Abstracts: Hemiptera

Altincicek *et al.*, 2008, Wounding-mediated gene expression and accelerated viviparous reproduction of the pea aphid Acyrthosiphon pisum.

Piercing of the pea aphid *Acyrthosiphon pisum* with a bacteria-contaminated needle elicits lysozyme-like activity in the haemolymph but no detectable activities against live bacteria. No homologues of known antimicrobial peptides were found in cDNA library generated by using the suppression subtractive hybridization method or in over 90 000 public expressed sequence tag (EST) sequences, but lysozyme genes have recently been described in pea aphid. Production of viviparous offspring was significantly accelerated upon wounding.

Pea aphid showed weakened immune system. No homologues of known antimicrobial peptides were found. The presence of insect defensins in other Hemiptera and in the basal apterygote insect *Thermobia domestica* (Altincicek Vilcinskas, 2007) suggests that at least this type of antimicrobial peptides may have been lost during aphid evolution. Interestingly, the observation that pierced aphids showed a limited capacity to seal their wound by haemolymph coagulation and melanization agrees with the finding that an encapsulation response of pea aphid to the parasitoid wasp *Aphidius ervi* is either very weak or non-existent (Oliver et al., 2005).

Regards to weakened immunity of pea aphids: (1) Aphids and relatives of Hemiptera share the unique ability to exploit exclusively phloem sap as diet, which is usually sterile (Douglas, 2006). Thus, the risk of encountering pathogens in their diet is limited. (2) Aphids harbour primary symbionts that are vertically transmitted and located intracellularly, as well as secondary symbionts that are both vertically and horizontally transmitted and also survive extracellularly in the insect haemolymph where they face the host's antimicrobial defences (Moran Dunbar, 2006; Haine, 2008). It is possible that the symbionts provide protection, *e.g.* pea aphid has been reported to be protected against fungal pathogens by the facultative symbiotic Gram-negative bacterium *Regiella insecticola* (Scarborough et al., 2005) and also against the parasitoid wasp *Aphidius ervi* by the facultative symbiotic Gram-negative bacterium *Hamiltonella defensa* (Oliver et al., 2005). This may further explain why only lysozyme-like activity is present in the haemolymph, as lysozymes target mainly Gram-positive bacteria, whereas aphid symbionts belong to Gram-negative bacteria. (3) As immune responses are costly because they require investment of re-

sources which are shared with other fitness-relevant traits (Rolff Siva-Jothy, 2003; Schmidt-Hempel, 2005; Freitak et al., 2007), it is reasonable that aphids increase terminal reproductive investment in response to a putative survival threat such as an immune challenge.

Gerardo *et al.*, 2010, Immunity and other defenses in pea aphids, Acyrthosiphon pisum. Pea aphids appear to be missing genes present in insect genomes characterized to date and thought critical for recognition, signaling and killing of microbes. In line with results of gene annotation, experimental analyses designed to characterize immune response through the isolation of RNA transcripts and proteins from immune-challenged pea aphids uncovered few immune-related products. Gene expression studies, however, indicated some expression of immune and stress-related genes.

In the fruit fly *Drosophila melanogaster*, recognition of an invasive microbe leads to signal production via four pathways (Toll, IMD, JNK, and JAK/STAT) (Boutros *et al.*, 2002). Each pathway is activated in response to particular pathogens (Dionne *et al.*, 2008). Signaling triggers the production of multitude effectors, including, most notably, antimicrobial peptides (AMPs). In insect genomes annotated to date, these pathways appear well conserved, with most of the key components found across flies (*Drosophila spp.*) (Sackton *et al.*, 2007), mosquitoes (*Aedes aegypti*, *Anopheles gambiae*) (Waterhouse *et al.*, 2007; Christophides *et al.*, 2002), bees (*Apis mellifera*) (Evans *et al.*, 2006) and beetles (*Tribolium castaneum*) (Zou *et al.*, 2007).

The cellular component of pea aphids' innate immune response may also be different to that seen in other insects. While many insects encapsulate parasitoid wasp larvae, smothering them to death with hemocytes, aphids appear not to have this layer of protection (Bensadia *et al.*, 2006; Carver *et al.*, 1988). Aphids, however, appear to recruit some hemocytes to parasitoid eggs, suggesting that cellular immunity may play an alternative, though possibly more limited, role (Bensadia *et al.*, 2006).

There is evidence that pea aphid has some defense systems common to other arthropods, *e.g.*, the Toll and JAK/STAT signaling pathways, HSPs, ProPO. However, several of the genes thought central to arthropod innate immunity are missing in pea aphid, including PGRPs, the IMD signaling pathway, defensins, c-type lysozymes.

The failure of finding aphid homologs to many insect immune genes can be resulted from large evolutionary distance between pea aphid and taxa used as reference (divided 100 million years ago). However, similar homology-search based method successfully detected immune-related genes in even more divergent insects. Another explanation for lack of immune genes is that pea aphid mount an alternative but equal immunity. However, functional analysis, together with Altincicek *et al.*, 2008, found little evidence for an alternative response to *E. coli* infection.

Altincicek *et al.*, 2008 proposed three hypotheses on the ecological success of pea aphid with the

