



HAL
open science

An eco-systems biology approach for modeling tritrophic networks reveals the influence of dietary amino acids on symbiont dynamics of *Bemisia tabaci*

Gopinath Selvaraj, Diego Santos-Garcia, Netta Mozes-Daube, Shlomit Medina, Einat Zchori-Fein, Shiri Freilich

► To cite this version:

Gopinath Selvaraj, Diego Santos-Garcia, Netta Mozes-Daube, Shlomit Medina, Einat Zchori-Fein, et al.. An eco-systems biology approach for modeling tritrophic networks reveals the influence of dietary amino acids on symbiont dynamics of *Bemisia tabaci*. *FEMS Microbiology Ecology*, In press, 10.1093/femsec/fiab117. hal-03330169

HAL Id: hal-03330169

<https://hal.science/hal-03330169v1>

Submitted on 31 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An eco-systems biology based modeling of tritrophic networks reveals the influence of dietary amino acids on symbiont dynamics of *Bemisia tabaci*

Gopinath Selvaraj^{1,2}, Diego Santos-Garcia³, Netta Mozes-Daube², Shlomit Medina¹,
Einat Zchori-Fein² and Shiri Freilich¹

¹Institute of Plant Sciences, Newe Ya'ar Research Center, The Agricultural Research Organization, Ramat Yishay, Israel.

²Institute of Plant Protection, Newe Ya'ar Research Center, The Agricultural Research Organization, Ramat Yishay, Israel.

³Laboratoire de Biométrie et Biologie Evolutive (UMR5558), Université Lyon, Centre National de la Recherche Scientifique (CNRS), 43 Boulevard du 11 Novembre 1918, 69622 Villeurbanne, France

Abstract

Background

Metabolic conversions allow organisms to produce a set of essential metabolites from the available nutrients in an environment, frequently requiring metabolic exchanges among co-inhabiting organisms. Here, we applied genomic-based metabolic simulations for exploring tri-trophic interactions between a sap-feeding insect (*Bemisia tabaci*), its host-plants and its symbiotic (primary, co-primary, secondary) bacteria. The relative simplicity of this ecosystem allows functional simulations that capture both environment and community-level metabolic conversions. Simulations explored the metabolic capacities of the insect together with different symbionts under environments representing several natural phloem contents.

Results

Outcomes depict a puzzle of three-layer origins (plant, insect and symbionts) for the source of essential metabolites across natural habitats and stratify the insect-symbiont interactions behind the phenotypic plasticity that enable the polyphagous life-style. Focusing on amino acid content of the phloem, predictions mapped the potential consumption and production of the insect and its symbionts in different environments. In parallel to simulations, natural and artificial feeding experiments provide a supporting evidence for an environment-based effect on symbiont dynamics. Based on simulations, decrease in relative abundance of a symbiont can be associated with a loss of fitness advantage due to an environmental excess in amino acids whose production used to depend on the symbiont.

Conclusion

The study demonstrates that genomic-based predictions can bridge environment and community dynamics and lead to the design of symbiont manipulation strategies.

[250 word limit; 263 – our abstract]

Key words:

[Einat's comments: Need words that do not appear in the title. Should thus be decided after we agree on the title]

Introduction:

The metabolic capacity of individual organisms to fulfill their nutritional needs in natural environments is a fundamental evolutionary driving force and a key determinant of the function of the entire ecosystem. Using these capacities, organisms can produce a set of metabolites required for their growth and development from the available nutrients in an environment. Typically, the metabolic capacities of an organism *per se* are not sufficient to meet all of its nutritional requirements. Towards fulfilling these requirements, species are frequently involved in metabolic exchanges with other co-inhabiting organisms, which delineate many of the biological interactions and consequently shape corresponding evolutionary adaptations [1]. Such interactions may occur among any living entities including plants, animals and microbes, connecting various trophic strata [2, 3].

Among these relations, insect-bacterial symbiont interactions were well studied in the past two decades [4-7]. In many insect groups, symbiotic interactions became obligatory when an insect host became specialized to feed on nutritionally challenging sources like plant xylem and phloem saps [8-10]. Such an interaction between an ancestral insect species from the suborder Auchenorrhyncha of the order Hemiptera and the bacterium *Candidatus* Sulcia muelleri is estimated to have emerged 270 millions of years ago, marking the beginning of congruent evolution of trophic interactions between insect and vertically transmitted symbionts [8]. Consequently, throughout the evolution, bacterial symbionts have been providing their insect hosts a range of fitness benefits from protection against biotic and abiotic stressors to nutritional amendments [11]. Complementary interactions between the insect and symbionts are an essential trophic adaptation serving the production of the full spectrum of cellular building blocks bestowing fitness advantage for symbiont-harboring insects [12, 13]. One such well-studied insect-symbiont model system is the phloem sap-feeding sweet potato whitefly, *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodidae) [13-17]. The obligate/primary symbiont of this whitefly, *Candidatus* Portiera aleyrodidarum (hereafter *Portiera*) is harbored in designated host cells called 'bacteriocytes' [15, 18, 19]. In

addition to the primary symbiont, significant portion of *B. tabaci* population is co-fixated with either *Candidatus Hamiltonella defensa* (hereafter *Hamiltonella*) or *Candidatus Arsenophonus* (hereafter *Arsenophonus*) in the bacteriocytes [7, 15, 17, 18, 20]. All three symbionts are suggested to complement host's nutrition: *Portiera* is required for the essential amino acids and carotenoids [13, 19]; *Hamiltonella* and *Arsenophonus* produce cofactors and vitamins [7, 18, 19]. Like essential amino acids, the cofactors and the vitamins are highly crucial for the growth and development of insects [21, 22], and sometimes causative metabolites in the evolution of obligatory mutualism between insect and symbionts [23, 24]. Both *Hamiltonella* and *Arsenophonus*, by displaying characteristics of primary symbionts including vertical transmission, progressing genomic reduction and complementing indispensable nutrients [7, 18], can be considered as co-primary. *B. tabaci* can co-harbor additional facultative/secondary symbiotic bacterial genera including *Candidatus Rickettsia* (hereafter *Rickettsia*), *Candidatus Wolbachia* (hereafter *Wolbachia*) and *Candidatus Cardinium* (hereafter *Cardinium*), which form varying combinations and inhabit several whitefly tissues including the bacteriocyte [6, 7, 15]. Together, the whitefly and its bacterial symbionts form a complex termed 'holobiont' - a symbiotic assemblage of host and microbes that live, develop and evolve together [25], with intricate metabolic exchanges spreading along wide range of plant hosts [14, 16, 17].

Trophic interactions are increasingly explored by applying genomics-based approaches, mainly such that focus on exploring metabolic capacities. The genomics data hoard the information not only on the enzymatic profiles, but also on the corresponding gaps in terms of missing genes or enzymes in metabolic pathways. The parallel exploration of co-inhabiting organisms in a holobiont can aid in assembling puzzle pieces of biological interactions among them [13, 14, 16, 26-28]. Computational tools are being developed for processing genomics data into metabolic network presentation, allowing prediction of trophic interactions in systems involving multiple species [28-31]. Metabolic network modeling were previously used to study interactions among microorganisms [16, 27, 28, 30, 32]. On a

