) (Supplementary Fig. 5a and Supplementary Data 2). In addition, the target discovery rate of high eigenvector, HDLB, and LDHB with two cut-off values is lower than that of HDHB nodes (Supplementary Data 2). Given that PPIN-1 and PPIN-2 utilized effectors from three different pathogens, we also investigated whether a particular node targeted with more than

---

*Fig. 2* Network analyses of nodes in various layers of AI-1\(_{\text{MAIN}}\). **a** Schematic illustration of network layering using the weighted \(k\)-shell decomposition method. Connected hypothetical network (left; gray nodes) and decomposed network into three shells (right; \(k = 1, k = 2, \text{and } k = 3\) in green, red, and black colors) are shown. **b** Distribution of average degree of each shell from the innermost of the network (core) designated as 1 to the periphery of the network denoted as 1000 in AI-1\(_{\text{MAIN}}\). Effector targets and non-targets are shown in red and blue nodes, respectively (\(r^2 = 0.67\) and Mann-Whitney-Wilcoxon Test \(P < 2.2 \times 10^{-16}\)). **c** Average degree (Welch’s t-test \(P = 1.57 \times 10^{-14}\)) and **d** average betweenness (Welch’s t-test \(P = 4.27 \times 10^{-12}\)) for internal layers AI-1\(_{\text{MAIN}}\) proteins (red) and peripheral layers AI-1\(_{\text{MAIN}}\) proteins (blue) are plotted. **e** Distribution of average information centrality (IC) for each shell starting from the core of the network in AI-1\(_{\text{MAIN}}\) (\(r^2 = 0.82\) and Mann-Whitney-Wilcoxon test \(P < 2.2 \times 10^{-16}\)).
one effector from the same pathogen or different pathogens could be used as a predictive indicator. However, we did not observe any correlation between the number of unique effectors interacting with a particular node and its degree in AI-1\textsubscript{MAIN} (Supplementary Fig. 5b). In fact, the hub with the highest number of connections in AI-1\textsubscript{MAIN} is targeted by only a single effector. Taken together, we concluded that centrality measures such as degree, betweenness, and eigenvector are thus of limited use to comprehensively analyze inter-species interactome datasets.

**Structural features of nodes in Arabidopsis interactome.**

Recently, \(k\)-shell decomposition analysis was shown to outperform other known centrality measures including degree, betweenness, and PageRank in network-based analyses and for the identification of the most influential proteins in the network\textsuperscript{14}. While the unweighted \(k\)-shell decomposition analysis considers all edges equally\textsuperscript{16}, we used a weighted \(k\)-shell decomposition method to understand the topological properties of AI-1\textsubscript{MAIN}\textsuperscript{37} (Fig. 2a). We defined the internal and peripheral layers (shells) for AI-1\textsubscript{MAIN} nodes that reside within the one-third and two-third layers, respectively (Supplementary Data 3). We observed a power-law correlation between the average degree and shell depth \((r^2 = 0.67\) and Mann–Whitney–Wilcoxon Test \(P < 2.2 \times 10^{-16}\)) (Fig. 2b and Supplementary Fig. 6). We also demonstrated that the nodes located in the vicinity of the network core (internal layers AI-1\textsubscript{MAIN} nodes) possess significantly higher average degree and betweenness in comparison to the nodes distributing in the periphery of the network (Fig. 2c, d, \(P = 1.57 \times 10^{-14}\) and \(P = 4.27 \times 10^{-12}\), respectively). These data indicate that the nodes residing within the internal layers are possibly better information spreaders. To substantiate this, we measured the information centrality (IC), an index that focuses on how information might flow through many different paths\textsuperscript{13}. While we observed a strong power-law correlation between IC and shell depth \((r^2 = 0.82\) and Mann–Whitney–Wilcoxon Test \(P < 2.2 \times 10^{-16}\)), we also showed that the average IC of nodes...
present in the internal layers of AI-1\textsubscript{MAIN} is significantly higher than that of proteins in the remaining network (Fig. 2e and Supplementary Fig. 7, $P < 2.2 \times 10^{-16}$). These data indicate that the proteins closer to the network core are poised to be the most active spreaders of information.