possibility of lacking a strong immunity. First, aphids feeds on plant sap which is often sterile, leading to reduced risk for encountering pathogens. However, aphids are capable of acquiring pathogenic bacteria from the surface of their host plants' leaves (Stavrinides *et al.*, 2009), and aphids become host to a diverse assemblage of bacteria and fungi under stressful conditions (Nakabachi *et al.*, 2003). Furthermore, *Sitophilus* weevils, which when challenged with *E. coli* significantly up-regulate immune genes (Anselme *et al.*, 2008), spend their entire larval and nymph stages within sterile cereal grains, indicating that a sterile diet is not likely to explain the absence of antibacterial defenses in aphids. Second, aphid symbionts may provide protection against pathogens, *e.g.* pea aphid has been reported to be protected against fungal pathogens by the facultative symbiotic Gram-negative bacterium *Regiella insecticola* (Scarborough et al., 2005) and also against the parasitoid wasp *Aphidius ervi* by the facultative symbiotic Gram-negative bacterium *Hamiltonella defensa* (Oliver et al., 2005). This seems plausible regards to the cost of immune gene expression versus the benefit of protection by the secondary endosymbionts. However, it does not explain how the secondary endosymbionts (as Gram-negative bacteria), often present in aphid hemolymph, are themselves perceived and controlled by the aphid immune system. Third, aphids may invest in terminal reproduction in response to an immune challenge, rather than in a costly immune response, as Altincicek *et al.*, 2008 found increased viviparous offspring production upon wounding. Such an increase has been found in many invertebrates including *Biomphalaria* snails (Minchella *et al.*, 1981; Minchella *et al.*, 1985), *Acheta* crickets (Adamo *et al.*, 1999), *Daphnia* waterfleas (Chadwick *et al.*, 2005), and *Drosophila* flies (Polak *et al.*, 1998). Even without immune challenge, these insects also tends to invest most resources towards rapid, early onset reproduction (r-selection), and such organisms may specifically invest less in costly immune responses (Zuk *et al.*, 2002; Miller *et al.*, 2007). However, this may not sufficient for explaining weak immunity of pea aphids, as r-selected taxa such as *Drosophila* still mount complex immune responses. Furthermore, aphids do not increase their reproductive effort in the face of all immune challenges: fungal infection reduces the number of offspring pea aphid produce within 24 hours of inoculation (Baverstock *et al.*, 2006), and response to stabbing with bacteria seems to be specific to the aphid genotype and to the location of the stab.

Kim and Lee, 2017, Insect symbiosis and immunity: the bean bug-Burkholderia interaction as a case study.

Primary symbionts refers to maternally transmitted obligated symbionts. They associate with host at evolutionary time scale and become vital to host survival. The adaptive evolution of primary symbionts is accompanied with dramatic reduction in genome size and loss of genes essential for free-living (Moran, 2003; McCutcheon and Moran, 2012). Secondary or facultative symbionts are vertically/horizontally transmitted, and are not essential for host survival. Their association with hosts is recent, and still

possess free-living ability without collapsed genomes.

Bean bug *Riptortus pedestris* is a member of order Hemiptera feeding on plant sap. The midgut of bean bug is divided into morphologically distinct regions called M1, M2, M3, M4B and M4, where M4 is symbiotic. Symbiotic organ in M4 has two rows of crypts, whose lumens are densely colonized by betaproteobacterial symbionts of genus *Burkholderia*. The symbionts of bean bug are acquired orally from rhizosphere environment during early nymphal stage. *Burkholderia* is a soil bacterium. They retain free-living ability after association with bean bug and are easily cultured in labs.

nder bacterial challenges, symbiotic bean bugs exhibits better survival than aposymbiotic (Kim *et al.*, 2015). This better survival retains after inhibition of cellular immunity, indicating stronger humoral immunity induced by *Burkholderia*. Without bacterial challenges, antimicrobial peptide (AMP) (riptocin, rip-defensin and rip-thanatin) expression in fat body of symbiotic and aposymbiotic bean bugs is similar. However, AMP expression significantly increases in symbiotic bean bugs compared with aposymbiotic ones (Kim *et al.*, 2015).

Burkholderia is Gram-negative bacterium. Its cell envelop consists of inner- and outer membranes. Lipopolysaccharide (LPS) is located at the outer part of outer membrane. It is composed of lipid A embedded in outer membrane and oligosaccharide connecting with O-antigen. In symbiotic *Burkholderia*, O-antigen is lost when compared with free-living *Burkholderia* (Kim *et al.*, 2005). However, lipid A and oligosaccharide are retained. Besides, symbiotic *Burkholderia* are more susceptible to detergent than free-living ones, indicating compromised cell membrane integrity.

Free-living *Burkholderia* are resistant to antimicrobial activity of bean bug haemolymph (Loutet and Valvano, 2011), while symbiotic ones are highly susceptible (Kim *et al.*, 2015). When bean bug AMPs are purified, symbiotic *Burkholderia* are more susceptible to riptoctin and rip-defensin than free-living ones (Kim *et al.*, 2015). Ultimately, injected symbiont *Burkholderia* are removed much faster by bean bug immunity than free-living ones (Kim *et al.*, 2015).

Systemic injection of symbiotic *Burkholderia*, free-living *Burkholderia* and *Escherichia coli* triggers similar level of AMP expression in fat body. However, in midgut M4 region, the expression of AMPs is similar in symbiotic and aposymbiotic bean bugs. The AMP expression in M4 region is lower than basal expression of AMPs in fat body (Kim *et al.*, 2015). These indicate potential immune privilege of symbiotic organ.

The susceptibility of *Burkholderia* to bean bug immunity could be an advantage for easy management of symbiotic population. During nymphal stage, the size of symbiont population increases. However, pattern of transient decrease of symbiotic population is observed prior to moulting period in each instar stage. This transient decrease is corresponding to increase of antimicrobial activity of symbiont organ,

including up-regulated expression of c-type lysosome and riptocin (Kim *et al.*, 2014). Another mechanism for symbiont population management in bean bugs is related to M4B region of midgut. M4B region of symbiotic bean bug exhibits strong antimicrobial activity, while in aposymbiotic bean bug, little antimicrobial activity is exhibited. Besides, died bacterial symbionts present in M4B region. These indicate that the antimicrobial activity of M4B midgut is induced by *Burkholderia* (Kim *et al.*, 2013). One of components responsible for antimicrobial activity of M4B region is cathepsin-L-like protease (Byeon *et al.*, 2015), which is highly and preferentially expressed in M4B region of symbiotic bean bugs (Futahashi *et al.*, 2013). Antimicrobial activity exhibited by M4B and M4 region of bean bugs is only effective for symbiont *Burkholderia*, but free-living *Burkholderia* are resistant (Byeon *et al.*, 2015; Kim *et al.*, 2013).

Husink and McCutcheon, 2016, Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis.