multitrophic scale, many studies have explored the symbiotic interactions of sap-feeding insects and their symbionts in general [4, 12, 33, 34], and *B. tabaci* and its symbionts in particular [13, 14]. Though complementing the nutritional needs of the insect host is considered a crucial factor in symbiont fixation [13, 14], to date, the effect of host plant (phloem) environment (the set of metabolites serving as nutritional base for any biological activities and interactions) on insect-symbiont metabolic profile was not yet considered. Most phloem-feeders, including *B. tabaci*, exclusively depend on the phloem's nutritional quality. Relative to other plant parts, the latter is relatively lower in essential amino acids, and its amino acid composition significantly varies from plant to plant [35, 36]. Here, we compiled the nutritional composition of phloem from a range of host plants based on a literature survey. Using genomic-based simulation, we aimed at deciphering the interactions between *B. tabaci* and its bacterial symbionts under specific plant phloem environments using computational network modeling approaches. Simulations, using expansion algorithm [9, 37], were recently applied for exploring interactions between symbionts of *B. tabaci* in bacteriocyte-like environment [16]. Briefly, the expansion algorithm generates a set of all possible metabolites that can be produced given a set of starting compounds (collectively termed as "environment" as described above) and a set of feasible reactions (based on enzymes drawn from genomic data of participating organisms). When networks were completed by combining metabolic capacities of two or more different species (which are possibly from same ecosystem and believed to be interacting), this defines the complementary synthesis of metabolites that can be predicted as outcome of metabolic interactions. In the current study, we took a similar approach to explore insect host-symbiont interactions in phloem-like environments. This extension of the simulations for exploring tri-trophic level interactions among insect, its symbionts and plant is enabled by the availability of complete genomic data of both the whitefly (*B. tabaci*) and its symbionts in addition to documented nutritional compositions of the phloem content of many host plants, forming a well-defined eco-system model. In order to demonstrate the possible significance of genomic-based simulations in deciphering the role of

environment in shaping community structure of symbionts, the metabolic predictions were correlated with experimental data on the dynamics of symbionts in the whitefly under different diet conditions.

Results:

Assessment of amino acid profiles in the phloem of various plants

Since the biochemical compositions of plant phloem directly affect the fitness of sap-feeding insects [38, 39], the amounts and composition of specific amino acids are expected to influence the dynamics of holobionts formed by insects and their symbionts. To provide a base-line assessment of this starting point, a comparison of multiple plant phloem was performed based on literature survey of experimental data. As the majority of studies focused on the amino acid content of the phloem and as literature emphasizes the role of interactions with symbionts in complementing amino acid synthesis [16, 35, 36], we have focused on the amino acids content exclusively and used it hereafter for further analyses. Overall, phloem composition data were collected from 16 independent studies covering 36 plants from 16 families (Table 1). The compiled data set is consistent with previous reports describing phloem contents as poor in essential amino acids and rich in a few specific non-essential ones [35, 36] (Fig. 1.). Among essential amino acids, threonine is most abundant (especially higher in three of Rutaceae members and melon), followed by arginine which is above 10% in poinsettia, cotton and orange jasmine (Fig 1). Phloem content data were clustered into five major groups. Clustering pattern clearly demonstrates a taxonomic signature: cluster I – mainly plants from the Fabaceae family whose phloem is rich in asparagine; cluster II - most plants from the Solanaceae together with few others whose phloem is rich in glutamine; cluster III – Rutaceae family, rich in proline; cluster IV – mainly Poaceae, rich in glutamate; cluster V – Cucurbitaceae and Euphorbiaceae, rich in glycine. Deviating from the taxonomic grouping are one tomato and one potato samples (sly2 and stu2 respectively), both within the Solanaceae, that are not grouped in the glutamine-rich cluster II. Overall, despite the independence of the studies, the diversity of methods used (Table 1), and natural variations in

phloem content during growth and following environmental perturbations, a clear taxonomic signature emerges, where phloem contents are clustered mainly according to the phylogeny of the plant and different families show characteristic differential abundance pattern of amino acids.

Genomics-based predictions of complementary interactions between *B. tabaci* and its primary symbiont *Portiera* in different phloem environments

In a 'holobiont' comprising sap-feeding insect and its symbionts, despite variations in the available environmental resources (here, phloem content), metabolic conversions should enable the production of all essential cellular building blocks (crucial metabolites for the survival and growth of an organism). Here, we explore this metabolic robustness by simulating metabolic activity of one of the sap-feeding insect, *B. tabaci* together with its primary symbiont *Portiera* in different plant phloem environments (Fig. 2). Simulations make use of a network representation of the genomic data – the enzyme coding genes, for predicting the influence of environmental inputs (phloem content) on the production of essential components [16]. In order to represent the natural diversity of habitats that can host the polyphagous whitefly *B. tabaci*, two plants from each of five clusters (Fig. 1) were selected and their phloem content served as approximate environments in this analysis. Complementary interactions between the whitefly and its symbionts were predicted by comparing metabolic capacities of the whitefly alone, to those of a holobiont system, considering the whitefly together with its symbiont under those ten phloem contents (Fig. 2). The synthesis of most essential amino acids including valine, leucine, isoleucine, phenylalanine, tryptophan, lysine, arginine and histidine depends on insect-symbiont interactions under most phloem environments. This is in accordance with previous studies demonstrating that *Portiera* complements the synthesis of essential amino acids to its insect host [13, 14, 18, 19]. *B. tabaci* can independently synthesize the essential amino acid methionine using cysteine as a precursor through successive transformation catalyzed by cystathionine gamma-lyase (Enzyme commission (EC) number: 4.4.1.1), cysteine-S-

conjugate beta-lyase (old: 4.4.1.8/4.4.1.6 & new: 4.4.1.13), homocysteine S-methyltransferase (2.1.1.10). In addition, *Portiera* can also complement methionine production through the end reaction in the above pathway by an alternate enzyme, cobalamin-independent methionine synthase - MetE (2.1.1.14) [13, 16]. The non-essential amino acid tyrosine can be produced independently by *B. tabaci* only when phenylalanine is present in the phloem (for example in tomato and pumpkin - Sly1_P and Cpe; Fig. 2). This independent production of tyrosine by *B. tabaci* is exclusively phenylalanine dependent. The enzyme involved in the direct conversion of phenylalanine to tyrosine is phenylalanine 4-monooxygenase (1.14.16.1), as shown in other insect systems [12, 40], but unlike the pathway in weevil (a type of beetle), in which its symbiont produces tyrosine by phenylalanine-independent pathway [41]. In the absence or deficiency of phenylalanine in the phloem, tyrosine can be produced by complementary interactions with *Portiera*, where it still depends on *B. tabaci* to convert chorismate to prephenate by Chorismate mutase (5.4.99.5), a key step in phenylalanine synthesis. The current analysis illustrates the ability of *B. tabaci* and its symbionts to produce all the required elements across a range of natural environments (phloem) presented by different host plants.

In silico simulations further compared this *B. tabaci-Portiera* complementation pattern, reflecting long evolved mutualism, to an artificial complementation pattern formed between *Portiera* and another insect – the pea aphid *Acyrtosiphon pisum*. Like *B. tabaci*, *A. pisum* is a sap-feeding insect whose nutrition relies on mutualistic interactions with *Buchnera* - an evolutionary-adapted, primary symbiont [4]. Unlike the *B. tabaci-Portiera* complementation profiles, the artificial *A. pisum-Portiera* holobiont could not co-synthesize all amino acids (Fig. 2). The four amino acids that cannot be produced are histidine, lysine, tryptophan and tyrosine. In histidine biosynthetic pathway, histidinol-phosphatase (3.1.3.15) and histidinol dehydrogenase (1.1.1.23) that are found in *B. tabaci*, are missing in *A. pisum*, to which, *Buchnera* supplies the histidine to fulfill the demand. In lysine biosynthesis pathway, there are three enzymes are missing in *Portiera* including 4-hydroxy-tetrahydrodipicolinate reductase (1.17.1.8), diaminopimelate epimerase (5.1.1.7) and diaminopimelate decarboxylase

(4.1.1.20), all found in *B. tabaci*, but absent in *A. pisum*. Unlike *Portiera*, *Buchnera* can able to synthesize and complement lysine [4]. In the phenylalanine biosynthesis pathway, *Portiera*'s genome lacks the enzyme which transforms chorismate to prephenate, and relies on the *B. tabaci* enzyme chorismate mutase, for complementing the production of the amino acid. In contrast, in the *A. pisum*-*Buchnera* holobiont, the enzyme is found in the symbiont and not in the insect. Hence in the artificial *A. pisum*-*Portiera* holobiont, there is no predicted pathway for the production of tyrosine when phenylalanine is missing in the phloem. All the above enzymes from *B. tabaci* (except those from histidine pathway) are acquired by horizontal gene transfer from other bacteria during evolution to compensate for the gene loss in *Portiera* [13]. Over all, this comparison between natural and artificial tritrophic interactions demonstrates an orchestration of metabolism evolving to ensure the full production of cellular building blocks (here amino acids) in dynamic environments.