**Effector targets in AI-1\textsubscript{MAIN} by k-shell analysis targets.** Since the internal layers of AI-1\textsubscript{MAIN} are enriched with nodes corresponding to influential spreaders of information, we thus predicted that effectors preferentially target nodes distributing in the vicinity of network core. Towards this, we demonstrated that nodes present in the internal layers of AI-1\textsubscript{MAIN} are significantly enriched with effector targets compared to those located in the periphery of network (Fig. 3a, b, hypergeometric $P = 2.61 \times 10^{-48}$) with 33\% discovery rate of effector targets (Supplementary Data 2, $P = 3.01 \times 10^{-50}$). No enrichment of effector targets was
observed in randomly generated networks (Fig. 3c, d and Supplementary Data 2). In concordance with these results, we next showed that nodes that reside in the internal layers of AI-1.MAIN are enriched in CND phenotypes, and depleted in ESN phenotypes (hypergeometric \( P = 0.05 \), Fig. 3e and Supplementary Data 2) compared to the proteins in the periphery of the network. However, we did not observe any enrichment of these phenotypic groups in the internal layers of two independent random networks (Supplementary Fig. 8a and b and Supplementary Data 2). These results indicate that the weighted \( k \)-shell decomposition analysis surpasses other centrality measures for effector target discovery.

Previously, we performed a phenotypic mapping experiment of 124 Arabidopsis mutants corresponding to effector targets. In that study, we showed that 63 effector targets display disease-related phenotypes\(^{27}\), suggesting an almost equal chance (51%) to obtain immune-related phenotype or no phenotype for a mutant corresponding to an effector target. Remarkably, we demonstrated that the nodes located in the internal layers of AI-1.MAIN are enriched and depleted in immune-related phenotypes and no immune-related phenotypes, respectively (Fig. 3f and Supplementary Data 2, hypergeometric \( P = 2.552 \times 10^{-6} \)). This enrichment of immune-related phenotypes was absent in the internal layers of both random networks (Supplementary Fig. 9c and d and Supplementary Data 2). Intriguingly, we did not observe any correlation between the average effector degree (interacting degree of an effector to host proteins) and the proximity of the network core (Supplementary Fig. 9). Collectively, our data suggest that nodes located closer to the core of the network are targeted by effectors. Moreover, these nodes are enriched with CND and immune-related phenotypes.

**Discovery of the most influential nodes in CSI\(^{LRR} \)**

LRR-RKs control plant growth and immunity by detecting and responding to ‘self’ and ‘non-self’ signals in the extracellular space. These surface localized receptors can act as pattern recognition receptors by sensing MAMPs, thereby controlling MTI\(^{36-40}\). Since a small subset of LRR-RKs have been shown to be targeted directly by pathogen effectors, we extended our weighted \( k \)-shell decomposition and functional analyses to identify both effector targets as well as the most influential spreaders of information in CSI\(^{LRR} \). Using our approach, we assigned 35 LRR-RKs to the one-third internal shells (or the internal layers) of CSI\(^{LRR} \) (Fig. 4a and Supplementary Data 3), and we postulated that these receptors are likely to be the most influential spreaders of information. Towards this, we performed additional network-centric analyses. We observed strong power-law correlations between the shell depth and the average degree (\( r^2 = 0.9 \)) and the average IC (\( r^2 = 0.93 \)) in CSI\(^{LRR} \) (Fig. 4b, c and Supplementary Fig. 10a, \( P = 2.43 \times 10^{-15} \), respectively). As in AI-1.MAIN, the average degree value of the nodes located in the internal layers of CSI\(^{LRR} \) was significantly higher than that of their peripheral counterparts (Supplementary Fig. 10b, \( P < 0.001 \)). Similar to AI-1.MAIN\(^{7} \) and Compound-Potential Target Network in cardiovascular disease\(^{35} \), we discovered a significant overlap of nodes between high degree and high betweenness in CSI\(^{LRR} \) network (Supplementary Fig. 10c, \( r^2 = 0.71 \)). Thus, although generated by independent methods, CSI\(^{LRR} \) and AI-1.MAIN share an overall similar network architecture based on centrality measures and weighted \( k \)-shell decomposition analyses.