Citrus mealybug *Planococcus citri* has two bacterial endosymbionts with an unusual nested arrangement: the γ-proteobacterium *Moranella endobia* lives in the cytoplasm of the β-proteobacterium *Tremblaya princeps*. To test the stability of this three-way symbiosis, host and symbiont genomes for five diverse mealybug species were sequenced. β-proteobacterial genomes from diverse mealybug species are *Tremblaya* with similar genome sizes, while γ-proteobacteria are from different clades with different genome sizes. Therefore, it is inferred that *Tremblaya* is the result of a single infection in the ancestor of mealybugs, while the γ-proteobacterial symbionts result from multiple replacements of inferred different ages from related but distinct bacterial lineages.

Three scenario of the order and timing of the γ-proteobacterial infections are proposed. In idosyncratic scenario, there was a single γ-proteobacterial acquisition in the ancestor of the Pseudococcinae that has evolved idiosyncratically as mealybugs diversified over time, leading to seemingly unrelated genome structures and coding capacities. In independent scenario, the γ-proteobacterial infections occurred independently, each establishing symbioses inside *Tremblaya* in completely unrelated and separate events. In replacement scenario, there was a single γ-proteobacterial acquisition in the Pseudococcinae ancestor that has been replaced in some mealybug lineages over time.

The idosyncratic scenario can be discarded as phylogenetics of γ-proteobacterial symbionts reveals that they have originated from clearly distinct and well-supported bacterial lineages. The independent and replacement scenarios are more difficult to tell apart. Under the independent scenario, *Tremblaya* may experience two rounds of genome corruption: one after association with mealybug ancestor, and one after infection of γ-proteobacteria. Therefore, one should expect diverse genome sizes in β- and γ-proteobacteria. Conserved genome sizes of β-proteobacteria and diverse sizes of γ-proteobacterial genomes favor the replacement scenario.

Two reasons why γ-proteobacteria end with living inside β-proteobacteria are proposed. The first is that it was easier to use the established transport system between the insect cell and *Tremblaya* than to evolve a new one. The second is that the insect immune system likely does not target *Tremblaya* cells, and so the *Tremblaya* cytoplasm is an ideal hiding place for a newly arrived symbiont.

Gil *et al.*, 2017, Tremblaya phenacola PPER: an evolutionary beta-gamma-proteobacterium collage.

Bougainvillea mealybug *Phenacoccus peruvianus* (PPER) harbors single betaproteobacterial symbiont *Tremblaya phenacola*. The genome of *Tremblaya phenacola* PPER is highly rearranged, in contrast to the high genomic stability of all previously sequenced *Tremblaya* lineages, with an almost absolute synteny conservation among *Tremblaya phenacola* strains (McCutcheon and von Dohlen, 2011; Husnik and McCutcheon, 2016), and a single inversion in *Tremblaya phenacola* in *Phenacoccus avenae* (PAVE) (Husnik and McCutcheon, 2016). Chromosome rearrangements cause perturbations in GC skew, which have a deleterious impact upon the replication system (Rocha, 2004). Therefore, although bacterial chromosomes can undergo many rearrangements at the beginning of an endosymbiotic relationship (see the marked example of '*Candidatus* Sodalis pierantonius' SOPE; Oakeson et al., 2014), long-term endosymbionts tend to present a typical GC skew, an indication that it is recovered with evolutionary time. Contrary to the PAVE genome, with a typical GC-skew pattern (Rocha, 2008), PPER genome presents a non-polarized and highly disrupted GC skew, except in the most syntenic region between both genomes, containing most ribosomal protein genes, suggesting that the chimeric genomic architecture is not stabilized.

The *Tremblaya phenacola* PPER genome contains 192 different CDSs, 188 with an assigned function. There are only four duplicated genes inside repeats (*rpsU*, *hisG*, *prmC* and $\mathrm{TPPER}_0$0169/220), *andtwohavetwohomologs*(*infAandrlmE*).*Itpossessesasingleribosomaloperonandacompletesetof*

In annotated CDSs, 102 CDSs appear to be of betaproteobacterial origin, but another 80 appear to belong to a gammaproteobacterium. Furthermore, there is a relationship between the taxonomic affiliation of each identified CDS and their G+C content. Generally, genes with gammaproteobacterial assignation have lower G+C values than betaproteobacterial assignation (Agashe and Shankar, 2014), which is consistent with genes in *Tremblaya phenacola* PPER. *Tremblaya phenacola* PPER genes not assigned to any category have a wide range in G+C content, and most of them have very short length. There is also differences in codon usage depending on the beta or gammaproteobacterial assignation in genes of *Tremblaya phenacola* PPER genes. The distribution of gamma or beta genes along the *Tremblaya phenacola* PPER genome is not random: most contigs contain only genes of one taxonomic origin, some others change the gene affiliation in the middle, and only one contig is completely intermixed.

The functional distribution of *Tremblaya phenacola* PPER genes is not random either. The transcrip-

tional machinery and the ribosomes are of betaproteobacterial origin, while aminoacyl-tRNA synthetases (not the complete set, as in other mealybugs) appear to be of gammaproteobacterial origin. The only exception is serS (a pseudogene in several *Tremblaya princeps* strains; Husnik and McCutcheon, 2016), which gave no clear affiliation. Except for iscSUA (involved in (Fe–S) cluster assembly), genes devoted to tRNA maturation are also of gammaproteobacterial origin. This pattern is similar to the nested endosymbiotic consortia from pseudococcinae mealybugs, *Tremblaya* has retained most of its own transcriptional and translational machinery except for aminoacyl-tRNA synthetases, which must be provided by the gammaendosymbiont. Furthermore, all maintained subunits of the DNA polymerase (also preserved in other *Tremblaya princeps*) are of beta origin. However, the other proteins involved in DNA replication (helicase and ligase) are of gamma origin; the first one has been preserved in all other *Tremblaya* genomes sequenced, while the second is absent in all of them. Genes involved in translation initiation (*infA*, *infB* and *infC*) and elongation (*fusA* and *tufA*) are of beta origin, although there is an additional gammaproteobacterial *infA*. Genes involved in translation termination (*prfA*, *prfB* and *prmC*), ribosome recycling (*frr*) and degradation of proteins stalled during translation (*smpB*), as well as N-formyltransferase (*fmt*) and peptide deformylase (*def*) are of gamma origin.