Genomics-based predictions of complementary interactions between *B. tabaci* and its co-primary and secondary symbionts

In addition to its primary symbiont *Portiera*, *B. tabaci* harbors varying combinations of co-primary/secondary symbionts, mainly, *Hamiltonella*, *Arsenophonus*, *Rickettsia*, *Wolbachia* and *Cardinium*, though not all together [17]. In accordance with previous reports [13, 14, 19], the genomics based analysis demonstrated the capacity of *Portiera* to complement the production of essential amino acids in plant phloem - *B. tabaci*'s natural environment (Fig. 2). Using the same simulation procedure, we tested the metabolic performances of holobionts formed by *B. tabaci* and each of its secondary symbionts. Genomic reconstructions relied on previously published metabolic networks of all symbionts [16] with the exception of *Arsenophonus* genome that was assembled *de novo* (Methods). Simulations were carried in several natural phloem environments and outcomes are shown in phloem contents representing broad bean and cotton (Alf - Fabaceae and Ghi2 – Malvaceae, respectively; Fig 3). Unlike the primary symbiont *Portiera*, none of the co-primary/secondary symbionts complements the whitefly's production of amino acids, considering all phloem environments

tested (Fig. 3). However, *Arsenophonus* and *Hamiltonella* possess some independent pathways for the production of few essential and non-essential amino acids (Fig. 3).

We further screened for the capacity of these holobiont combinations to produce additional metabolites that are considered essential cellular building blocks including lipids, co-factors and vitamins [30, 42]. Simulations reflected the reported role of different symbionts. Though *Hamiltonella* and *Arsenophonus* do not complement the synthesis of essential amino acids, both are found to synthesize co-factors and vitamins, including the previously suggested synthesis of riboflavin, FAD and NAD [18]. Overall, *Hamiltonella* and *Arsenophonus* (co-primary symbionts) are predicted to have broader production profile of essential cellular building blocks in comparison to *Rickettsia*, *Wolbachia* and *Cardinium* (secondary ones). The symbionts' production profile can be associated with the well-documented distribution profiles as widely-preferred to intermittently-preferred ones in *B. tabaci* populations and sub-populations. The co-primary symbiont *Hamiltonella* or *Arsenophonus* is fixed together with *Portiera* in most of the *B. tabaci* populations, unlike the sporadically distributed secondary symbionts [17, 20]. Except in very rare cases, *Hamiltonella* and *Arsenophonus* have exclusive distribution pattern and are not found together, possibly reflecting their functional overlap that can be associated with competition for the common resources and not providing sufficient fitness benefit that is justifiable in terms of energetic cost for co-hosting both [7, 17, 20]. The stratification of symbionts based on their metabolic profiles detected in the present study (Fig. 3) is also well correlated with the previously reported physical distribution patterns of the symbionts in the insect host with co-primary symbionts located exclusively inside the bacteriocytes and the secondary ones typically distributed sparsely in different tissues and rarely inside the bacteriocytes [7, 15]. These production profiles, also consistent in other plants (Example: Additional file 1: Figure S2), suggest that metabolic capacities play a role in determining the distribution profile of symbionts and symbiont fixation is oriented towards reaping maximum benefits for the 'holobiont' during the evolution.

Testing the influence of dietary variations on symbiont dynamics

Both *Hamiltonella* and *Arsenophomus* possess independent pathways for the production of few amino acids (Fig. 3). The relevance of these pathways depends on the phloem content. For example, though *Hamiltonella* possess an independent pathway for asparagine production (Fig. 3), such capacity is likely to be relevant in the asparagine-poor phloem of cotton (~5%) and not in the asparagine-rich phloem of broad bean (~50%) (Percentage compositions of plant phloems are in Additional file 2). To explore the impact of variations in phloem content on community dynamics, the relative abundance of *B. tabaci* symbionts was assessed following rearing on two different host plants – broad bean and cotton. These specific plants were chosen based on the literature review, which shows that the phloem content of broad bean is dominated by asparagine, whereas that of cotton is more diverse with several amino acids, but, in relatively low concentrations (Fig. 1). *B. tabaci* MEAM1 biotype species from a population carrying *Hamiltonella* and *Rickettsia* (B-PHR) or just *Hamiltonella* (B-PH) were transferred from cotton host to broad bean for long (77 days, about 3 generations; B-PHR line) or short (20 days, one generation; B-PH) terms of rearing (Methods). The change in *B. tabaci* line is due to the loss of *Rickettsia* in the lines maintained in the lab (Methods). Transfer to the broad bean plant required about a month of adjustment in order to recover to the size of the original population. In both experiments, the relative abundances of *Portiera* and *Hamiltonella* was reduced by approximately two fold following the transfer to broad bean host, though the reduction is significant only in the short-term experiment; *Rickettsia* density remained unchanged (Fig. 4). Considering the composition of phloem content in broad bean (Fig. 1), it was hypothesized that asparagine is a dominant element in conferring the observed reduction in the relative abundance of *Portiera* and *Hamiltonella*. A possible explanation for the reduction of *Hamiltonella* can be related to the transfer to the asparagine-rich environment, where the asparagine producing *Hamiltonella* loses its competitive advantage (Fig. 5). Reduction in *Portiera* can be related to possible dependency of *Portiera* on *Hamiltonella* for cofactors [13] and presumably other intricate exchanges, as part of an adaptation process towards a fixed co-inhabitation inside the

bacteriocytes. To directly explore the impact of asparagine concentrations on community dynamics, the relative abundance of *B. tabaci* symbionts was assessed in controlled artificial-feeding systems where media were augmented with seven different amino acids, but, single amino acid at a time (Methods). Out of seven amino acids tested, only asparagine has induced a significant reduction of about two fold in the relative abundance of *Portiera* and *Hamiltonella* (Fig. 4), supporting it as the causative factor in inducing reduction in the plant system. The only other modification in community structure was induced by the amino acid serine, leading to a reduction in the relative abundance of *Rickettsia*, but not of *Portiera* and *Hamiltonella*. Similar to the reduction in *Hamiltonella* in asparagine supplemented medium, the serine-induced reduction in *Rickettsia* can be related to a loss in competitive advantage as *Rickettsia* can able to produce serine when absent or scarce in the phloem (Additional file 1: Figure S2).

Discussion:

The current study attempts to delineate multitrophic networks in a natural ecosystem by coupling genomics-based computational methods with experimental-based knowledge. The chosen tritrophic system explored here consists of a plant host, a sap-feeding insect (the whitefly, *B. tabaci*) and its symbionts (primary, co-primary and secondary). The plant *via* its phloem content serves as the sole nutrimental source for the insect and the symbionts functioning as single unit, termed holobiont. The relative simplicity of this ecosystem in comparison to other natural habitat allows to cover the two crucial components required for simulating metabolic activity: (1) environment – available nutrients; (2) set of reactions – inferred from enzyme coding genes of the holobiont. These components were retrieved from experimental-based reports on the metabolic profiles of phloem content of different plants (environment) and from the available genome data of the holobiont (reactions), respectively. The extensive literature survey of phloem content supports characteristic differences in the amino acid compositions among plant families.

