Are ecological communities the seat of endosymbiont horizontal transfer and diversification? A case study with soil arthropod community.

Manisha Gupta[1], Rajbir Kaur[1], Ankita Gupta[2], and Rhitoban Raychoudhury[3]

[1]IISER Mohali
[2]NBAIR
[3]Indian Institute of Science Education and Research Mohali

November 21, 2020

Abstract

Maternally inherited endosymbionts are one of the most abundant bacteria infecting arthropods and show extensive horizontal transfer. Such widespread distribution and extensive recombination among these endosymbionts could be an outcome of horizontal transfer as for such genetic exchanges to occur their hosts should come in contact. One such level of biological organization where different hosts can do that is the ecological community. Despite various studies focusing on known model species and specific ecological interactions among hosts, reports on community wide endosymbiont data are rare. To better understand endosymbiont spread, we investigated the incidence, diversity, extent of horizontal transfer and recombination of three such endosymbionts (Wolbachia, Cardinium and Arsenophonus) in a specific soil arthropod community. Wolbachia strain characterization was done using multiple genes whereas single 16S rRNA gene was used for Cardinium and Arsenophonus. Amongst 3509 individual host arthropods belonging to 390 morphospecies, 12.05% were infected with Wolbachia, 2.82% with Cardinium and 2.05% with Arsenophonus. Phylogenetic incongruence between host and endosymbiont indicated extensive horizontal transfer of endosymbionts within this community. Three cases of recombination between Wolbachia supergroups and eight incidences of within supergroup genetic exchange were also found. Statistical tests of similarity indicated supergroup A Wolbachia and Cardinium to show a pattern consistent with rapid horizontal transfer within the community. However same tests done for super group B Wolbachia and Arsenophonus did not show similar patterns. We highlight the importance of extensive community wide studies for a better understanding of the spread of endosymbionts across global arthropod communities.

Introduction:

Maternally-inherited endosymbionts infecting arthropods are one of the most diverse and abundant of all bacteria infecting them. About two-third of terrestrial arthropods are infected with at least one maternally-inherited endosymbiont (Hilgenboecker, Hammerstein, Schlattmann, Telschow, & Werren, 2008; Weinert, Araujo-Jnr, Ahmed, & Welch, 2015). These endosymbionts play crucial role in the ecology and evolution of their hosts (Gebiola et al., 2017; Semiatizki et al., 2020). The most abundant of these are *Wolbachia, Cardinium, Arsenophonus, Rickettsia* and *Spiroplama.* Out of these, *Wolbachia* remains the most widely distributed endosymbiont with an incidence rates of 16-66% (Hilgenboecker et al., 2008; Werren, Windsor, & Guo, 1995) and 18 different clades (supergroup A to R) reported in different hosts across the world (Landmann, 2019). Incidence of the other endosymbionts varies from 4-10% (Duron et al., 2008; Zchori-Fein & Perlman, 2004).

The key factor explaining this abundance has been their ability to jump from one host to the other, in spite of being vertically transmitted from mother to offspring (Werren, Baldo, & Clark, 2008). As a result, these endosymbionts rarely show congruence with the host phylogeny, indicating substantial horizontal transfer

across evolutionary timescales to taxonomically unrelated hosts (Shoemaker et al., 2002; Werren, Zhang, & Guo, 1995). This is also evident from the occurrence of similar endosymbiont strains in taxonomically unrelated hosts and conversely, the presence of divergent strains in closely related hosts (Vavre, Fleury, Lepetit, Fouillet, & Bouletreau, 1999). Individual arthropods can harbor multiple strains of one endosymbiont as well as multiple strains of different endosymbionts (Zélé et al., 2018) which perhaps indicates that different endosymbionts can use the same host to spread across different arthropod communities.

Another key feature of endosymbionts is the pervasive recombination seen in their genomes (Ellegaard, Klasson, Naslund, Bourtzis, & Andersson, 2013). This has been particularly well documented in *Wolbachia*(Malloch & Fenton, 2005) as well as other endosymbionts (Mouton et al., 2012). The level of recombination is so extensive that single gene sequences are unable to properly reflect the evolutionary history of a strain. Unsurprisingly, this has necessitated the development of multi locus strain typing (MLST) system (Maiden et al., 1998). The results of such MLST surveys revealed the extent of recombination to be similar to those of pathogenic free-living bacteria (Yahara et al., 2016). This is surprising because, unlike free-living bacteria, most endosymbionts cannot survive outside the host. Therefore, this extensive recombination must be an outcome of the horizontal transfer of strains across arthropod communities, since, for recombination to happen, two endosymbionts must come in contact within one host cytoplasm. As new strains are horizontally transferred to novel hosts; they encounter resident endosymbionts and thereby increase opportunities of recombination between them. Evidence for such recombination is also well documented. The parasitoid wasp *Nasonia* and its host*Protocalliphora* show the presence of a very similar recombinant *Wolbachia* (Werren & Bartos, 2001). Recombinant *Wolbachia*strains have also been reported in *Anastrepha* fruit flies and their parasitoid braconid wasps (Mascarenhas, Prezotto, Perondini, Marino, & Selivon, 2016).