Like all other mealybug endosymbionts, *Tremblaya phenacola* PPER mediates essential amino acid synthesis. As in most studied pseudococcinae mealybugs, all genes retained for the biosynthesis of methionine, threonine, isoleucine, leucine and valine, and the production of phenylalanine from chorismate are of betaproteobacterial origin, while the pathways for the production of chorismate and lysine retain the same patchwork pattern. Histidine biosynthesis is an exception, as PPER has only retained genes of gammaproteobacterial origin. The cysteine biosynthetic pathway is more complete in PPER, with all genes of gamma origin. Regarding tryptophan biosynthesis, dominated by gammaproteobacterial genes in previously analyzed mealybugs' endosymbiotic consortia, in *Tremblaya phenacola* PPER the first step is performed by beta proteins, while the rest of the genes are of gamma origin, a similar pattern to that found in other insect's endosymbiotic consortia (that is, *Serratia/Buchnera* in lachninae aphids and some *Carsonella*/secondary systems in psyllids; Lamelas *et al.*, 2011; Sloan and Moran, 2012; Manzano-Marín *et al.*, 2016).

Tremblaya phenacola PPER genome goes beyond what could be considered a standard horizontal gene transfer event, and rather resembles the complete fusion of two genomes to form a new chimeric organism. Independent phylogenomic analyses of two concatenations of the *Tremblaya phenacola* PPER genes assigned as beta or gammaproteobacterial placed beta origin genes in *Tremblaya phenacola* clade, while the gammaproteobacterial genes were placed into the *Sodalis*-allied clade (Husnik and McCutcheon, 2016) as a sister species of '*Candidatus* Mikella endobia', nested gamma-endosymbiont of the pseudoccinae

mealybug *Paracoccus marginatus*.

How could the genomic fusion have ocurred? Although HGT is uncommon in modern endosymbionts, it is an extended phenomenon in flowering-plant mitochondria (Sanchez-Puerta, 2014), derived from an ancestral -proteobacterial endosymbiont (Andersson *et al.*, 1998). The most notable case corresponds to *Amborella trichopoda*, whose mitochondrial DNA has incorporated the complete mitochondrial genomes of three green algae and one moss, plus two mitochondrial genome equivalents from other angiosperms (Rice *et al.*, 2013). Such a high frequency of HGT has been explained by mitochondrial fusion and subsequent genomes fusion and rearrangements, mediated by homologous recombination systems (Maréchal and Brisson, 2010). Something similar might have occurred in *Tremblaya phenacola* PPER. On the basis of current evidences, the ancestor of all *Tremblaya* probably had a reduced genome (Husnik and McCutcheon, 2016). In the lineage driving to *Tremblaya phenacola* PPER, a gammaproteobacterium must have entered the consortium and, instead of replacing *Tremblaya phenacola* (as in the tribe Rhizoecini and genus *Rastrococcus*; Gruwell *et al.*, 2010), or establishing a nested endosymbiosis (as in the *Tremblaya phenacola* clade; reviewed by Husnik and McCutcheon, 2016), a cellular fusion event must have occurred, followed by genomic fusion. It cannot be discarded that a nested endosymbiosis preceded the cellular and genomic fusions. Because this phenomenon implies the existence of a DNA recombination machinery, the most plausible hypothesis is that such genes were present in the genome of the gammaproteobacterial donor, similarly to what has been described in citrus mealybug (López-Madrigal *et al.*, 2013). In fact, most mealybugs' gamma-endosymbionts that have been completely sequenced (McCutcheon and von Dohlen, 2011; López-Madrigal *et al.*, 2013; Husnik and McCutcheon, 2016) or screened for homologous recombination genes (López-Madrigal *et al.*, 2015) present a more or less complete recombination machinery. Transposable elements might also facilitate a fusion process. Some authors suggest that in arthropod intracellular environments, the possibility of two bacteria co-infecting the same cell generates an 'intracellular arena' where distantly related bacterial lineages can exchange mobile elements (Duron, 2013). However, although insertion sequences are frequent in early endosymbiotic stages (Latorre and Manzano-Marín, 2016), they have not been identified in any sequenced mealybugs' gamma-endosymbiont, and no indication of their former presence in *Tremblaya phenacola* PPER. After the fusion, the chimeric genome must have undergone massive gene loss, getting rid of almost all redundant and non-essential genes. The initial presence of homologs might have accelerated gene losses through recombination until DNA recombination genes disappeared. The remnant repeats involved in intrachromosomal recombination might have been maintained due to the loss of such genes, leading to the current, complex genome organization.

Szabo *et al.*, 2017, Convergent patterns in the evolution of mealybug symbioses involving

different intrabacterial symbiosis.

Manna mealybug *Trabutina mannipara* contains a betaproteobacterial symbiont *Tremblaya*, which contains a gammaproteobacterial symbiont *Trabutinella endobia*. Genomic sequences of *Tremblaya* are highly syntenic and harbored nearly identical sets of genes in the manna/citrus mealybug in accordance with a monophyletic origin of the outer symbionts among mealybugs. The genome of the intrabacterial symbiont *Trabutinella* shows only minimal synteny with that of *Moranella* from citrus mealybug, which is consistent with the distinct evolutionary origin of these symbionts. Although *Trabutinella* genome is much smaller than *Moranella* genome (McCutcheon von Dohlen, 2011), *Trabutinella* genome is likely still in the process of reduction as indicated by the presence of 27 pseudogenes (Moran and Bennett, 2014).

Genes of bacterial origin are found in manna mealybug genome. These genes are involved in synthesis of essential amino-acid, biotin and riboflavin. Many laterally acquired genes in genome of manna/citrus mealybug appear as sisters in phylogenetic tree, indicating that they share a common origin and were present in ancestral mealybugs before diversification of manna/citrus mealybug.

Symbiotic partners of the manna mealybug and citrus mealybug systems partition the synthesis of essential amino acids in a highly similar manner. In most of the essential amino acid production pathways, exactly the same steps are carried out by the inner or the outer symbiont in manna mealybug and citrus mealybug, despite the independent origin of the intrabacterial symbionts. A similar, yet far less complex situation has been observed among members of Auchenorrhyncha, where *Sulcia* synthesizes eight or seven essential amino acids while the remaining two or three are produced by different co-symbionts in different lineages, for instance. by *Baumannia* in sharpshooters, *Hodgkinia* in cicadas and *Zinderia* in spittlebugs (McCutcheon and Moran, 2010; Bennett and Moran, 2013).