Based on the network topological similarity concept between CSI\(^{LRR} \) and AI-1.MAIN, we predicted that the nodes present in the internal layers of CSI\(^{LRR} \) should be associated with CND and immune-related phenotypes including MTI. Our analysis was limited by the dearth of LRR-RKs for which a clearly defined function has been assigned in the literature. However, we found that BRI1-associated receptor kinase 1 (BAK1), the most interconnected node in CSI\(^{LRR} \)\(^{16} \), is located in the core of CSI\(^{LRR} \). BAK1 acts as a major coreceptor for a range of ligand binding receptors that regulate MTI and plant development, and is, therefore, also a functional hub\(^{38} \). It is worth noting that none of these LRR-RKs have been previously associated with ESN phenotypes, further suggesting the roles of this set of LRR-RKs in stress responses.

In addition to the roles of BAK1\(^{38} \), the functions of 22 other LRR-RKs including somatic embryogenesis receptor kinase (SERK)\(^{36} \), BAK1-interacting LRR-RKs (BIRs)\(^{39} \), Brassinosteroid insensitive 1 (BRI1)-LIKE (BRLs)\(^{41} \), ERECTA (ER)\(^{40} \), ER-like (ERLIs)\(^{42} \), flg22-induced receptor-like kinase 1 (FRK1)\(^{29,42} \), Impaired Oomycete Susceptibility 1 (IOS1)\(^{43} \), Receptor Protein Kinase 1 (RPK1)\(^{44} \), Senescence-Associated Receptor-Like Kinase (SARK)\(^{45} \), Articulation Point Executive (APEX)\(^{16} \), Flagellin Sensitive 2 (FLS2)\(^{46} \), HAESA Like (HSL2)\(^{47} \), Strubellig Receptor Family 3 (SRF3)\(^{48} \), and PSY1-receptor (PSY1R)\(^{49} \) have been previously proposed in MTI as well as other biotic and abiotic stresses (CND phenotypes). To further substantiate the potential functions of these 35 LRR-RKs in CND as well as immune-related phenotypes, we compared them with an MTI subnetwork\(^{16} \). This immune-related module was derived through a community analysis in CSI\(^{LRR} \). We found that LRR-RKs located in the internal layers of CSI\(^{LRR} \) constitute 66% of the MTI subnetwork (Fig. 4a).

Given the overwhelming enrichment of LRR-RKs corresponding to CSI\(^{LRR} \) internal layers with CND and immune-related phenotypes, we further hypothesized that these sets of LRR-RKs are potential targets of pathogen effectors. To test this, we
performed a pairwise Yeast two-hybrid (Y2H) experiment and tested cytoplasmic domains of 20 LRR-RKs against 31 effectors from *Pseudomonas syringae pv. tomato* DC3000 (Pto DC3000). We recapitulated the interaction of BAK1 with HopAB2, originally discovered in split-ubiquitin system. Moreover, we also found seven additional LRR-RKs interacting with nine effectors (Fig. 4d, e; 40% effector discovery rate, \(P < 2.2 \times 10^{-16}\)). In contrast, a parallel experiment involving LRR-RKs that distribute in the peripheral layers of CSI LRR showed no significant enrichment of effector target discovery rate (6.25%; Fig. 4e). We further validated these inter-species interactions by employing split-YFP system in Arabidopsis cells, an independent confirmatory method (Fig. 4f). Thus, we expected the internal layers CSILRR LRR-RKs to be the converging points of effectors from diverse pathogens. Indeed, FLS2 was previously demonstrated to associate with a bacterial effector, AvrPto in a co-immunoprecipitation assay. Moreover, three NSP-Interacting Kinases, NIK1, NIK2, and NIK3, were previously shown as virulence targets of the geminivirus nuclear shuttle protein (NSP), further substantiating the discovery rate of effector targets located within the internal layers of CSILRR.