Horizontal transfer, therefore, can explain at least two major characteristics of endosymbionts, their wide distribution as well as the recombinant nature of their genomes. A major question that emerge from this is, at what level of biological organization are these horizontal transfers taking place? A relatively simple way to uncover this level is to enumerate specific ecological interactions where close associations between the two interacting arthropods have been implicated in horizontal transfer. These include host-parasite, host-parasitoid, prey-predator and other ecological relationships. Examples where host-parasitoid interactions have been implicated for such transfer include the presence of similar *Wolbachia* strains among frugivorous *Drosophila* and their hymenopteran parasitoid (Vavre et al., 1999), *Nasonia vitripennis* and *Muscidifuraxuniraptor* sharing similar *Wolbachia* with their fly host*Protocalliphora* (Baudry, Bartos, Emerson, Whitworth, & Werren, 2003), transmission of *Wolbachia* into whitefly via parasitoid wasps (Ahmed, Breinholt, & Kawahara, 2016). Another such ecological association which can lead to endosymbiont transfer is prey-predator relationships like the predatory mite *Metaseiulus occidentalis*and its prey *Tetranychus urticae* (spider mite) sharing similar endosymbionts (Hoy & Jeyaprakash, 2005). Parasites like mites can also facilitate the transfer of *Wolbachia* to *Drosophila* host populations (Brown & Lloyd, 2015). These can also be host plant mediated transfer of *Cardinium* to different leaf hopper species (Gonella et al., 2015) as well as horizontal transfer of*Wolbachia* in whitefly *via* cotton leaves (Li et al., 2017).

It is clear from these examples that these horizontal transfers are taking place when two hosts are coming together to perform a particular ecological function. The endosymbiont present within these hosts are then serendipitously getting transferred from one host to the other. Therefore, to understand the dynamics of the spread of endosymbionts through horizontal transfer, one needs to look at the level where most of these ecological associations are taking place, which is within a particular ecological community. A well-defined ecological community will have several diverse host taxa interacting with each other, thereby, facilitating horizontal transfer. Moreover, many host taxa can belong to many different ecological communities (Morrow, Frommer, Shearman, & Riegler, 2014). This cosmopolitan nature of a few host taxa will further facilitate the spread of endosymbionts from one ecological community to another, almost like spreading through a metacommunity (Brown, Mihaljevic, Des Marteaux, & Hrček, 2020). Therefore, investigating endosymbiont diversity and horizontal transfer within specific ecological communities seems logical. Yet, there are very few studies that have taken this approach and instead focus mainly on endosymbiont spread within a particular habitat (Stahlhut et al., 2010), or in a specific genus (Baldo et al., 2008; Raychoudhury, Baldo, Oliveira, &

Werren, 2009) or within specific taxa (Ahmed et al., 2016). Amongst community-wide surveys, Kittayapong, Jamnongluk, Thipaksorn, Milne, and Sindhusake (2003), demonstrated *Wolbachia* strain diversity within rice field arthropod community. Sintupachee, Milne, Poonchaisri, Baimai, and Kittayapong (2006), reported plant-mediated horizontal transfer among arthropod community found on pumpkin leaves. Most of these studies are based on single gene phylogenies which makes identification of recombination difficult. An important corollary of this view of within-community horizontal transfer of endosymbionts can lead to another important hypothesis about sequence diversity of the endosymbionts themselves. If endosymbionts are rapidly undergoing horizontal transfer within a particular ecological community then very similar bacterial strains would be found among the arthropod hosts of that community. This would make these bacteria more closely related to each other, than expected, resulting in lower than expected pairwise sequence divergence among them. This lower than expected levels of sequence divergence can serve as a signature of recent and relatively rapid community-wide horizontal transfer of resident endosymbionts.

In the present study, we try to answer whether such relatively rapid horizontal transfer and resulting recombination are happening within the endosymbionts of a diverse soil arthropod community. Three major endosymbionts, *Wolbachia, Cardinium* and *Arsenophonus* , were selected and screened across arthropod hosts. We investigated *Wolbachia* sequence diversity using the well-established MLST scheme (Baldo et al., 2006) and also identified specific recombination events. We also investigated *Cardinium* and *Arsenophonus*incidence but with single gene sequences. A statistical model was then used to test whether the endosymbiont found within this community are more closely related to each other than expected.

Materials and Methods:

Sample collection and morphological identification of soil arthropods:

To access soil arthropod biodiversity, sampling was done in October and November 2015 (a post-monsoon season) from a relatively undisturbed land ($220 \times 70 m^2$) near the vicinity of the host institution (30?39'N 76?43'E, Mohali, Punjab, India; Figure S1A). The study area was naturally divided into roughly 8 quadrants by plantations of poplar (Figure S1B). Five randomly selected quadrants were sampled by collecting leaf litter and pitfall traps. These two independent sampling methods were used in tandem, since pitfall traps is biased for surface-active taxa, whereas leaf litter method is biased towards less active taxa, resulting in a more comprehensive sampling of the resident species (Olson, 1991; Querner & Bruckner, 2010). Two parallel transects, each 30m long, were marked across each quadrant using a rope (Figure S1B). Each of these transects were marked at 10m and two alternative marked points were sampled for leaf litter, while pitfall traps were placed at the other two ends (Figure S1C). Samples from each type of collections were later combined. In total, 20 collections, each from leaf litter and pitfalls, were obtained.