A conceivable scenario explaining the observed similarities between the manna mealybug and citrus mealybug symbioses would be that before manna and citrus mealybug diverged, the *Tremblaya* ancestor was already infected by an (intra-)bacterial symbiont in ancestral mealybugs and this ancient association would have facilitated reduction of the *Tremblaya* genome and has shaped its gene repertoire. The inner symbiont might have been subsequently replaced in the ancestor of citrus and/or manna mealybug, with the new symbiont taking over the functions required by *Tremblaya* and at the same time allowing for loss of further genes from the *Tremblaya* genome, which would account for the observed differences between the two systems. This scenario is favoured over alternative scenarios such as ancient associations of *Tremblaya* with another bacteriocyte-associated symbiont because of the high level of congruence between the loss of genes in different *Tremblaya* strains at intermediate steps of essential amino-acid synthesis pathways.

Bublitz *et al.*, 2019, Peptidoglycan production by an insect-bacterial mosaic citrus mealybug, Tremblaya, Mornaella.

Citrus mealybug *Planococcus citri* has two bacterial symbionts: *Tremblaya princeps* that lives in bacteriocytes of mealybug, and *Moranella endobia* that lives in *Tremblaya*.

Genomics predicts a complete pathway for synthesis of peptidoglycan, composed of *Moranella* genes and several genes that were horizontally transferred from bacteria to the nuclear genome of mealybugs. No related genes are found in *Tremblaya*. Peptidoglycan constitutes are detected in whole insect preparation, and the peptidoglycan-specific molecule D-Ala ss specifically localized at the *Moranella* periphery. A peptidoglycan-targeting antibiotic specifically affects the *Moranella* cell envelope. A peptidoglycan-related horizontal gene transfer of Alphaproteobacterial origin is localized to the *Moranella* cytoplasm.

Peptidoglycan-based cell wall is an ancient and defining feature of bacteria, and peptidoglycan biosynthesis pathway is often highly conserved in bacterial genomes. There are two exceptions of this pattern. The first example of peptidoglycan-related horizontal gene transfers comes from the chromatophore of the rhizarian protist *Paulinella chromatophora*. *Paulinella* chromatophore has a peptidoglycan layer (Kies, 1974), which is encoded primarily on the chromatophore genome with the exception of one bacterial horizontal gene transfer to the host protist genome (Nowack *et al.*, 2016). The second example comes from the group of photosynthetic eukaryotes whose ancestor formed the original endosymbiosis with the cyanobacterium that became the chloroplast. This group, called the Archaeplastida, includes land plants, red algae, green algae, and glaucophyte algae (Lane and Archibald, 2008; McFadden, 2001). Many archaeplastidal nuclear genomes encode some peptidoglycan-related endosymbiosis gene transfers and horizontal gene transfers (van Baren *et al.*, 2016; Sato and Takano, 2017), but these genes do not always seem to work together to form a functional peptidoglycan layer at the chloroplast periphery. A chloroplast-localized peptidoglycan layer has been verified using fluorescently labeled D-Ala in the moss *Physcomitrella patens* (Hirano *et al.*, 2016), and possible chloroplast peptidoglycan layers have been observed by EM in glaucophytes (Schenk, 1970). But in the land plant *Arabidopsis thaliana*, which retains some peptidoglycan-related genes on its nuclear genome, no peptidoglycan layer exists at the chloroplast periphery and at least one peptidoglycan-related enzyme has been coopted for a different function (Garcia *et al.*, 2008). These results serve as a cautionary note about inferring function from genomics alone: gene presence is not a reliable predictor of biological function (Doolittle, 2013).

One important remaining question is the source of D-Ala and D-Glu in *Moranella*'s peptidoglycan, as homologs of Alr (alanine racemase) and MurI (glutamate racemase) do not present as horizontal gene transfers on mealybug genome or present in *Moranella* genome. These activities might be moonlighted by other genes. For example, GlyA and MetC have been shown to moonlight as alanine racemases in *Chlamydia trachomatis* and *Escherichia coli*, respectively (De Benedetti *et al.*, 2014; Kang *et al.*, 2011; Otten *et al.*, 2018), and eukaryotic homologs for these genes present on the mealybug genome. Similarly,

DapF has been shown to moonlight as a glutamate racemase in *Chlamydia trachomatis* (Liechti *et al.*, 2018), and this gene exists as an horizontal gene transfer of alphaproteobacterial origin on citrus mealybug genome (Husnik *et al.*, 2013). It is also possible that the source of D-Ala and D-Glu is not from these putatively moonlighting enzymes at all, but rather from either the plant sap diet of the insect or from D-amino acids in citrus mealybug produced from normal insect biochemistry. D-amino acids have been found in both plants (Robinson, 1976) and insects (Auclair and Patton, 1950; Corrigan and Srinivasan, 1966; Corrigan,1969), although the levels of these compounds have not been measured in citrus mealybug.

MurF encoded by horizontal transferred genes in mealybug genome is localized specifically in *Moranella* cytoplasm. Importing enzyme/mRNA into *Moranella* cytoplasm for peptidoglycan synthesis may help avoid immune responses. Other insects with long-term endosymbionts devote resources to scavenging peptidoglycan fragments in order to prevent continuous immune activation (Maire *et al.*, 2019). By sequestering peptidoglycan production to inside of *Moranella*, citrus mealybug may avoid the need for such contingency pathways, at least until *Moranella* cells are recycled near the end of the mealybug's life (Kono *et al.*, 2008).

Most peptidoglycan-related genes on mealybug genome (horizontal transferred) function in the cytoplasmic part of peptidoglycan synthesis, whereas the peptidoglycan-related genes retained by *Moranella* all code for inner membrane- or periplasm-associated proteins. It indicates that genes functioning in *Moranella* cytoplasm are more likely to be transferred to insect genome, and this may reflect the mechanisms through which proteins/RNA encoded by insect are transported into symbionts.