Outcome of the simulation depicts a puzzle of three-layer origins (plant, insect and symbionts) of amino acids considering different host plants. Most of the non-essential amino acids are either available in the phloem or can be synthesized by *B. tabaci*; the production of most essential amino acids is carried by primary symbiont, *Portiera*; the production of cofactors, vitamins and lipids is carried by the co-primary symbionts (either *Hamiltonella* or *Arsenophonus*); remaining secondary symbionts are predicted to synthesize few components. The genomic based predictions for *B. tabaci* holobiont system are consistent with reports on both the supply of phloem nutrients and bacterial symbionts for producing the entire spectrum of required components [10, 11]. The metabolic robustness of the insect-symbiont unit thus reflected by its capacity of producing the full set of amino acids across all natural environments tested. This demonstration of the holobiont as self-sufficient in terms of amino acid production, may partially explain the phenotypic plasticity behind the polyphagous life-style of many sap-feeding insects [35]. This plasticity is not only useful during alternation of plant hosts, but, can also be an adaptive feature to changes in the amino acid composition of the same host plant, for example, during various developmental phases [43]. Simulations provide two lines of supporting evidence for the role of functional adaptation in fixating the symbiotic interactions: (1) the primary symbiont *Portiera* is uniquely adapted to its host *B. tabaci* but not to other sap-feeding insects, reflecting the long-term host-specific evolutionary orchestration of metabolism [44, 45]; (2) the two co-primary symbionts – *Hamiltonella* and *Arsenophonus*, fulfill similar functional roles in this holobiont system, providing a possible explanation for their mutual exclusive distribution pattern [17, 20, 46].

In addition to capturing previously reported system-level interactions among the insect and its primary and co-primary symbionts [7, 13, 16, 18, 19], predictions highlight alternative routes for the production of some of the amino acids that can be independently synthesized by symbionts, in parallel to the insect (Fig. 3). Such redundant pathways might carry fitness advantage by providing an additional route for controlling the production rate of key compounds. The alternative production routes carried by the co-primary/secondary

symbionts might reflect a need for compensating low efficiency production of these compounds by the insect and/or its primary symbiont. Notably, during long-term maternal transmission primary symbionts tend to undergo extensive fixation of deleterious mutations, which limits the efficiency of their gene products [47, 48]. Such compensations by alternate routes bestow a fitness advantage to the holobiont. For example, synthesis of asparagine by *Hamiltonella*, may fulfill the demand of the holobiont in light of the low activity of the asparagine synthase (EC: 6.3.5.4) of *B. tabaci* [14].

Both whole plant and artificial feeding experiments showed that the level of specific amino acids in the diet effects the dynamics of the symbiont community (Fig. 4). In particular, the relative abundance of *Hamiltonella* and *Portiera*, but not *Rickettsia* was shown to be negatively correlated with asparagine levels (Fig. 4). The decreased relative abundance of *Hamiltonella* in the presence of asparagine excess can be related to a loss in fitness advantage due to the bacterium's inherent capacity to produce this amino acid. The parallel reduction in the relative abundance of *Portiera* can be explained by the intricate interactions between the primary and co-primary symbionts. Recent works are providing increasing evidence for an unique and direct cell to cell communication that allows bacterial species to exchange of metabolites, proteins and even plasmids [49]. Several metabolic dependencies of *Portiera* on *Hamiltonella* were previously suggested, including the supply of cofactors [13], as well as other key metabolites such as amino acyl-tRNAs [45] and 10-formyltetrahydrofolate (Fig. 3), indispensable for the initiation of protein synthesis and purine biosynthesis in prokaryotes [50]. All of these dependencies, which are due to the extremely reduced genome of *Portiera* may explain the associated reduction along with *Hamiltonella*. Other than asparagine, serine was the the only amino acid (out of seven tested) for which the supply in excess has induced a change in the relative abundance of a symbiont; a high amount of serine was correlated with a significant reduction in the level of *Rickettsia* (Fig. 4). The same line of reasoning applies here. Serine can be produced by *Rickettsia*, but not by *Portiera* and *Hamiltonella* (Additional file 1: Figure S2), hence its supply in excess might challenge the fitness advantage. Both

asparagine and serine are common non-essential amino acids available in plant phloem, which might play a central role in the community dynamics of symbionts. However, the non-essential amino acid glutamine does not induce a change in any of the symbiont levels, though like asparagine, it can be produced by *Hamiltonella*. Hence, our analysis cannot lead to a formulation of consistent paradigm connecting the environment and community dynamics.

Conclusions:

Over all, the results demonstrate the usefulness of coupling genomics with environment (nutritional source) data and subsequent network simulations for depicting metabolic exchanges in natural ecosystems. This metabolic construction allowed capturing known interactions yet providing an ecosystem-level perspectives comparing the metabolic profiles of different holobionts in different environments. Such wide-scale comparative analysis can pinpoint at environmental components that might provide fitness advantage to specific symbionts. Correlating predictions with experimental results provides possible interpretation of observations, though at the moment no clear ecological principle can be deduced. Better coverage of environment, including quantitative information on metabolite content together with transcriptomic data and a wider array of experimental results on community dynamics under a range of conditions is likely to promote the formulation of generic fitness rules. Ultimately, the present study envisions a future-oriented approach for establishing ecosystems level functional knowledge from available 'omics (genomics and metabolomics) data and paves way for *in silico* systematic exploration of complex multitrophic interactions.

Materials and methods:

Compilation and clustering of amino acid content from the phloem of different plants

Phloem content of 36 plant species from 16 families were used for this analyses. The quantitative data for each phloem content were collected and standardized (Table 1), and raw data are available in Additional file 2. Heatmap of the normalized phloem contents was generated using pheatmap R package [51] with default parameters ("Euclidean" measure for distance calculation and "complete" for clustering method). Confidence level of branches in

the dendrogram was determined using pvclust R package [52]. The confidence scores for each branch were shown in Additional file 1: Figure S1.

Genome annotations and metabolic reconstruction

Information on the genomic sequence and number of identified EC accessions for symbionts and insects are indicated in Table 2. The final EC lists of each insect and symbiont are provided in Additional file 3. Briefly, metabolic reconstructions of *B. tabaci*: *Portiera*, *Hamiltonella*, *Rickettsia*, *Wolbachia* and *Cardinium* were retrieved from our previous work [16]; annotations are based on the IMG/M pipeline of the JGI [53]. Metabolic reconstruction of the *B. tabaci* symbiont *Arsenophonus* was based *de novo* genome assembly and annotation as described in the sub-section below, and was based on tools implanted in the pipeline. Genome annotations of the *A. pisum* symbiont *Buchnera* were directly downloaded from the JGI's IMG/M website. EC annotations for the genomes of *A. pisum* and *B. tabaci* were retrieved from the Kyoto Encyclopedia of Genes and Genomes database [54]. Annotations for *A. pisum* were directly retrieved from the 'genes' directory (organism code api). Annotations for *B. tabaci* were generated with the KEGG annotation tool GhostKOALA [55] using protein coding sequences as input. Sequences were downloaded from the whitefly genome database (<ftp://www.whiteflygenomics.org/pub/MEAM1/MEAM1/v1.2/>).

DNA extraction, sequencing, assembly and annotation of the genome sequence of *B. tabaci* symbiont *Arsenophonus*

Total genomic DNA was extracted from bacteriomes dissected from laboratory population of *Bemisia*. MED species (Q2 biotype) whose cytotype strain was determined as AWR: *Arsenophonus* sp., *Wolbachia* sp. and *Rickettsia* sp. Fourth instar nymphs (“red eye pupae”) were detached from the leaf and transferred to a 2% agar plate. Amplified genomic DNA was obtained using single bacteriomes. In brief, single bacteriomes were dissected with glass micro-needles, transferred to 10 µl of cold lysis buffer (400 mM KOH, 10 mM EDTA, and 100 mM DTT), and left for 10 min on ice. Before Whole Genome Amplification (WGA),

10 µl of neutralization buffer (400 mM HCl, 600 mM Tris-HCl pH 7.5) was added to lysed bacteriomes. Finally, ten WGA were performed according to manufacturer instructions (Illustra GenomiPhi V2, GE Healthcare), pooled, and sent for NextSeq 2x150 bp (Illumina) sequencing at DNA Services Facility (University of Illinois, Chicago) and MinION R9.4.1 (Oxford Nanopore Technologies) sequencing at Darwin Bioprospecting.