Leaf litter was collected from an area of roughly $0.09 m^2$ (Figure S1D), and immediately placed in a plastic bag. Additionally, each leaf litter sample was accompanied by a soil sample of an area $282 cm^3$ immediately below the leaf litter (Sabu & Shiju, 2010). Samples were weighed so that each sample roughly had the same weight (500-600gm) and were then settled in a series of Tullgren funnel with 100W light source. The emerging arthropods were collected in a 50 ml beaker, with absolute ethanol, continuously for the next 4-6 days or until no arthropod samples emerged. Emerged arthropods were collected every 24 hours and 100 ml of fresh absolute ethanol was added to the collection beaker. The pitfall traps were settled on the ground by placing a 250ml beaker with 50ml of absolute ethanol (Figure S1E). Samples were collected every second day with a replacement of fresh ethanol.

Arthropods obtained from each of pitfall and leaf litter samples (20 samples each) were sorted individually. Detailed (dorsal and lateral) views of each individual arthropod obtained were photographed under a stereomicroscope (M205C, Leica Microsystems) with scale varying from 0.2-2mm. These were then sorted according to their morphology (into morphospecies) and provisionally identified till order level. A total of 3509 individuals were sampled, which were categorized into 390 distinct morphospecies. Several diversity indices were calculated with EstimateS v9.1.0 (Colwell, 2013) like ACE (Abundance coverage estimator; Chao, Hwang, Chen, & Kuo, 2000), Chao1 (Chao, 1984), ICE (Incidence coverage estimator), Chao2 and

Jack 1 and 2 (Smith & van Belle, 1984).

For morphospecies which had more than three individuals, DNA was extracted from a single individual by either the HiPurA^{TM} insect DNA purification kit (HIMEDIA) or by using the Phenol-Chloroform-Isoamyalcohol (PCI) method. In PCI method, samples were crushed in 200µl lysis buffer containing 10mM each of Tris-HCL (pH 8.0), EDTA (pH 8.0) and NaCl. DNA was precipitated using isopropanol and dissolved in 1X TE (pH 8.0). For morphospecies which only had single individuals, a different nondestructive extraction protocol was used (Rowley et al., 2007). Whole individuals were incubated at 60°C in 100-400µl of Guanidinium thiocyanate (GuSCN) based extraction buffer (GuSCN, 0.1M, Tris-HCL, 0.2M EDTA with Triton x-100) for 1-4hr. Then individuals were removed for storage and the DNA remaining in the buffer was precipitated using isopropanol. Extracted DNA was quantified using the NanoDrop^{TM} 2000 spectrophotometer (Thermo Fisher Scientific) and PCR-suitability was accessed by running a PCR reaction using $28S$ primers (Table S2).

The morphospecies were barcoded using (Hebert, Cywinska, Ball, & Dewaard, 2003) the ˜600bp of the mitochondrial *CO1* gene (Table S2). 2-20ng/µl of extracted DNA was used in 20µl PCR reactions with an initial denaturation step at 95°C for 3 minutes, 39 cycles of denaturation (95°C, 45 seconds), annealing (51-56°C for 45 seconds), extension (72°C, 1 minute) and a final extension at 72°C for 10 minutes. PCR products were visualized on 1% agarose gels and then cleaned with Exonuclease I and Shrimp alkaline Phosphatase (New England Biolabs Inc.). PCR products were sequenced using BigDye® Terminator v3.1 cycle sequencing kit. Initially, only the forward strand was sequenced and if any base ambiguity was observed then the reverse strand was also sequenced.

CO1 sequences obtained were identified through the NCBI (Johnson et al., 2008) and BOLD databases (Ratnasingham & Hebert, 2007) by BLAST (last performed in August 2019). The best hit obtained was used to check the provisionally identified morphospecies. If both databases yielded the same hit then it was determined to have been identified. If they yielded different hits then the taxonomic identification was moved down to the level common in these two hits. These results were further cross-referenced with the photographic data to finally build up a repository of identified morphospecies with their taxonomic ranks (Table S3). Only unique morphospecies were included in further analysis after analyzing the *CO1* sequences.

Endosymbiont screening and phylogenetic analysis:

All the 390 morphospecies were screened for the three endosymbionts- *Wolbachia* , *Cardinium* and *Arsenophonus* . Incidence of each of these endosymbionts was estimated using primers specific to them (Table S2). The Multi Locus Strain Typing (MLST) system was used (Baldo et al., 2006) to identify and characterize the *Wolbachia* infections. For *Cardinium* and *Arsenophonus,* $16S$ *rRNA* gene was amplified using specific primers (Table S2).

To test for the presence of *Wolbachia* , the *wspec* primers were used. Samples positive for *wspec* were then sequenced for one of the MLST genes, usually *fbpA* , to identify single *Wolbachia* infections by inspecting the chromatograms for multiple peaks. Samples with multiple *Wolbachia* infection were not processed further as assigning a particular sequence to a particular *Wolbachia* would have been impossible. Resultant allele sequences from MLST genes were compared with existing sequences in PubMLST database (Jolley, Bray, & Maiden, 2018) to identify their allele profiles (number assigned to each unique sequence) and ST (new strain type as defined by the combination of five MLST allele profiles). Sequences that did not have a match in the PubMLST database were submitted to the database for curation. Sequences obtained from this study were deposited in NCBI and BOLD database (Table S3).