Host takeover of endosymbiont peptidoglycan production can be an important step in the regulation of endosymbiont cell division and potentially further integration with the host organism (de Vries and Gould, 2018). In moss, knocking out a peptidoglycan-related horizontal gene transfer on the nuclear genome results in enlarged chloroplasts (Machida *et al.*, 2006), and treatment with various peptidoglycan-targeting antibiotics results in fewer and larger chloroplasts per host cell (Katayama *et al.*, 2003). Together these data suggest that the movement of peptidoglycan-related genes from organelle genome to the host is a way for hosts to regulate organelle division (de Vries and Gould, 2018; Katayama *et al.*, 2003; Machida *et al.*, 2006). In citrus mealybugs, *Tremblaya* was acquired before *Moranella* (Hardy *et al.*, 2008; Thao *et al.*, 2002), and so the host insect must have found a way of controlling *Tremblaya* as the sole endosymbiont prior to the acquisition of *Moranella*. It is tempting to speculate that peptidoglycan-related horizontal gene transfers have been retained on the insect genome as a way of controlling the cell division of a bacterium that lives inside of another bacterium inside of insect cells.

Husink *et al.*, Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis.

The smallest reported bacterial genome belongs to *Tremblaya princeps*, a symbiont of *Planococcus citri* mealybugs (PCIT). *Tremblaya* PCIT not only has a 139 kb genome, but possesses its own bacterial endosymbiont, *Moranella endobia*. Genome and transcriptome sequencing, including genome sequencing from *Tremblaya* symbiont of *Phenacoccus avenae* (PAVE), which lacks intracellular bacteria, reveals that the extreme genomic degeneracy of *Tremblaya* PCIT likely resulted from acquiring *Moranella* as an endosymbiont. In addition, at least 22 expressed horizontally transferred genes from multiple diverse bacteria to the mealybug genome likely complement missing symbiont genes. However, none of these horizontally transferred genes are from *Tremblaya*, showing that genome reduction in this symbiont has not been enabled by gene transfer to the host nucleus.

Acquisition of *Moranella* symbiont may trigger extreme genome degeneracy in *Tremblaya* PCIT. Genome reduction in *Tremblaya* PAVE occurs to a degree consistent with other previously reported tiny symbiont genomes, and *Tremblaya* PCIT gene set is an almost perfect subset of *Tremblaya* PAVE. These results suggest that much of the reductive genome evolution observed in *Tremblaya* (down to approximately 170 kb) occurred before the acquisition of *Moranella* in the common ancestor of *Planococcus citri* and *Phenacoccus avenae* and that the extreme genomic degeneracy observed in *Tremblaya* PCIT (from 170 kb to 140 kb) was likely due to the acquisition of *Moranella* by *Tremblaya* at some point in the lineage leading to *Planococcus citri*. This scenario is consistent with studies showing that massive and rapid gene loss can occur in bacteria that transition to a symbiotic lifestyle (Mira *et al.*, 2001; Moran Mira, 2001; Nilsson *et al.*, 2005), after which gene loss slows, and gross genomic changes become infrequent, even over hundreds of millions of years (McCutcheon and Moran, 2010; Tamas *et al.*, 2002; van Ham *et al.*, 2003).

Pathways for translation, synthesis of essential amino acids, vitamins and peptidoglycan in *Tremblaya* PCIT are complemented by *Moranella*, mealybug genes originated from bacteria-to-mealybug horizontal gene transfers (HGTs) and mealybug genes of eukaryotic origin. In PCIT, ten HTGs group closely with other alphaproteobacterial sequences in phylogenetic trees, and nine HTGs from Gammaproteobacteria, two from Bacteroidetes, and one that is phylogenetically unresolved. The majority of these HGTs are not present in *Tremblaya* and *Moranella* genomes.

The presence of a large number of HTGs involved in peptidoglycan production and recycling is consistent with the hypothesis that cell lysis is the mechanism used to share gene products between *Moranella* and *Tremblaya* PCIT (Koga *et al.*, 2013; McCutcheon and von Dohlen, 2011). This idea was initially suggested based on a lack of transporters encoded on the *Moranella* genome combined with the large number of gene products or metabolites involved in essential amino acid biosynthesis and translation that would need to pass between *Moranella* and *Tremblaya* PCIT for the symbiosis to function (McCutcheon and

von Dohlen, 2011). Subsequent electron microscopy on mealybugs closely related to PCIT showed that although most gammaproteobacterial cells infecting the *Tremblaya* cytoplasm were rod shaped, some were amorphous blobs seemingly in a state of degeneration (Koga *et al.*, 2013). THe results suggest a plausible mechanism for how the insect host controls this process: by differentially controlling the expression of the horizontally transferred genes, the host could regulate the cell wall stability of *Moranella*. Increasing the expression of *murABCDE* genes would increase the integrity of *Moranella*'s cell wall, and increasing the expression of *mltD/amiD* would tend to decrease *Moranella*'s cell wall strength. As *Tremblaya* PCIT encodes no cell-envelope-related genes and likely uses host-derived membranes to define its cytoplasm, it would be unaffected by changes in gene expression related to peptidoglycan biosynthesis. This hypothesis is testable, because the levels of *Tremblaya* and *Moranella* are uncoupled in mealybugs closely related to PCIT; in males in particular, *Moranella* levels drop to undetectable levels while *Tremblaya* persists (Kono *et al.*, 2008). In situations where *Moranella* is reduced with respect to *Tremblaya*, low expression of *murABCDEF* and increased expression of *mltD/amiD* would be expected.

McCutcheon and Dohlen, 2011, An interdependent metabolic patchwork in the nested symbiosis of mealybug.

Citrus mealybug *Planococcus citri* represents a nested symbiosis system: a betaproteobacteria *Candidatus* Tremblaya princeps lives inside citrus mealybug, while a gammaproteobacteria *Candidatus* Moranella endobia lives in cytoplasm of *Tremblaya*.

Tremblaya genome is extremely small (0.14 Mbp) and degenerated (121 proteins). A 7-kbp region of *Tremblaya* genome exists two orientations within single insect host, while *Tremblaya* lacks genes involved in recombination.