Illumina raw reads were quality screened with FASTQC v0.11.8 and trimmed/clipped using Trimmomatic v0.36 paired-end mode (ILLUMINACLIP:TruSeq2-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50). Second, a Kraken2 custom database [56], including *B. tabaci* MEAM1 and all available whitefly's symbionts genomes, was used to select Illumina high quality and Nanopore raw reads classified as *Arsenophonus*. *Arsenophonus* Nanopore reads were trimmed/corrected and assembled with Canu v1.7.1 (genomeSize=1.2m minReadLength=500) [57]. Third, *Arsenophonus* Illumina and Nanopore high quality reads were assembled with Spades v3.12.0 (single-cell and careful mode turned on), using the obtained Canu assembly as untrusted contigs [58]. Obtained contigs were scaffolded with SSPACE-LongRead [59] and gap-filled with Gapfiller v1.10 [60] using Nanopore and Illumina high quality reads, respectively. Finally, Usearch v10 (-usearch_global -query_cov 0.8 -strand both -id 0.96) [61] was used to globally align Illumina high quality and Nanopore raw reads against the contigs. Recovered reads were used, together with previously selected reads, to scaffold/gapfill the contigs. These steps were repeated until no more reads or scaffolds/gaps were recovered/filled. The final assembly was corrected with Bowtie2 v2.3.5.1 [62] and Pilon v.1.21 (- -jumps) [63] using Illumina high quality reads. Total assembly size is 1.86 Mbp with 328 contigs (N50 - 10.7 Kbp; L50 - 44 Kbp; average Illumina and Nanopore coverage of 93X and 5X, respectively).

Genome assembly quality was assessed with CheckM v1.0.8 [56] obtaining a 91.96% completeness value (similar to other *Arsenophonus* with close genome sizes) (Additional file 4). Genome annotation was performed with prokka v1.12 (---gram neg --mincontiglen 500 -rfam) [57], InterProScan v5.31 [58] and Interpro database v70.0 [59], and PRIAM v2.0 [60]

with default parameters and the enzyme-specific profiles release of January 2018. Overall 2412 CDS were predicted mapped to 704 unique EC accessions (Table 2). RAW Illumina and Nanopore sequences and *Arsenophonus* genome assembly were deposited at ENA of EMBL-EBI under project number PRJEB35780.

Metabolic activity simulations of tritrophic network

Metabolic activity simulations were carried using the expansion algorithm as described in [16]. Briefly, this algorithm can predict the active metabolic network (expanded) given a pre-defined set of substrates and reactions. The pre-defined set of substrates represents available environmental components, which are phloem amino acid compositions in this study. The pre-defined list of reactions is composed by the sets of reactions encoded by the genomes of either the insect or each symbiont or by both together. Although the analysis is qualitative [16], a threshold of 5% was set in order to select amino acids in plant phloem whose quantity presumably fulfilling the insect's nutrition. In accordance with the previous reports [35, 36], this threshold allows to exclude most essential amino acids that are assumed to be nutritionally limiting for sap-feeding insects; yet the threshold is permissive enough to reflect amino acid diversity in phloem content.

Determination of the effect of nutrients on symbiont dynamics

Insect rearing

Bemisia tabaci MEAM1 colonies were reared on cotton (*Gossypium hirsutum* 'Acala') under standard greenhouse conditions: 26 ± 2 C, 60% relative humidity, and a photoperiod of 14L:10D. At the commencement of the study, the presence of *Hamiltonella* and *Rickettsia* was fixed in the population [64, 65], but during the work, the *Rickettsia* lost completely.

Plant feeding experiments

From the ongoing colony on cotton plants, *B. tabaci* individuals were transferred to broad bean (*Vicia faba*) and also to fresh cotton plants located in isolated cages (dimensions: W24.5 x D24.5 x H63 cm; mesh size: 160 µm). Following transfer, whiteflies were grown in

both plants for short (20 days; about one generation) and long (77 days; about 3 generations) growth periods. The long term experiment was performed while the *B. tabaci* line contained *Portiera*, *Hamiltonella* and *Rickettsia*, whereas the short term experiments was done while the *B. tabaci* line contained only *Portiera* and *Hamiltonella*.

Artificial feeding experiment

To provide the whiteflies with known concentrations of various amino acids, *B. tabaci* adult females (24 to 72 hrs old) were collected from ongoing colony on cotton plants and transferred to a clean 200 ml glass bottle with wide mouth. Experiments were carried out in glass bottles in order to allow free movement of the whiteflies unlike patterns observed in plastic tubes, possibly due to electrostatic force. Otherwise, the artificial feeding scheme generally followed as described in Thompson (2006) [66]. Briefly, the open neck of the bottle was covered with two parafilm layers that contained liquid medium from where the whiteflies feed (Additional file 1: Figure S3). Successful feeding behavior in this set up was confirmed initially with the use of food-grade colorants in the medium, which was further avoided in actual experiments. The bottle was covered with black cloth on the sides so that the light can pass only through the parafilm layers. The basic medium contained 20% sucrose in filter sterilized (0.22 μ syringe filter) tap water (which was prepared in bulk to be used for all the experiments). The medium was supplemented with different amino acids at specific concentration, a single amino acid at a time. The amino acid concentrations were chosen based on the work of Thompson (2006), which are non-lethal to the whiteflies [66]. No toxic effect in the tested concentrations of amino acids or controls was observed (Additional file 1: Table S1). The whiteflies were allowed to feed for three days, after which they were collected for further analysis.

Relative quantification of symbionts

The relative abundance of the symbionts in the whiteflies was determined using real-time quantitative PCR, following the protocol described by Caspi-Fluger *et al.*, (2011) [64]. Individual whiteflies were collected and then subjected to genomic DNA extraction using

Nucleospin Tissue kit (Machery-Nachel, Germany). DNA quantification of individual samples was done using Qbit 2.0 (Invitrogen Corporation, CA, USA) and the concentrations of the samples were equalized before the amplifications. Amplifications were performed with primers designed for *Portiera*, *Hamiltonella* and *Rickettsia* (Table 3) using 1× Fast SYBR Green PCR mix (Applied Biosystems) and a StepOnePlus real-time system (Applied Biosystems). The cycling conditions were: 20s activation at 95°C, 40 cycles of 3s at 95°C, 30s at 60°C and 25s at 72°C. Melting ramp from 60-99°C, rising by 1°C at each step, and waiting 5s after each step. Channel source 470 nm, detector 510 nm. An insect β -actin gene was used as the reference gene and the relative abundance of the symbionts was assessed using the $2^{-\Delta\Delta C_T}$ method [67].

Table 1. Description of sources of data from which the content of phloem of different plants was retrieved. Actual quantitative data of these phloem contents are available in Additional file 2 and visualized in Fig. 1.