Sequences were aligned with Sequencher 5.2.4 (Gene Codes Corporation) and manually edited with BioEdit v. 7.2.5 (Hall, 1999). DNA sequence evolution models were computed using MEGA7 (Kumar, Stecher, & Tamura, 2016). GTR+g (general time reversible model with γ-distributed rate variation) was found to be the best model for all *CO1* phylogenetic trees. Bayesian phylogeny was constructed for *CO1* sequences using MrBayes v3.2.5 (Ronquist et al., 2012). Each phylogenetic analysis was run at least twice and was accepted only if there was no change in the major branching order (Figure S2). Phylogenetic trees were visualized

and edited with Figtree v1.4.2 (Rambaut, 2009).

Maximum likelihood phylogenetic trees of *Wolbachia, Cardinium and Arsenophonus* were constructed in MEGA7 with 1000 bootstrap replicates for each. The suitable substitution models obtained were T92+g+i (Tamura 3-parameter with γ-distributed rate variation and proportion of invariable sites) for concatenated MLST dataset, T92+g for *gatB* ,*hcpA* , *ftsZ* , *fbpA* gene, HKY+g (Hasegawa, Kishino, and Yano) for *coxA* and K2+g (Kimura 2-parameter) for*Cardinium* and *Arsenophonus* .

To account for the frequent recombination seen in *Wolbachia*genomes, ClonalFrame v2.1 (Didelot & Falush, 2007) was used to infer phylogeny from multilocus sequence data. ClonalFrame was run for 3 x 10^5 iterations with the first 50% iterations discarded as burn-ins. Estimates of recombination rate was also obtained.

Identifying horizontal transfers, recombination events and test of endosymbiont similarity:

To test for horizontal transmission of endosymbionts across the soil arthropod hosts, two sets of analyses were done. The first was a visual estimation of the horizontal transfer obtained by comparing host and endosymbiont phylogenies. The second was a quantitative estimation of the correlation between the pairwise distance matrices of hosts and their corresponding endosymbionts. The Spearman method (r) of Mantel test for correlation (Legendre & Legendre, 2012) was determined in R v1.2.5. A total of 5 different correlations were computed with their corresponding host sequences: all the *Wolbachia* infected samples (33), only A supergroup *Wolbachia* infected samples (16), only B supergroup *Wolbachia* infected samples (15), only*Cardinium* infected samples (8) and *Arsenophonus* infected samples (7).

To identify and visualize recombination events within the*Wolbachia* concatenated MLST sequences (2079bp), SplitsTree v4.14.8 (Huson, Kloepper, & Bryant, 2008) was used to construct phylogenetic network using uncorrected p -distance and Neighbor-net method (Bryant & Moulton, 2004). To statistically evaluate presence of recombination, Φ test (Bruen, Philippe, & Bryant, 2006) was used in SplitsTree v4.14.8. These identified events were then evaluated using RDP4 v4.97 (Martin, Murrell, Golden, Khoosal, & Muhire, 2015) which has several suits of programs to detect and identify recombination events (like RDP, GENECONV, MaxChi, SiScan, BootScan, Chimaera and 3Seq). A recombination event was considered significant if it was shown to be statistically significant by at least three or more detection methods. Recombination breakpoints were also identified with RDP v4.97.

To test whether the endosymbionts are largely being transferred within the community, and thereby have less sequence divergence than expected, we used the model by Baldo et al. (2008) to test the similarity of endosymbionts within community. Briefly, the mean pairwise distance of endosymbionts presents in soil arthropod community were compared with an equivalent number of pairwise distances randomly selected from the database for 10000 iterations. These iterations were computed to give a null distribution for comparison with the soil endosymbiont sequence data by Wilcox rank-sum test with continuity correction and 95% confidence interval (performed in R v1.2.5). From our sampling, we found 17 unique ST's for *Wolbachia* supergroup A and 15 for*Wolbachia* supergroup B infections. Whereas, in the PubMLST database there were 228 unique ST's for A supergroup and 252 ST's for B supergroup (last checked May 2019). For the 9 samples infected with*Cardinium* and 8 for *Arsenophonus* , 248 and 228 homologues, respectively, were obtained from NCBI (last checked May, 2020). Pairwise distance was calculated using MEGA7 (Kumar *et al* ., 2015) and then corrected with Jukes and Cantor model in DNAsp v5.10.01 (Librado & Rozas, 2009). Density plots for endosymbiont divergence from the soil arthropods and the databases were plotted in R (http://R-project.org).

Results:

Morphospecies diversity and endosymbiont infection frequencies:

A total of 3509 individual arthropods were collected and sorted into 390 different morphospecies. Out of these 198 morphospecies were exclusively obtained from pitfall traps, 123 from leaf litter sampling and 69 morphospecies were obtained from both. To evaluate whether the sampling method employed yielded a significant proportion of the community diversity, we computed rarefaction analysis with EstimateS. These

provided estimates ranging from 858 (±0), obtained from Incidence coverage estimator to 600 (±32.97), obtained through Jack1 (Table S1). This indicates our sampling could capture 45-65% of the possible morphospecies in the community (Figure 1A). This is within expectations when compared to similar studies (Rhoades et al., 2017; Weller & Bossart, 2017).

The taxonomic identification of the 390 morphospecies obtained were primarily done by visual inspection and confirmed by using their *CO1* sequences in BOLD and NCBI database (Table S3). These belonged to seven classes, 24 orders, 118 families and 198 genera of arthropods. We were able to amplify *CO1* gene for 314 morphospecies. This was probably due to nucleic acid degradation as they were brought out of storage many times for visual identification, sorting and photography. Most of these samples were of single individuals (190 morphospecies) which prevented DNA extraction from additional samples.