Immune-related functions of newly identified LRR-RKs. In addition to the known CND and immune phenotypes for 22 LRR-RKs, we aimed to characterize the roles of seven additional LRR-RKs in plant immunity (MTI and ETS). We obtained loss-of-function mutants corresponding to NIK1, NIK2, NIK3, SRF6,
SRF9, RPK1, and APEX and demonstrated the lack of transcript accumulation in these mutants. We hypothesized that mutants corresponding to these seven LRR-RKs would manifest CND and immune-related phenotypes. To test this, we subjected the mutants corresponding to these seven LRR-RKs to infection with either the fully virulent bacterial pathogen *Pto* DC3000 or with *Pto* DC3000 hrcC−, a mutant strain that lacks a functional type-III secretion system required for effector protein delivery into host cells. While we reproducibly observed a significant increase in the virulence of *Pto* DC3000 on the *skf9*, *apex*, *sr6-2*, *rpk1*, and *nik3* mutants compared with wild-type Col-0 plants (Fig. 5), no significant difference in bacterial growth was observed when plants were infected with *Pto* DC3000 hrcC− except for *nik3* (Fig. 5). These results indicate that SRF9, APEX, SRF6, and RPK1 LRR-RK receptors have an MTI-independent function and negatively regulate the virulence activities of one or more effectors. In comparison to wild-type plants, we observed a significant reduction of *Pto* DC3000 hrcC− growth in the *nik1* and *nik2* mutants, whereas *Pto* DC3000 growth was unaffected (Fig. 5). Thus, NIK1 and NIK2 negatively regulate the induction of MTI. Overall, we demonstrated the positive and negative contributions of these seven LRR-RKs in MTI as well as ETS under diverse physiological conditions. While the molecular mechanisms by which these newly identified LRR-RKs contribute to plant defense are focal points of future research, here we discovered novel players of plant immunity in *CSI*<sup>LRR</sup> using network biology-based approaches.

**Discussion**

In the last 15 years, interactome mapping in diverse organisms led to the development of several premises in network biology, scale-free network architecture, nodes’ connectivity, and the centrality-lethality rule are applied to discover novel components in diverse systems. In this study, we performed an in-depth network analyses on two unrelated experimental interactomes and revealed their topological features. We determined that high degree and high betweenness nodes are enriched and depleted in conditional and essential phenotypes, respectively. Additional noteworthy findings concern another widely known network model implying that highly connected and central nodes are targets of diverse pathogens. Instead, we demonstrated that nodes with increased connectivity that are located closer to the network core are the preferred targets of pathogen attack compared to the proteins that reside in the network periphery. Finally, we identified previously known as well as novel LRR-RKs involved in MTI and ETS.