Tremblaya retains genes involved in essential amino acid synthesis, but does not have complete pathways of its own. *Moranella* complements several essential amino acid synthesis genes lost in *Tremblaya*. However, it is unclear how transport of metabolites occurs between cosymbionts. *Tremblaya* genome encodes no predicted transporters. *Moranella* encodes a handful of proteins involved in membrane transportation, but none are specific for amino acids or their precursors. Some components of the Sec translocation machinery are present in the *Moranella* genome, and it is possible that these are used to transport some proteins across *Moranella*'s inner membrane. A search for signal peptides in the *Moranella* proteome revealed 27 proteins with N-terminal secretory signal peptides; however, none was involved in essential amino acid biosynthesis.

Tremblaya is missing several gene homologs for translation-related functions that are often retained in other highly reduced bacteria genomes, including all aminoacyl-tRNA synthetases, translational release factors. As translation machinery is significantly different in eukaryotes and bacteria, it seems unlikely

that the missing translation-related genes in *Tremblaya* are complemented by host. Horizontal gene transfer from bacteria to host might be the solution, although no transfer of functional genes between symbiont and host has been found in another two insects, pea aphid (Nikoh *et al.*, 2010) and human body louse (Kirkness *et al.*, 2010).

The nested structure of the mealybug symbionts is likely controlled by the host. There are at least two morphological forms of *Moranella*: a reproductive form in which cells were small in size and in the process of dividing, and a degenerative phase in which cells became unevenly shaped and elongated (Buchner, 1965). The particular Moranella form was dependent on the life stage of the insect and seemed to be synchronized within a bacteriocyte (Buchner, 1965). Furthermore, the infection levels of Tremblaya and Moranella are uncoupled in mealybugs (Kono *et al.*, 2008). During male development, the number of Moranella cells relative to *Tremblaya* cells drops significantly as the insects age, whereas in female insects, the levels of the two symbionts remain roughly equivalent over the entire life cycle (Kono *et al.*, 2008). Given that Tremblaya has an extremely limited coding capacity that is largely devoted to essential amino acid biosynthesis and translation, and given that only seven genes are of completely unknown function, it seems impossible that *Tremblaya* itself controls any structural aspect of the symbiosis. Likewise, the *Moranella* genome does not encode any genes involved in traditional infective strategies and does not indicate any obvious pathway by which it could be an active participant involved in seeking out the *Tremblaya* cytoplasm. Thus, it seems likely that the host is largely in control of the structure and organization of this bacteria-within-a-bacterium symbiosis.

Tremblaya survives with highly reduced genome with loss of genes thought to be essential for survival (*e.g.* translation). The missing activities can be complemented by several mechanisms: (1) gene products or metabolites of either host or bacterial origin imported from the host; (2) gene products or metabolites imported directly from the other symbionts if present; (3) genetic coadaptations to the loss of genes within the reduced genome itself; (4) the direct use of *Moranella* gene products as a result of a simple, passive mechanism such as *Moranella* cell lysis within the cell membrane system of *Tremblaya*.

Tremblaya genome is extremely small, but low gene dense. During the shift from a free-living to an obligate intracellular lifestyle, where the constant exposure to the stable and rich environment of the host cell combined with a severe reduction in population size (and subsequent reduction in the efficacy of purifying selection) allows large numbers of pseudogenes to accumulate (Ochman *et al.*, 2006; Andersson *et al.*, 2001). These pseudogenes are eventually purged from the genome through mutational patterns favoring deletions (Mira *et al.*, 2001), leading to small gene-dense genomes such as those from insect nutritional symbionts. A possible explanation is that *Tremblaya* undergone genome reduction after association with mealybug, and acquisition of *Moranella* leads to further genome reduction. Basal

lineages of mealybugs in the same subfamily as citrus mealybug seem to contain *Tremblaya* without the intracellular gammaproteobacterial endosymbiont (Hardy *et al.*, 2008; Thao *et al.*, 2002), indicating that *Moranella* was acquired after the establishment of *Tremblaya*. The patterns of gene pseudogenization also fit this hypothesis, as most pseudogenized *Tremblaya* genes have functional *Moranella* homologs.

Gomez-Polo *et al.*, 2017, An exceptional family: Ophiocordyceps-allied fungus dominates the microbiome of soft scale insects (Hemiptera Sternorrhyncha: Coccidae).

Ribosomal genes from seven soft scale (Coccidae) species showed high prevalence of an *Ophiocordyceps*-allied fungal symbiont, which is from an lineage widely known as entomopathogenic. The *Ophiocordyceps*-allied fungus from soft scales is closely related to fungi described from other hemipterans, and they appear to be monophyletic, although the phylogenies of the *Ophiocordyceps*-allied fungi and their hosts do not appear to be congruent. Microscopic observations show that the fungal cells are lemon-shaped, are distributed throughout the host's body and are present in the eggs, suggesting vertical transmission.

Deng *et al.*, 2021, The ubiquity and development-related abundance dynamics of Ophiocordyceps fungi in soft scale insects.

Nuclear ribosomal internal transcribed spacer (ITS) gene fragment was used to analyze the diversity of fungal communities in 28 soft scale (Coccidae) species. Coccidae-associated *Ophiocordyceps* fungi (COF) were prevalent in all 28 tested species with high relative abundance. Meanwhile, the first and second instars of *C. japonicus* had high relative abundance of COF, while the relative abundances in other stages were low, ranging from 0.68% to 2.07%. The result of fluorescent in situ hybridization showed that the COF were widely present in **hemolymph** and vertically transmitted from mother to offspring.

Chong Moran, 2016, Intraspecific genetic variation in hosts affects regulation of obligate heritable symbionts.