Out of 390 morphospecies screened, approximately 47 (12.05%) morphospecies were found to be infected with *Wolbachia* . Among these, 38.30% of them belonged to Hymenopterans, 25.53% to Hemiptera, 12.77% to Diptera, 8.51% each to Araneae and Coleoptera, 4.26% to Orthoptera and 2.13% to Sarcoptiformes (Figure 1B). Two morphospecies, morph0081 and morph0085 (both Hymenoptera- Platygastridae) had multiple *Wolbachia* infections and therefore not included for further analysis. There were nine infected morphospecies for which we were unable to amplify all the five MLST genes probably because of the above-mentioned DNA quality issues. We proceeded with 36 unique host-*Wolbachia* combinations and 34 unique ST's for which we had amplified all the five MLST genes. When resultant 180 allele profiles were compared with existing sequences in PubMLST database, we found 77 new allelic profiles (14 each for *gatB* and *coxA* , 27 for*hcpA* , 12 for *ftsZ* and 10 for *fbpA*) with 30 new ST's (Table 1). For the strains labelled ST-N1 and ST-N2, unique ST could not be assigned through PubMLST, as due to DNA quality issues as only one strand of *gatB* (ST-N1, ST-N2) and *ftsZ* (ST-N1) could be amplified. As PubMLST requires chromatogram information from both strands, these were manually labelled as ST-N1 and ST-N2.

Phylogenetic analysis of MLST data using ClonalFrame showed 17 *Wolbachia* strains to cluster with known *Wolbachia*supergroup A and 15 with B supergroup while two clustered with supergroup F (Figure 2A). Supergroup A infections were predominantly found in Hymenoptera (70.5%) whereas Hemipterans (73.3%) had mostly B supergroup infections. Such taxonomic bias of *Wolbachia*supergroup has been noted previously in dipterans (Stahlhut et al., 2010), bees (Gerth, Röthe, & Bleidorn, 2013), ants (Russell et al., 2009) and in lepidopterans (Ilinsky & Kosterin, 2017).

Eleven (2.82%) of the morphospecies had *Cardinium* infections with four (33%) each from Araneae and Hymenoptera, and one each from Entomobryomorpha, Mesostigmata and Psocodea (Figure 2B). All 11 *Cardinium* strains found in this study clustered with group A *Cardinium* strains (Nakamura et al., 2009). Three morphospecies,*i.e* . morph0085 (Hymenoptera- Platygastridae), morph0152 (Hymenoptera- *Dicroscelio* sp.), morph0171 (Hymenoptera-*Trichopria* sp.) were found to be infected with both *Wolbachia* and *Cardinium* . Eight morphospecies (2.05%) had*Arsenophonus* infections with two each from Hemiptera and Hymenoptera and one each from Diptera, Entomobryomorpha, Psocodea and Thysanoptera (Figure 2C). Two morphospecies, morph0294 (Hymenoptera- Platygastridae) and morph0329 (Hemiptera- *Balclutha*) were found to be infected with both *Wolbachia* and *Arsenophonus.*Morph0085 had multiple *Wolbachia* as well as also had*Cardinium* infections, whereas Morph0328 (Psocodea-*Embidopsocus*) had both *Cardinium* and *Arsenophonus* . Therefore, multiple endosymbionts can use the same host to spread across different communities (Russell et al., 2012; Zhao, Chen, Ge, Gotoh, & Hong, 2013).

Horizontal Transfer of endosymbiont strains:

To reveal the extent of horizontal transfer events of endosymbionts across their host taxa, a qualitative assessment of phylogenetic congruency was done with host and their corresponding bacterial infections. As figure 3 reveals there is extensive horizontal transfer of the endosymbionts within the soil arthropods. A Mantel test (r) of correlation between pairwise distance of host and their corresponding endosymbiont also showed no significant correlations (Figure S5).

If endosymbionts are first moving around the host taxa of this particular community then very similar

bacterial strains would be found in taxonomically distant soil arthropods. This is precisely what we found with two distinct *Wolbachia* strains. ST-541 and ST-559 were each found in two distinct taxonomically unrelated hosts (Table 1). Morph0001 (Orthoptera- *Neonemobius*) and morph0098 (Hemiptera-*Phorodon*) were found to be infected with *Wolbachia* ST-541, whereas ST-559 was found in both morph0213 (Hemiptera-*Heteropsylla*) as well as morph0220 (Hemiptera- Delphacidae). Again, the possibility remains that these transfers could have happened independently and not correlated with the hosts being members of a particular community. But this assumes a non-parsimonious explanation that two independent events would converge on the transfer of the same *Wolbachia* ST in two different hosts.

As these bacteria are transferred around to different hosts, they are also coming in contact with each other. Whether this leads to stable multiple infections is not known, but this obviously creates opportunities for genetic exchange where the two interacting bacteria are now in a single host cytoplasm. Moreover, such co-infections can trigger selection whereby only a single endosymbiont can remain within a host. Such flux seems to be a key feature of endosymbiont dynamics, especially with *Wolbachia* , where loss is 1.5 times higher than acquisition of new infections (Bailly-Bechet et al., 2017).