We showed that both AI-1<sub>MAIN</sub> and CSI<sup>LRR</sup> displayed properties of a scale-free network (Fig. 4 and Supplementary Fig. 1). Since the birth of this theory, however, several seminal studies have outlined sentinel importance or presented contradicting views of the scale-free property. An important question, however, is whether the power law distribution of nodes is a consequence of a specific technology bias, for example, yeast two-hybrid (Y2H) vs. affinity purification with mass spectrometry (AP-MS). Irrespective to the choice of research methods, dozens of large-scale interactomes in both prokaryotes and eukaryotes have been reported to exhibit scale-free properties. In Arabidopsis, AI-1<sub>MAIN</sub> was generated using GAL4-based Y2H method by employing over 8000 ORFs. While a systems-level method by employing over 8000 ORFs, we showed previously known as well as novel LRR-RKs were targeted by effectors from diverse pathogens. While hubs and bottlenecks are remarkable predictors of pathogen targets, they only make up for a small fraction of nodes in a scale-free network, i.e., 6.5% in AI-1<sub>MAIN</sub>. These data indicate that hubs and bottlenecks can predict pathogen effectors in Arabidopsis with high significance as shown for human–viral or human–bacterial interactomes, but the predictive power of these centrality measures is very low. Given that infectious organisms require the hosts to remain viable for their growth and reproduction, a very recent report suggests that interactome connectivity directly relates to pathogen fitness during infection. According to this tenant, pathogens rearrange host interactomes instead of dismantling network integrity to alter cellular physiology for their benefits. Thus, we expected that...
Methods

Network analyses. The centrality measures in both Arabidopsis Interactome version 1 "main screen" (AI-1MAIN) and Cell Surface Interactome (CSI-LRR) were analyzed using Networkx package and Python 2.7.10. Briefly, we calculated degree, the n number of edges of a particular node, and degree distribution of a network is defined as $n_i/n$. Betweenness, the number of shortest paths that pass through a node $v_i$ is analyzed as $g(v_i) = \sum_{st} \sigma_{st}(v_i)/\sigma_{st}$, where $\sigma_{st}$ is the sum of shortest paths from node $s$ to node $t$; $t$ and $\sigma_{st}(v_i)$ is the number of paths that pass through $v_i$. Eigenvector, a measure of the influence of a node in a network, $x_i$ of node $i$, is calculated as $x_i=1/\lambda \sum_{j \neq i} \xi_{ij} w_{ij}$ for each node we computed its degree, betweenness, and eigenvector as described above. Hence, we selected two cut-offs for each case.

Degree: 50 and 25
Betweenness: 0.025 and 0.01
Eigenvector: 0.1 and 0.01

IC calculates the flow of information between two nodes in a connected network. IC was computed as described previously. Briefly, IC (i) for node i in a graph G is calculated as $IC(i) = \frac{1}{\sum_{j \neq i} T_{ij}}$.

Here, $n$ is the total count of nodes and $T_{ij}=\{(r_i+r_j-r_{ij})-1\} r_{ij}$ is a component of R matrix. $D$ is a weighted degree diagonal matrix for each node, and $\beta$ is a matrix consisting 1 for all elements. Therefore, $R=(\{D-\beta\})^{-1}$. Mathematically, $T_{ij}$ is well-defined as infinite. Hence, $\frac{1}{D} = 0$.

Weighted k-shell decomposition is performed as described in Fig. 2 and Wei et al. 179. Briefly, the generation of shells process is defined by the weight of both degree of a node and its edges and calculated as $K_i = \alpha w_i + (1 - \alpha) \sum_{j \in T_i} w_{ij}$, where $\alpha$ is a set of neighboring nodes of $i$, $w_i$ is the weight of the edge that is defined as $w_{ij} = K_i + K_j$. The value of $\alpha$ can be set on a spectrum of 0 through 1 with 0 and determining high edge or high degree favorability in k-shell decomposition calculation, respectively. We performed k-shell decomposition using a range of cut-off values, i.e., 0, 0.5, and 1.0.
Table 1 List of the primers used in this study

| Gene   | Name of primer     | Sequence 5'-3'   |
|--------|--------------------|------------------|
| NIK1   | SALK_017538-LP     | GACAAAAACATGACAGGGTGGG |
| NIK2   | SALK_017538-RP     | CGTATTTCTGTTGTGCCT   |
| NIK3   | SALK_044363-LP     | CCAAGAAGAAGAAAAACCAAGCC |
| NIK3   | SALK_044363-RP     | GAAAGGTTATCAATGTCGCTC |
| SRF6-1 | SALK_054337_LP     | AAGGCTTACCGTCGAGATTC |
| SRF6-2 | SALK_054337_RP     | TGGGATATTAAACGGTCTG |
| SRF9   | SALK_077702_LP     | GTTGCTGACCATGGAAAGTC |
| SRF9   | SALK_014459-RP     | TCCCTGCTCACCATAACGAG |
| RPK1   | SALK_005054_LP     | CTACGTCAACAAGGTTGGG |
| APEX   | SALK_055240_LP     | GCATAAGCCATTTTCCCAAAC |
| APEX   | SALK_055240_RP     | TCATGGAAACCTTACCAGTTC |