The extent of intraspecific variation in the regulation of a mutually obligate symbiosis, between the pea aphid *Acyrthosiphon pisum* and its maternally transmitted symbiont *Buchnera aphidicola*, using experimental crosses to identify effects of host genotypes. Symbiont titer, as the ratio of genomic copy numbers of symbiont and host, as well as developmental time and fecundity of hosts, were measured. There was a large (¿10-fold) range in symbiont titer among genetically distinct aphid lines harboring the same *Buchnera* haplotype. Aphid clones also vary in fitness, measured as developmental time and fecundity, and genetically based variation in titer is correlated with host fitness, with higher titers corresponding to lower reproductive rates of hosts. The results show that obligate symbiosis is not static but instead is subject to short-term evolutionary dynamics, potentially reflecting coevolutionary interactions between host and symbiont.

Henry *et al.*, 2013, Horizontally Transmitted Symbionts and Host Colonization of Eco-

logical Niches.

Facultative or "secondary" symbionts are common in eukaryotes, particularly insects. While not essential for host survival, they often provide significant fitness benefits. It has been hypothesized that secondary symbionts form a "horizontal gene pool" shuttling adaptive genes among host lineages in an analogous manner to plasmids and other mobile genetic elements in bacteria. However, we do not know whether the distributions of symbionts across host populations reflect random acquisitions followed by vertical inheritance or whether the associations have occurred repeatedly in a manner consistent with a dynamic horizontal gene pool. Here we explore these questions using the phylogenetic and ecological distributions of secondary symbionts carried by 1,104 pea aphids, *Acyrthosiphon pisum*. We find that not only is horizontal transfer common, but it is also associated with aphid lineages colonizing new ecological niches, including novel plant species and climatic regions. Moreover, aphids that share the same ecologies worldwide have independently acquired related symbiont genotypes, suggesting significant involvement of symbionts in their host's adaptation to different niches. We conclude that the secondary symbiont community forms a horizontal gene pool that influences the adaptation and distribution of their insect hosts. These findings highlight the importance of symbiotic microorganisms in the radiation of eukaryotes.

Hosokawa *et al.*, 2007, Obigate symbiont involved in pest status of host insect.

A pest stinkbug species, *Megacopta punctatissima*, performed well on crop legumes, while a closely related non-pest species, *Megacopta cribraria*, suffered low egg hatch rate on the plants. When their obligate gut symbiotic bacteria were experimentally exchanged between the species, their performance on the crop legumes was completely reversed: the pest species suffered low egg hatch rate, whereas the non-pest species restored normal egg hatch rate and showed good performance. The low egg hatch rates were attributed to nymphal mortality before or upon hatching, which were associated with the symbiont from the non-pest stinkbug irrespective of the host insect species.

Couret *et al.*, 2019, Even obligate symbioses show signs of ecological contingency: Impacts of symbiosis for an invasive stinkbug are mediated by host plant context.

Many species interactions are dependent on environmental context, yet the benefits of obligate, mutualistic microbial symbioses to their hosts are typically assumed to be universal across environments. We directly tested this assumption, focusing on the symbiosis between the sap-feeding insect *Megacopta cribraria* and its primary bacterial symbiont *Candidatus* Ishikawaella capsulata. We assessed host development time, survival, and body size in the presence and absence of the symbiont on two alternative host plants and in the insects' new invasive range. We found that association with the symbiont was critical for host survival to adulthood when reared on either host plant, with few individuals surviving in the absence of symbiosis. Developmental differences between hosts with and without microbial symbionts, however,

were mediated by the host plants on which the insects were reared. Our results support the hypothesis that benefits associated with this host–microbe interaction are environmentally contingent, though given that few individuals survive to adulthood without their symbionts, this may have minimal impact on ecological dynamics and current evolutionary trajectories of these partners.

Arp *et al.*, 2016, Annotation of the Asian citrus psyllid genome reveals a reduced innate immune system. A genome-wide analysis of immune genes in Asian citrus psyllids *Diaphorina citri* was presented. β-1,3-glucan binding protein and IMD signaling (Imd, Dredd and Relish) are absent. Antimicrobial peptides are absent.

Sloan *et al.*, 2014, Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genome in sap-feeding insects.
Endosymbionts typically have experienced extreme genome reduction, but the role of host in this process remains unclear. *Carsonella ruddii* is a vertically-transmitted gammaproteobacterial endosymbiont widely-present in psyllid bacteriomes. It lacks genes involved in DNA replication, transcription and translation. There are three possible mechanisms for the exceptional gene loss: (1) modification in cellular processes or selection for multifunctional proteins; (2) compensation from additional endosymbionts, as observed in psyllid *Ctenarytaina eucalypti*; (3) compensation from host-coding proteins. mRNA-seq data of psyllid *Pachypsylla venusta* revealed that host genes that were up-regulated in bacteriome complement amino acid synthesis pathways that are absent/incomplete in endosymbionts. Draft genome of host revealed horizontal gene transferrs (HGTs) from bacteria of diverse lineage.

Salcedo-Porras *et al.*, 2019, *Rhodnius prolixus*: identification of missing components of the IMD immune signaling pathway and functional characterization of its role in eliminating bacteria. Previously reported missing components of IMD pathway were found in *Rhodnius prolixus*. They were involved in response to infection with Gram-negative bacteria. RNAi revealed the role of IMD pathway in regulating antimicrobial peptides (AMPs).

Kwak *et al.*, 2022, Chromosomal-level assembly of *Bactericera cockerelli* reveals rampant gene family expansions impacting genome structure, function and insect-microbe-plant-interactions.

Owen *et al.*, 2020, Hemiptera phylogenomic resources: tree-based orthology prediction and conserved exon identification.

Ma *et al.*, 2020, JNK pathway plays a key role in the immune system of the pea aphid and is regulated by microRNA-184.

Tomizawa *et al.*, 2020, Numerous peptidoglycan recognition protein genes expressed in the bacteriome of the green rice leafhopper *Nephotettix cincticeps* (Hemiptera, Cocadellidae).

Ma *et al.*, 2021, Comparative analysis of *Adelphocoris suturalis* Jakovlev (Hemiptera: Mirodae) immune responses to fungal and bacterial pathogens.

Yu *et al.*, 2021, Characterization of PGRP-LB and immune deficiency in the white-backed planthopper *Sogatella furcifera* (Hemiptera: Delphacidae).

Nishide *et al.*, 2019, Functional crosstalk across IMD and Toll pathways: insight into the evolution of incomplete immune cascades.