Thus, this phenomenon of horizontal transfer should also create another opportunity where endosymbionts can potentially undergo recombination with each other since they are now in the same host cytoplasm.

Recombination events between endosymbiont strains:

Recombination in endosymbiont genomes is pervasive and such events significantly add to the diversification of these bacteria (Jiggins, von Der Schulenburg, Hurst, & Majerus, 2001). To check for incidence of recombination, we first analyzed the overall rates of recombination in the *Wolbachia* sequences with both ClonalFrame and RDP4. Both analyses showed a rate of nucleotide substitutions due to recombination/point mutation (r/m) of around 2.4 (95% confidence interval between 1.4- 3.7) which represents intermediate rates of recombination (Vos & Didelot, 2009). This also indicates that recombination introduces twice more nucleotide substitutions as compared to point mutation in the *Wolbachia* dataset. Unsurprisingly, the Φ test in SplitsTree also showed significant evidence of recombination (p <0.001) for the same *Wolbachia* sequences (Figure S3). However, for *Cardinium* and *Arsenophonus* , RDP4 did not indicate any evidence of recombination. This was probably due to the use of a single gene (16*S rRNA* gene) for these two bacteria.

To enumerate the recombination events within the *Wolbachia*sequences, we first looked at phylogenetic trees to check if single gene phylogenies of all the 5 MLST genes (Figure S4) differ significantly with the concatenated MLST trees (Figure 4). The next level of analysis was to use sliding window algorithms in RDP4 to locate recombination breakpoints wherever possible. All of these recombination events were then evaluated and confirmed manually. These analyses yielded several possible recombination events elaborated below.

Recombination between supergroups : Several cases of acquisition of a gene or gene segment from different supergroup were detected. Phylogenetic and network analysis of concatenated MLST dataset (Figure 4) showed *Wolbachia* ST-N2, infecting morph0343 (Hymenoptera- Encyrtidae), to cluster with B supergroup. But individual gene trees revealed that the *coxA* fragment of ST-N2 clusters with A supergroup (Figure 4) and has the allelic profile of 7. This phylogenetic disparity suggests that *coxA* gene of ST-N2 was acquired via recombination from a supergroup A *Wolbachia* . Curiously enough, *coxA* allele 7 is also found in two other *Wolbachia* infected hosts, ST-565 of morph0294 (Hymenoptera- Platygastridae) and ST-544 of morph0076 (Araneae- *Orthobula*), both with supergroup A infections (Table 1). Although it is impossible to know which *Wolbachia* strains originally underwent recombination and gave rise to the recombinant allele 7 of *coxA* , yet the presence of the same allele within the community suggests that the recombination event could have involved members within this ecological community.

Similarly, another case of recombination was observed where a B supergroup *Wolbachia* ST-560, of morph0214 (Hemiptera-*Muellerianella*), had the *coxA* gene fragment (allele profile 2) from the A supergroup (Figure 4). This recombinant *coxA* allele 2 also share sequence similarity with ST-550 and ST-571, where *coxA* alleles are different by only two base pairs (*coxA* allele profile 305) indicating that perhaps this is also

another case of recombination happening within the community.

Another case of recombination between supergroups was found with another MLST gene, *gatB* , but between supergroups A and F. The *Wolbachia* ST-552 (supergroup F), infecting morph0148 (Araneae- *Zelotes*), had a recombinant *gatB* , where the last 190 bp fragment came from the A supergroup. As the concatenated MLST tree (Figure 4) shows, ST-552 clusters with F supergroup, but the individual *gatB* gene tree shows it to be from the A supergroup. This 190 bp fragment differ by only one base pair with ST-544 infecting morph0076 (Araneae- *Orthobula*). This is also indicative of a possible recombination between these two *Wolbachia* STs belonging to two different supergroups.

Recombination within supergroups : The pervasive recombination necessitated the development of the MLST scheme for *Wolbachia* (Baldo et al., 2006) as single gene phylogenies were unable to properly represent the evolutionary history of a particular strain. In this scheme, alleles of any of the five different genes are given the same nomenclature if they share sequence identity. As table 1 shows, many of the morphospecies also share the same alleles. In fact, instead of the maximum possible number of unique alleles (180) that could have been present across the 5 MLST loci of the 36 infected morphospecies, there is only 136. This is indicative of acquisition of same alleles by recombination and are therefore, examples of within-supergroup recombination events whereby MLST fragments are exchanged across endosymbionts.

Next, we tried to identify intergenic (*i.e* ., within a particular MLST gene) recombination happening within a supergroup. Since, this detection is dependent on the algorithms present in RDP4 these estimates are inherently conservative. Most of these algorithms scans for above than expected sequence divergence in the given dataset. Therefore, recombination events happening between closely related strains and/or between regions with low variation will not be recorded as significant events by these algorithms.

There can be two types of intergenic recombination events. First, different MLST fragments (*e.g* ., between *coxA* and *gatB* of two different strains) can combine to form a chimeric gene and secondly, recombination can happen within the same MLST genes (*e.g* ., within *coxA* of two different strains). Our analysis did not find any examples of the former. This is unsurprising as all the MLST fragments are housekeeping genes and such chimeric variants will be under strong negative selection. However, eight instances of recombination within same MLST gene were found (Table 2), all within supergroup A.

How similar are endosymbiont strains within the community?