calculated by summing the degree of each node in the shell and dividing by the number of nodes presented in the shell using the following formula

\[
\text{Shell_avgDegree}_{\text{shell}} = \frac{\sum_{i=1}^{N \text{nodes}} \text{Degree}_i}{N \text{nodes}} \times \frac{1}{N - \text{nodes}}
\]

where \(N\text{nodes}\) is the number of nodes in the shell, \(N\) is the number of nodes in the dataset, \(\text{Degree}_i\) is the degree of the node \(i\), and \(N - \text{nodes}\) is the number of nodes removed from the dataset to calculate the shell.

Statistical analyses. We calculated hypergeometric test, linear regression (\(r^2\)), Mann–Whitney–Wilcoxon test, and Welch’s \(t\)-test using R version 3.3.1 as well as online Stat Trek tool. Briefly, hypergeometric test was performed to determine the enrichment of five phenotypic groups: (1) essential (ESN), (2) morphological (MRP), (3) cellular-biochemical (CLB), (4) conditional (CND), and (5) no phenotype (NPH) as described by Lloyd and Meinke34. In addition, we downloaded genome-wide phenotypes from TAIR1069 for each T-DNA insertion line. An intact WT was used in all experiments. The wild-type used in all experiments was Arabidopsis thaliana (Col-0) and the wild-type tester strain was S. cerevisiae (S. cerevisiae strain D17A, \(\text{CYH2}^+\)).

References

1. Kohli, P., Emborg, C. C., Bangalore, P. V., Kannan, P. & Mukhtar, M. S. Getting to the edge: protein dynamical networks as a new frontier in plant-microbe interactions. Front. Plant Sci. 5, 312 (2014).
2. Hu, J. X., Thomas, C. E. & Brunak, S. Network biology concepts in complex disease comorbidities. Nat. Rev. Genet. 17, 615–629 (2016).
3. Vidal, M., Cinzio, M. E. & Barabasi, A. L. Interactome networks and human disease. Cell 144, 986–998 (2011).
4. McCormack, M. E., Lopez, J. A., Crocker, T. A. & Mukhtar, M. S. Making the right connections: network biology and plant immune system dynamics. Curr. Plant Biol. 5, 2–12 (2016).

Data availability. All supporting data from this study are available from the article and Supplementary Information files, or from the corresponding author upon reasonable request. Moreover, the weighted-k-shell algorithm implemented in Java language and can be accessed at http://cgojo.cs.umd.edu.

Received: 17 February 2018 Accepted: 11 May 2018
Published online: 13 June 2018
Windram, O., Penfold, C. A. & Denby, K. J. Network modeling to understand plant immunity. *Curr. Opin. Struct. Biol.* **44**, 201–210 (2017).

Arabidopsis Interactome Mapping, C. Evidence for network evolution in an Arabidopsis interactome map. *Science* **333**, 601–607 (2011).

Barabasi, A. L. Network science. *Philos. Trans. A Math. Phys. Eng. Sci.* **371**, 20120375 (2015).

Windram, O., Penfold, C. A. & Denby, K. J. Network modeling to understand plant immunity. *Annu. Rev. Phytopathol.* **52**, 93–111 (2014).

Barzel, B. & Barabasi, A. L. Universality in network dynamics. *Nat. Phys.* **9**, 673–681 (2013).

Yang, K. K. et al. Cross-disciplinary network comparison: matchmaking between hairballs. *Cell Syst.* **2**, 147–157 (2016).