Family	Plant Latin name	Plant common name	Short name [Ref] - Method [†]
Amaranthaceae	<i>Spinacia oleracea</i>	Spinach	Sol [68] - C
	<i>Chenopodium hybridum</i>	Maple-leaved goosefoot	Chy[35] - B

	<i>Chenopodium album</i>	Pigweed	Cal[35] - B
Asteraceae	<i>Arctium minus</i>	Lesser burdock	Ami [35] - B
	<i>Cirsium arvense</i>	Creeping thistle	Car[35] - B
Brassicaceae	<i>Arabidopsis thaliana</i>	Thale cress/mouse-ear cress	Ath[35] - B
Cucurbitaceae	<i>Cucurbita pepo</i>	Pumpkin	Cpe[69] - B
	<i>Cucumis melo</i>	Melon	Cme[70] - B
Euphorbiaceae	<i>Euphorbia pulcherrima</i>	Poinsettia	Epu[69] - B
Fabaceae	<i>Vicia faba</i>	Broad bean	Vfa1 [36] - A Vfa2 [35] - B
	<i>Trifolium pratense</i>	Red clover	Tpr[36] - A
	<i>Medicago sativa</i>	Alfalfa	Msa[36] - A
	<i>Pisum sativum</i>	Pea	Psa[36] - A
	<i>Trifolium repens</i>	White clover	Tre [35] - B
Lamiaceae	<i>Lamium purpureum</i>	Red dead-nettle	Lpu[35] - B
Malvaceae	<i>Gossypium hirsutum</i>	Cotton*	Ghi1 [71] - B
			Ghi2 [71] - B
Papaveraceae	<i>Papaver dubium</i>	Long-headed poppy	Pdu[35] - B
Plantaginaceae	<i>Plantago lanceolata</i>	Ribwort plantain	Pla[35] - B
Poaceae	<i>Zea mays</i>	Maize/Corn	Zma[72] - C
	<i>Triticum aestivum</i>	Wheat	Tae [73] - C
	<i>Oriza sativa</i>	Rice	Ori [74] - C
	<i>Hordeum vulgare</i>	Barley	Hvu[75] - C
Polygonaceae	<i>Rumex acetosa</i>	Common sorrel	Rac[35] - B
	<i>Rumex obtusifolius</i>	Broad-leaved dock	Rob [35] - B
	<i>Rumex acetosella</i>	Red sorrel	Rla[35] - B
Rutaceae	<i>Citrus × sinensis</i>	Sweet orange1	Csi1 [76] - D
		Sweet orange2	Csi2 [77] - D
	<i>Murraya paniculata</i>	Orange jasmine	Mpa[77] - D
	<i>Bergera koenigii</i>	Curry leaf tree	Bko[77] - D
Solanaceae	<i>Solanum lycopersicum</i>	Tomato	Sly1[78] - A
			Sly2[78] - B
	<i>Nicotiana tabacum</i>	Tobacco	Nta[79] - E
	<i>Solanum tuberosum</i>	Potato**	Stu1 [43] - B
Stu2 [43] - B			
<i>Solanum nigrum</i>	Black nightshade	Sni[35] - B	
Tropaeolaceae	<i>Tropaeolum majus</i>	Garden Nasturtium	Tma[35] - B
Violaceae	<i>Viola arvensis</i>	Field pansy	Var[35] - B

[†]Description of methods: A - stylet exudation; B - EDTA-enhanced exudation; C - insect laser technique; D - centrifugation; E - shallow incision and collection with microcapillary tubes.

*/**Plants from the same study following different treatments/developmental stages, respectively:

*Ghi1 and Ghi2 were sampled from cotton leaves with high and low nitrogen fertility, respectively;

**Stu1 and Stu2 were samples from potato leaves collected in 5th and 10th weeks, respectively.

Table 2. General details on the genomes of the primary, co-primary and secondary symbionts, and insect hosts used in the present study.

Symbiont genus	Host species	Accession	Number of ECs
<i>Portiera</i>	<i>Bemisia tabaci</i> MED-Q1-Spain	CP003835.1	100

<i>Cardinium</i>	<i>Bemisia tabaci</i> MED-Q1-Spain	GCA_000689375.1	112
<i>Hamiltonella</i>	<i>Bemisia tabaci</i> MEAM1-USA	GCA_002285855.1	434
<i>Arsenophonus</i>	<i>Bemisia tabaci</i> MED-Q2-Israel	PRJEB35780	704
<i>Rickettsia</i>	<i>Bemisia tabaci</i> MEAM1-USA	GCA_002285905.1	264
<i>Wolbachia</i>	<i>Bemisia tabaci</i> MED-Q2-Israel	GCA_900097055.1	253
<i>Buchnera</i>	<i>Acyrtosiphon pisum</i>	BA000003.2 (JGI)	298
Insect host		Accession	Number of ECs
<i>Acyrtosiphon pisum</i>		GCF_005508785.1	831
<i>Bemisia tabaci</i> MEAM1		GCA_001854935.1	893

Table 3. Description of the primers used for real-time qPCR.

Gene Primer set	Primer set	Nucleotide sequence	Expected size (bp)	Reference
<i>B. tabaci</i> β actin	Wf-B actin-F Wf-B actin-R	TCTTCCAGCCATCCTTCTTG CGGTGATTCCTTCTGCATT	~200	[80]
<i>Portiera</i> 16S rRNA	Port73-F Port266-R	GTGGGGAATAACGTACGG CTCAGTCCCAGTGTGGCTG	~195	[64]
<i>Hamiltonella</i> dnaK	dnaK-F dnaK-R	GGTTCAGAAAAAAGTGGCAG CGAGCGAAAGAGGAGTGAC	~200	[81]
<i>Rickettsia</i> gltA	glt375-F glt574-R	TGGTATTGCATCGCTTTGGG TTTCTTTAAGCACTGCAGCACG	~200	[64]

Acknowledgements:

The study was supported by the Israel Science Foundation, grant no. 484/17 and by a grant from the Ministry of Science & Technology, Israel & the France's Centre National pour la Recherche Scientifique (CNRS) to SF and EZF.

References

- Großkopf T, Soyer OS. Synthetic microbial communities. *Curr Opin Microbiol.* 2014;18:72-77.
- Turlings TCJ, Erb M. Tritrophic interactions mediated by herbivore-induced plant volatiles: mechanisms, ecological relevance, and application Potential. *Annu Rev Entomol.* 2018;63:433-452.
- Scherber C, Eisenhauer N, Weisser WW, Schmid B, Voigt W, Fischer M, et al. Bottom-up effects of plant diversity on multitrophic interactions in a biodiversity experiment. *Nature.* 2010;468:553-556.
- Douglas AE. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annu Rev Entomol.* 1998;43:17-37.
- Hansen AK, Moran NA. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proc Natl Acad Sci U S A.* 2011;108:2849.
- Marubayashi JM, Kliot A, Yuki VA, Rezende JAM, Krause-Sakate R, Pavan MA, et al. Diversity and localization of bacterial endosymbionts from whitefly species collected in Brazil. *PLoS ONE.* 2014;9:e108363.
- Santos-Garcia D, Juravel K, Freilich S, Zchori-Fein E, Latorre A, Moya A, et al. To B or not to B: Comparative genomics suggests *Arsenophonus* as a source of B vitamins in whiteflies. *Front Microbiol.* 2018;9:2254.
- Moran NA, McCutcheon JP, Nakabachi A. Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet.* 2008;42:165-190.