If an ecological community is the primary site of horizontal transfer of endosymbionts then the same (or very similar) bacterial strains would be found in multiple host taxa. This would result in a lower estimate of pairwise divergence among the endosymbionts present. Using the model from Baldo et al. (2008), we tested whether there is a significant reduction in the expected pairwise divergence of the endosymbionts found from the soil arthropod community. Mean pairwise distance was computed from three different sources: 1) endosymbionts within the community, 2) their expected value obtained from the equivalent number of pairwise distances randomly selected from the database, 3) all unique endosymbiont sequences obtained from PubMLST and NCBI database. These were computed separately for *Wolbachia* supergroup A, supergroup B, *Cardinium* and *Arsenophonus* . Results indicate that mean pairwise distance of *Wolbachia* supergroup A within the community (2.67%) was significantly less (Wilcox rank-sum test, $p<0.05$) than expected mean (3.54%; Table S4) and mean of all supergroup A strains in the PubMLST database (3.69%; Figure S6). In contrast, the mean pairwise distance of *Wolbachia* supergroup B strains within the community (4.17%) was significantly more (Wilcox rank-sum test, $p<0.05$) than expected mean (3.38%) and mean of B supergroup strains in the PubMLST database (3.43%). This higher than expected values for *Wolbachia* B supergroup strains can indicate presence of more divergent strains as compared to *Wolbachia* A supergroup within this community. However, when all the *Wolbachia* supergroup infections were taken together and their mean pairwise distance (8.68%) was compared with all such strains in the PubMLST database (8.66%), no significant difference was found (Wilcox rank-sum test, $p>0.05$). This perhaps indicates that although the soil arthropod community yielded several unique *Wolbachia* infections (Table 1), on average this still represents a subset of *Wolbachia* diversity reported till now. Similar to *Wolbachia* supergroup A, *Cardinium*

strains also showed similar trend where community pairwise distance (1.41%) was significantly less (Wilcox rank-sum test, $p<0.05$) than expected mean (2.48%) and mean of strains obtained from the database (2.01%). Whereas mean pairwise distance of *Arsenophonus* strains within the community (1.19%) was not significantly different ($p>0.05$) from the expected mean (1.38%) as well as from mean of strains obtained from the database (1.55%). Thus, *Wolbachia* supergroup A as well as *Cardinium* strains within the community are more closely related among themselves (Table S4) but not *Wolbachia* supergroup B and *Arsenophonus* .

Discussion:

In this study, we evaluated sequence divergence and incidence of recombination in three major endosymbionts (*Wolbachia* , *Cardinium* and *Arsenophonus*) to answer whether the ecological community that they are a part of is the primary seat of their horizontal transfer and diversification. We used soil arthropod community because it is relatively insular and has a relatively high habitat endemicity of the residents. Our main goal was to assess whether community members facilitate the spread of endosymbionts as they themselves come in contact with each other for various ecological interactions. To properly assess whether ecological communities are indeed the seat of endosymbiont transfer and diversification, one needed to compare multi-gene phylogenies of such endosymbiont surveys from different communities. However, in spite our extensive literature surveys we could not find any such previous reports. Most surveys of arthropod communities concentrated on the hosts rather than on their endosymbionts (Gonçalves, Pereira, & Liu, 2012). Some studies like Kittayapong et al. (2003) and Sintupachee et al. (2006) did uncover the resident endosymbionts but mostly with single genes. This precluded a cogent comparison of endosymbiont diversity and incidence of recombination with the present study. Another set of studies did indeed sample endosymbionts with multi-gene sequences but concentrated on a few, and not all, host taxa within a community (Bing et al., 2014). Again, such studies are not ideal comparisons with the present one as these were biased towards a few host taxa. To partially overcome this problem, we used statistical models with extensive resampling. We observed that the supergroup A *Wolbachia* infections and *Cardinium* do indeed show less pairwise divergence, than expected, in accordance with our predictions. However, supergroup B *Wolbachia* and *Arsenophonus* infections did not show this pattern. In fact, the former shows more variation than expected whereas *Arsenophonus* shows no significant difference. This indicates that these endosymbionts have different propensity and/or rates of horizontal transfer within the community. We speculate what can be the reasons behind this.

Are some endosymbionts more prone to horizontal transfer and recombination?

One explanation for the observed patterns could be the relative ease with which supergroup A *Wolbachia* and *Cardinium* can undergo horizontal transfer and recombination as compared to supergroup B *Wolbachia* and *Arsenophonus* . This essentially means that the former two endosymbionts would encounter previously existing bacterial infections within their hosts which would increase the opportunity for recombination among the pre-existing and the new bacterial strains. Recombination would then create newer allele variants. This is indeed borne out by the results in table 1 which depicts the number of unique alleles found in this study among the *Wolbachia* infections. In all about 84% (71 out of a possible 85 alleles) of the A supergroup infection are unique. Whereas, about 75% (56 out of a possible 75 alleles) are unique in B supergroup *Wolbachia* infections. Furthermore, as indicated in table 2, the number of within supergroup recombination detected in the A supergroup strains (8 instances) far outnumber the B supergroup *Wolbachia* , where none were detected. This is in spite of horizontal transfer of the entire B supergroup ST's (ST-541 and ST-559) to taxonomically unrelated hosts (Table 1). An expected outcome of such pervasive horizontal transfer and resulting recombination would have been an increase in sequence diversity in the A supergroup strains, especially, if the source of recombination had been infections outside the community. This does not seem to be the case as the A supergroup infection show less than expected pairwise distance (2.67%) when compared with the B supergroup infections (4.17%). This indicates that the sources of recombination must be from infections within this community. In other words, the standing sequence variation of the A supergroup infections is being partitioned across the community-wide arthropod taxa into newer recombinants with resulting increase in allele diversity but not overall sequence divergence. Moreover, what follows from this relatively low pairwise divergence of the A supergroup infections is that this horizontal transfer and recombination must have been