Gao, J., Barzel, B. & Barabasi, A. L. Universal resilience patterns in complex networks. *Nature* **536**, 238 (2016).

Mishra, B., Sun, Y., Ahmad, H., Liu, X. & Mukhtar, M. S. Global temporal dynamic landscape of pathogen-mediated subversion of Arabidopsis innate immunity. *Sci. Rep.* **7**, 7849 (2017).

Pei, S., Muchnik, L., Andrade, J. S. J., Zheng, Z. & Makse, H. A. Searching for superspreaders of information in real-world social media. *Sci. Rep.* **4**, 5547 (2014).

Klockgether, K. et al. Arabidopsis G-protein interactome reveals connections to cell wall carbohydrates and morphogenesis. *Mol. Syst. Biol.* **7**, 532 (2011).

Yang, F. et al. Identifying pathogenicity of human variants via paralog-based yeast co-complementation. *PLoS Genet.* **13**, e1006779 (2017).

Jalili, M. et al. Evolution of centrality measurements for the detection of essential proteins in biological networks. *Front. Physiol.* **7**, 375 (2016).

Jeong, H., Mason, S. P., Barabasi, A. L. & Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **409**, 41–42 (2001).

Yu, H., Kim, P. M., Sprecher, E., Trifonov, V. & Gerstein, M. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* **3**, e59 (2007).

Das, J. et al. Exploring mechanisms of human disease through structurally resolved protein interactome networks. *Mol. Biosyst.* **10**, 9–17 (2014).

Crux Adenso, N., Munoz Giner, E., de Groot, N. S. & Torrent Burgas, M. Centrality in the host-pathogen interaction is associated with pathogen fitness during infection. *Nat. Commun.* **8**, 14092 (2017).

Pan, A., Lahiri, C., Rajendiran, A. & Shannumugam, B. Computational analysis of protein interaction networks for infectious diseases. *Brief Bioinform.* **17**, 517–526 (2016).

Mukhtar, M. S. et al. Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science* **336**, 596–601 (2011).

Wessling, R. et al. Convergent targeting of a common host protein-network by pathogen effectors from three kingdoms of life. *Curr. Biol.* **22**, 1359–1369 (2012).

Kuhner, S. et al. Proteome organization in a genome-reduced bacterium. *Science* **333**, 662–666 (2011).

Fontes, E. P., Santos, A. A., Luz, D. F., Waclawovsky, A. J. & Chory, J. The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. *Genes Dev.* **18**, 2545–2556 (2004).

Antal, M. & Balogh, L. Modeling belief systems with scale-free networks. *Neural Netw.* **22**, 1359–1371 (2009).

He, B. J. Scale-free brain activity: past, present, and future. *Trends Cogn. Sci.* **18**, 480–487 (2014).

Ros, T. et al. Neurofeedback tunes scale-free dynamics in spontaneous brain activity. *Cereb. Cortex* **27**, 4911–4922 (2017).

Wang, X. et al. BAK1 is not a target of the Pseudomonas syringae effector AvrPto. *Plant Syst. Microbiol. Interact.* **24**, 100–107 (2011).

Lee, I. C. et al. Age-dependent action of an ABA-inducible receptor kinase, RPK1, as a positive regulator of senescence in Arabidopsis leaves. *Plant Cell Physiol.* **52**, 651–662 (2011).

Jiang, Y. et al. Arabidopsis ABA receptor-like kinases interact with yeast raised not to target the common signaling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. *Cell Host Microbe* **4**, 17–27 (2008).

Rajendiran, A., Lahiri, C., Shannumugam, B., Kumar, A. & Mukhtar, M. S. Computational analysis at Strubbekliger Rezervor Kinas 3 drives immune-triggered incompatibilities between Arabidopsis thaliana accessions. *Nat. Genet.* **42**, 1135–1139 (2010).