9. Zientz E, Dandekar T, Gross R. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiol Mol Biol Rev.* 2004;68:745.
10. Douglas AE. Phloem-sap feeding by animals: problems and solutions. *J Exp Bot.* 2006;57:747-754.
11. McLean AHC, Parker BJ, Hrček J, Henry LM, Godfray HCJ. Insect symbionts in food webs. *Philos Trans Royal Soc B.* 2016;371:20150325.
12. Hansen AK, Moran NA. The impact of microbial symbionts on host plant utilization by herbivorous insects. *Mol Ecol.* 2014;23:1473-1496.
13. Luan J-B, Chen W, Hasegawa DK, Simmons AM, Wintermantel WM, Ling K-S, et al. Metabolic coevolution in the bacterial symbiosis of whiteflies and related plant sap-feeding insects. *Genome Biol Evol.* 2015;7:2635-2647.
14. Ankrah NYD, Luan J, Douglas AE. Cooperative metabolism in a three-partner insect-bacterial symbiosis revealed by metabolic modeling. *J Bacteriol.* 2017;199:e00872-00816.
15. Gottlieb Y, Ghanim M, Gueguen G, Kontsedalov S, Vavre F, Fleury F, et al. Inherited intracellular ecosystem: symbiotic bacteria share bacteriocytes in whiteflies. *The FASEB Journal.* 2008;22:2591-2599.
16. Opatovsky I, Santos-Garcia D, Ruan Z, Lahav T, Ofaim S, Mouton L, et al. Modeling trophic dependencies and exchanges among insects' bacterial symbionts in a host-simulated environment. *BMC Genomics.* 2018;19:402.
17. Zchori-Fein E, Lahav T, Freilich S. Variations in the identity and complexity of endosymbiont combinations in whitefly hosts. *Front Microbiol.* 2014;5:310-310.
18. Rao Q, Rollat-Farnier P-A, Zhu D-T, Santos-Garcia D, Silva FJ, Moya A, et al. Genome reduction and potential metabolic complementation of the dual endosymbionts in the whitefly *Bemisia tabaci*. *BMC Genomics.* 2015;16:226.
19. Santos-Garcia D, Farnier P-A, Beitia F, Zchori-Fein E, Vavre F, Mouton L, et al. Complete genome sequence of “*Candidatus Portiera aleyrodidarum*” BT-QVLC, an obligate symbiont that supplies amino acids and carotenoids to *Bemisia tabaci*. *J Bacteriol.* 2012;194:6654-6655.
20. Skaljac M, Zanic K, Ban SG, Kontsedalov S, Ghanim M. Co-infection and localization of secondary symbionts in two whitefly species. *BMC Microbiol.* 2010;10:142.
21. Douglas AE. The B vitamin nutrition of insects: the contributions of diet, microbiome and horizontally acquired genes. *Curr Opin Insect Sci.* 2017;23:65-69.
22. House HL. Effects of different proportions of nutrients on insects. *Entomol Exp Appl.* 1969;12:651-669.
23. Nikoh N, Hosokawa T, Moriyama M, Oshima K, Hattori M, Fukatsu T. Evolutionary origin of insect–*Wolbachia* nutritional mutualism. *Proc Natl Acad Sci U S A.* 2014;111:10257.
24. Moriyama M, Nikoh N, Hosokawa T, Fukatsu T. Riboflavin provisioning underlies fitness contribution to its insect host. *mBio.* 2015;6:e01732-01715.
25. Rosenberg E, Zilber-Rosenberg I. Symbiosis and development: The hologenome concept. *Birth Defects Res C Embryo Today.* 2011;93:56-66.
26. Freilich S, Kreimer A, Meilijson I, Gophna U, Sharan R, Ruppim E. The large-scale organization of the bacterial network of ecological co-occurrence interactions. *Nucleic Acids Res.* 2010;38:3857-3868.
27. Levy R, Borenstein E. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proc Natl Acad Sci U S A.* 2013;110:12804.
28. Stolyar S, Van Dien S, Hillesland KL, Pinel N, Lie TJ, Leigh JA, et al. Metabolic modeling of a mutualistic microbial community. *Mol Syst Biol.* 2007;3:92.
29. Handorf T, Ebenhöf O, Heinrich R. Expanding metabolic networks: scopes of compounds, robustness, and evolution. *J Mol Evol.* 2005;61:498-512.
30. Kreimer A, Doron-Faigenboim A, Borenstein E, Freilich S. NetCmpt: a network-based tool for calculating the metabolic competition between bacterial species. *Bioinformatics.* 2012;28:2195-2197.
31. Duan G, Christian N, Schwachtje J, Walther D, Ebenhöf O. The metabolic interplay between plants and phytopathogens. *Metabolites.* 2013;3:1-23.
32. Röttgers L, Faust K. From hairballs to hypotheses—biological insights from microbial networks. *FEMS Microbiol Rev.* 2018;42:761-780.
33. Moran NA, Tran P, Gerardo NM. Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum *Bacteroidetes*. *Appl Environ Microbiol.* 2005;71:8802.

34. Gil R, Latorre A. Unity makes strength: a review on mutualistic symbiosis in representative insect clades. *Life*. 2019;9:21.
35. Wilkinson TL, Douglas AE. Phloem amino acids and the host plant range of the polyphagous aphid, *Aphis fabae*. *Entomol Exp Appl*. 2003;106:103-113.
36. Sandstrom J, Pettersson J. Amino acid composition of phloem sap and the relation to intraspecific variation in pea aphid (*Acyrtosiphon pisum*) performance. *J Insect Physiol*. 1994;40:947-955.
37. Ebenhoh O, Handorf T. Structural analysis of expanding metabolic networks. *Genome informatics*. 2004;15:35-45.
38. Corcuera LJ. Biochemical basis for the resistance of barley to aphids. *Phytochemistry*. 1993;33:741-747.
39. Salvucci ME, Wolfe GR, Hendrix DL. Effect of sucrose concentration on carbohydrate metabolism in *Bemisia argentifolii*: Biochemical mechanism and physiological role for trehalulose synthesis in the silverleaf whitefly. *J Insect Physiol*. 1997;43:457-464.
40. Santos-Garcia D, Silva FJ, Morin S, Dettner K, Kuechler SM. The all-rounder *Sodalis*: a new bacteriome-associated endosymbiont of the Lygaeoid bug *Henestaris halophilus* (Heteroptera: Henestarinae) and a critical examination of its evolution. *Genome Biol Evol*. 2017;9:2893-2910.
41. Anbutsu H, Moriyama M, Nikoh N, Hosokawa T, Futahashi R, Tanahashi M, et al. Small genome symbiont underlies cuticle hardness in beetles. *Proc Natl Acad Sci U S A*. 2017;114:E8382.
42. Freilich S, Kreimer A, Borenstein E, Yosef N, Sharan R, Gophna U, et al. Metabolic-network-driven analysis of bacterial ecological strategies. *Genome Biol*. 2009;10:R61.
43. Karley AJ, Douglas AE, Parker WE. Amino acid composition and nutritional quality of potato leaf phloem sap for aphids. *J Exp Biol*. 2002;205:3009-3018.
44. Jiang Z-F, Xia F, Johnson KW, Brown CD, Bartom E, Tuteja JH, et al. Comparison of the genome sequences of “*Candidatus Portiera aleyrodidarum*” primary endosymbionts of the whitefly *Bemisia tabaci* B and Q biotypes. *Appl Environ Microbiol*. 2013;79:1757-1759.
45. Santos-Garcia D, Vargas-Chavez C, Moya A, Latorre A, Silva FJ. Genome evolution in the primary endosymbiont of whiteflies sheds light on their divergence. *Genome Biol Evol*. 2015;7:873-888.
46. Kapantaidaki DE, Ovčarenko I, Fytrou N, Knott KE, Bourtzis K, Tsagkarakou A. Low levels of mitochondrial DNA and symbiont diversity in the worldwide agricultural pest, the greenhouse whitefly *Trialeurodes vaporariorum* (Hemiptera: Aleyrodidae). *J Hered*. 2014;106:80-92.
47. Koga R, Moran NA. Swapping symbionts in spittlebugs: evolutionary replacement of a reduced genome symbiont. *ISME J*. 2014;8:1237-1246.
48. Lambert JD, Moran NA. Deleterious mutations destabilize ribosomal RNA in endosymbiotic bacteria. *Proc Natl Acad Sci U S A*. 1998;95:4458.
49. Baidya AK, Bhattacharya S, Dubey GP, Mamou G, Ben-Yehuda S. Bacterial nanotubes: a conduit for intercellular molecular trade. *Curr Opin Microbiol*. 2018;42:1-6.
50. Whitehead TR, Park M, Rabinowitz JC. Distribution of 10-formyltetrahydrofolate synthetase in eubacteria. *J Bacteriol*. 1988;170:995-997.
51. Kolde R. Package 'pheatmap'. 2011.
52. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*. 2006;22:1540-1542.
53. Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Pillay M, et al. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res*. 2013;42:D560-D567.
54. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res*. 2014;42:D199-D205.
55. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol*. 2016;428:726-731.
56. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*. 2014;15:R46.

57. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*. 2017;27:722-736.
58. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455-477.
59. Boetzer M, Pirovano W. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics*. 2014;15:211.
60. Boetzer M, Pirovano W. Toward almost closed genomes with GapFiller. *Genome Biol*. 2012;13:R56.
61. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26:2460-2461.
62. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357-359.
63. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*. 2014;9:e112963.
64. Caspi-Fluger A, Inbar M, Mozes-Daube N, Mouton L, Hunter MS, Zchori-Fein E. *Rickettsia* 'in' and 'out': two different localization patterns of a bacterial symbiont in the same insect species. *PLoS ONE*. 2011;6:e21096.
65. Chiel E, Inbar M, Mozes-Daube N, White JA, Hunter MS, Zchori-Fein E. Assessments of fitness effects by the facultative symbiont *Rickettsia* in the sweetpotato whitefly (Hemiptera: Aleyrodidae). *Ann Entomol Soc Am*. 2009;102:413-418.
66. Thompson WMO. Influence of amino acids on cassava biotype *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae) when feeding on an artificial system. *J Entomol*. 2006;3:198-203.
67. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods*. 2001;25:402-408.
68. Riens B, Lohaus G, Heineke D, Heldt HW. Amino acid and sucrose content determined in the cytosolic, chloroplastic, and vacuolar compartments and in the phloem sap of spinach leaves. *Plant Physiol*. 1991;97: 227-233.
69. Byrne DN, Miller WB. Carbohydrate and amino acid composition of phloem sap and honeydew produced by *Bemisia tabaci*. *J Insect Physiol*. 1990;36:433-439
70. Isaacs R, Byrne DN, Hendrix DI. Feeding rates and carbohydrate metabolism by *Bemisia tabaci* (Homoptera: Aleyrodidae) on different quality phloem saps. *Physiol Entomol*. 1998;23:241-248.
71. Crafts-Brandner SJ. Plant nitrogen status rapidly alters amino acid metabolism and excretion in *Bemisia tabaci*. *J Insect Physiol*. 2002;48:33-41.
72. Ohshima T, Hayashi H, Chino M. Collection and chemical composition of pure phloem sap from *Zea mays* L. *Plant Cell Physiol*. 1990;31:735-737.
73. Hayashi H, Chino M. Collection of pure phloem sap from wheat and its chemical composition. *Plant Cell Physiol*. 1986;27:1387-1393.
74. Hayashi H, Chino M. Chemical composition of phloem sap from the uppermost internode of the rice plant. *Plant Cell Physiol*. 1990;31:247-251.
75. Winter H, Lohaus G, Heldt HW. Phloem transport of amino acids in relation to their cytosolic levels in barley leaves. *Plant Physiol*. 1992;99:996-1004.
76. Hijaz F, Killiny N. Collection and chemical composition of phloem sap from *Citrus sinensis* L. osbeck (sweet orange). *PLoS ONE*. 2014;9:e101830.
77. Killiny N. Metabolomic comparative analysis of the phloem sap of curry leaf tree (*Berbera koenegii*), orange jasmine (*Murraya paniculata*), and Valencia sweet orange (*Citrus sinensis*) supports their differential responses to Huanglongbing. *Plant Signal Behav*. 2016;11:e1249080.
78. Valle EM, Boggio SB, Heldt HW. Free amino acid composition of phloem sap and growing fruit of *Lycopersicon esculentum*. *Plant Cell Physiol*. 1998;39:458-461.
79. Hocking PJ. The composition of phloem exudate and xylem sap from tree tobacco (*Nicotiana glauca* Grah.). *Annals of Botany*. 1980; 45:633-643
80. Sinisterra XH, McKenzie CL, Hunter WB, Powell CA, Shatters RG. Differential transcriptional activity of plant-pathogenic begomoviruses in their whitefly vector (*Bemisia tabaci*, Gennadius: Hemiptera Aleyrodidae). *J Gen Virol*. 2005;86:1525-1532.

81. Moran NA, Degan PH, Santos SR, Dunbar HE, Ochman H. The players in a mutualistic symbiosis: Insects, bacteria, viruses, and virulence genes. PNAS USA. 2005;102:16919.

Figure Legend

Fig. 1. Heatmap of the amino acid composition (percentage) of different phloem contents from different plants. X-axis - short names of plants (see Table 1 for full names); Y-axis - amino acids. Major clusters are indicated with Roman numerals (I to V). Red bars above indicate branches with confidence level $\geq 95\%$ (Additional File 1: Figure S1).

Fig. 2. Profiles for predicted origin of amino acids in tritrophic plant-insect-symbiont interaction systems, considering a range of host plants. Each column represents the phloem of a different plant. Within each phloem (column) simulations predicted the source of each amino acid in two holobionts formed by the sap-feeding insects - *B. tabaci* (left) and *A. pisum* (right) - together with *B. tabaci*'s primary symbiont *Portiera*. Green cells – amino acids available in the phloem; black cells – amino acids synthesized by *B. tabaci*; pink cells – amino acids synthesized by complementary interaction between *Portiera* and the corresponding insect host. X-axis - short names of plants (see Table 1 for full names); Y-axis - amino acids.

Fig. 3. Profiles for predicted origin of cellular building blocks in tritrophic plant-insect-symbiont interaction systems, considering several *B. tabaci*'s symbionts. Simulations predicted the source of key cellular building blocks in holobionts formed by *B. tabaci* with its primary symbiont *Portiera*, co-primary symbionts *Hamiltonella* and *Arsenophonus*, and secondary symbionts *Rickettsia*, *Wolbachia* and *Cardinium*. Green cells – amino acids available in the phloem (of cotton or broad bean); black cells – cellular building blocks synthesized only by *B. tabaci*; grey cells – cellular building blocks synthesized only by the symbiont; brown cells – cellular building blocks synthesized by both insect and symbiont separately; pink cells – cellular building blocks synthesized by complementary interaction between the insect and symbiont. X-axis: *B. tabaci*'s symbionts; Y-axis: key cellular building blocks. UDAM - Undecaprenyl-diphospho-N-acetylmuramoyl-(N-acetylglucosamine)-L-alanyl-gamma-D-glutamyl-L-lysyl-D-alanyl-D-alanine.

Fig. 4. The relative abundances of *Portiera* (A), *Hamiltonella* (B) and *Rickettsia* (C) in *B. tabaci* host grown on natural and artificial diets. Whiteflies reared on cotton plants were either transferred to another host plant – broad bean for long or short growth term experiments (Exp-1 and Exp-2, respectively) or to an artificial minimal medium supplemented with a single amino acid at a time. Exp-1 and artificial feeding with amino acids asparagine, serine, tryptophan and tyrosine were carried out in two independent experiments, results are consistent and are pooled together. Other artificial feeding experiments were carried once. Exp-2 and experiments with amino acids lysine, histidine and glutamine were carried out

following the loss of the *Rickettsia* population. Relative abundance of each symbiont before and after media change was determined with qPCR. Bars represent mean±standard error. Numbers indicate the number of whitefly individuals. Significance of change in relative abundance was determined with Student's t-test at $p \leq 0.05$. * Significance of change $p \leq 0.05$; *** Significance of change $p \leq 0.001$.

Fig. 5. An illustration for a hypothetical scenario for the effect of phloem content on the dynamics of symbionts of sap-feeding insects. Symbionts, which benefit from a competitive advantage by producing a nutrient that is available in limited amount, may lose this advantage when the nutrient is available in excess leading to a reduction in population size. Since symbionts that are co-fixated in the bacteriocytes are intertwined in complex metabolic interactions (black arrows), their abundance is correlated.