Mosher, S. et al. The tyrosine-sulfated peptide receptors PSKR1 and PSY1R modify the immunity of Arabidopsis to biotrophic and necrotrophic pathogens in an antagonistic manner. *Plant J.* **73**, 471–482 (2013).

Xiang, Y. et al. BAK1 is not a target of the Pseudomonas syringae effector AvrPto. *Plant Syst. Microbiol. Interact.* **24**, 100–107 (2011).

Fontes, E. P., Santos, A. A., Luz, D. F., Waclawovsky, A. J. & Chory, J. The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. *Genes Dev.* **18**, 2545–2556 (2004).

Antal, M. & Balogh, L. Modeling belief systems with scale-free networks. *Neural Netw.* **22**, 1359–1371 (2009).

He, B. J. Scale-free brain activity: past, present, and future. *Trends Cogn. Sci.* **18**, 480–487 (2014).

Ros, T. et al. Neurofeedback tunes scale-free dynamics in spontaneous brain activity. *Cereb. Cortex* **27**, 4911–4922 (2017).

Strumpf, M. P. & Porter, M. A. Mathematics. Critical truths about power laws. *Science* **335**, 665–666 (2012).

Kuhner, S. et al. Proteome organization in a genome-reduced bacterium. *Science* **336**, 1235–1240 (2009).

Wang, H. et al. A complex-based reconstruction of the Saccharomyces cerevisiae interactome. *Mol. Cell. Proteomics* **8**, 1361–1381 (2009).

Zotenko, E., Mestre, J., O’Leary, D. P. & Przytycka, T. M. Why do hubs in the yeast protein interaction network tend to be essential: reexamining the connection between the network topology and essentiality. *PLoS Comput. Biol.* **4**, e1000140 (2008).

Yu, H. et al. High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104–110 (2008).

Kafri, R., Dahan, O., Levy, J. & Pilpel, Y. Preferential protection of protein interaction network hubs in yeast: evolved functionality of genetic redundancy. *Proc. Natl. Acad. Sci. USA* **105**, 1243–1248 (2008).

Ilia, A., Shannumugam, B., Lahiri, C., Kumar, A. & Mukhtar, M. Unraveling novel broad-spectrum antibacterial targets in food and waterborne pathogens using comparative genomics and protein interaction network analysis. *Infect. Genet. Evol.* **27**, 300–308 (2014).

Yang, H. et al. Insight into bacterial virulence mechanisms against host immune response via the Yersinia pestis-human protein-protein interaction network. *Infect. Immun.* **79**, 4413–4424 (2011).

Wuchty, S. Computational prediction of host-parasite protein interactions between P. falciparum and H. sapiens. *PLoS ONE* **6**, e26960 (2011).

Goh, K. I. et al. The human disease network. *Proc. Natl. Acad. Sci. USA* **104**, 8685–8690 (2007).

Taylor, I. W. et al. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat. Biotechnol.* **27**, 199–204 (2009).
Acknowledgements
This work was supported by the National Science Foundation (IOS-1557796) to M.S.M., the Alabama State funded Graduate Research Scholars Program (GRSP) to T.C.H, and grants from the Austrian Academy of Science through the Gregor Mendel Institute (Y. B.). Thanks to Dr. Brad Day for providing CD3-1089 and CD3-1096 vectors and positive controls. We greatly appreciate Mr. Benjamin Marsella for technical assistance, and Dr. K. Mukhtar and Ms. J. Lopez for critical reading the manuscript.

Author contributions
M.S.M. conceived the project. M.S.M. and H.A. performed network-based analyses. M.S.M. and Y.B. designed wet lab experiments. M.S.M, T.C.H., Y.S. and N.W. performed wet lab experiments. M.S.M., H.A. T.C.H., Y.S. and Y.B. performed statistical analyses. M.S.M. wrote the first draft of the manuscript. All the authors discussed the results and critically reviewed the manuscript and provided valuable comments/edits.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-04632-8.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.