recent or rapid enough for any post-recombination sequence variation to accumulate. This indicates that the A supergroup infections are either better at horizontal transfer across the community or are presently undergoing such rapid transfers as has been suggested by Werren et al (1995). On the other hand, the B supergroup *Wolbachia* infections show relatively diverged strains with low rates of recombination indicating much stable infections. Since, little is known about the biological characteristic of the different *Wolbachia* supergroups, other than sequence divergence, it is difficult to speculate whether there are supergroup specific effects on their hosts. For example, it is not known whether any supergroup infections exclusively infect any specific arthropod taxa or whether any supergroup make hosts more prone to horizontal transfer? Therefore, we concentrate on specific trophic interactions of the hosts themselves and try to explain why supergroup A infections show such extensive horizontal spread.

Are parasitoids serving as the conduit for the spread of endosymbionts?

Parasitoids can serve as a driving force for the horizontal transfer of endosymbionts (Haine, Pickup, & Cook, 2005) as their lifestyle entails close cellular and tissue contact with their host. Horizontal transfer involving parasitoids is generally unidirectional (from host to parasitoid) because they usually end up killing the host. But parasitoids can also act as phoretic vectors (Ahmed et al., 2015; Gehrer & Vorburger, 2012) and can transmit *Wolbachia* by sequential probing of infected and uninfected hosts. Moreover, horizontal transfers can also happen between parasitoids if infected and uninfected larval parasitoids share the same host (Huigens, De Almeida, Boons, Luck, & Stouthamer, 2004). Such habits can also facilitate multiple infections if parasitoids infecting same host have different endosymbiont infections. As these multiple infections come in close contact, they can undergo recombination. Such parasitoid-mediated horizontal transfer could be an explanation for the distribution of A supergroup *Wolbachia* infections in our sampling. Out of the 17 *Wolbachia* A supergroup STs found, nine STs were found in morphospecies which are parasitoid wasps (Hymenoptera). Seven STs were found from Platygastridae and one each from Bethylidae and Diapriidae (Table 1). Therefore, the comparatively higher incidence of recombination in *Wolbachia* supergroup A infections could be due to their presence in parasitoid hosts (Mascarenhas et al., 2016). Similarly, predators and parasites can also be conduits for the spread of endosymbionts as speculated by predation of infected *Armadillidium vulgare* by uninfected *Porcellio dilatatus*(Le Clec'h et al., 2013) and predatory mite *Metaseiulus occidentalis* and its prey, *Tetranychus urticae* (Hoy & Jeyaprakash, 2005). In soil arthropod community, we also observed incidences of *Wolbachia* recombination amongst predators. *Wolbachia* F supergroup ST-552, infecting morph0148 (Araneae-*Zelotes*), had a fragment of *gatB* gene similar to A supergroup ST-544 infecting morph0076 (Araneae- *Orthobula*). Also,*fbpA* gene of ST-544 was found to have probably recombined with ST-570 infecting Morph0375 (Coleoptera- *Monolepta sp.*). Apart from such trophic interactions within the community, if any member of a community interacts with individuals of other community then it is likely to spread endosymbionts from one community to other.

How are endosymbionts spreading from one ecological community to another?

Ecological communities are a diverse assemblage of many different species involved in a web of interactions with each other (Agrawal et al., 2007). However, rarely, are such communities isolated from each other. There are certain members which are relatively cosmopolitan and interconnect with members of multiple communities (Stireman & Singer, 2003) leading to a metacommunity-wide distribution (Brown et al., 2020). The distribution of these arthropods can lead to horizontal transfer of their resident endosymbionts to distinct ecological communities. Within the soil arthropod community, we have found one such example which can potentially be a source of horizontal transfer of endosymbionts across many other communities. The macropterous form of the planthopper *Nilaparvata lugens* (morph0111, BOLD ID SAEVG089-20, Table S3) was found from leaf litter sampling. *N . lugens* is a highly destructive pest of rice across tropical Asia and can also survive on other tropical grass species (Khan, Saxena, & Rueda, 1988). It is known to migrate long distances in search of actively growing rice plants (Riley, Smith, & Reynolds, 2003). The presence of *N .lugens* is unsurprising as our sampling season (October) coincided with the rice harvesting season in North India. *N . lugens*is known to be infected with several endosymbionts like *Wolbachia*and *Arsenophnonus* (Qu et al., 2013). In the present study, it was found to be infected with *Wolbachia* ST-163 from the B supergroup.

The same *Wolbachia* sequence type has also been reported from *N . lugens* from Southern China (Zhang, Han, & Hong, 2013). This indicates that such invasive pest species can potentially introduce their resident endosymbionts into many different arthropod communities.

Conversely, the presence of very similar endosymbionts in geographically distinct locations can indicate their spread from one ecological community to other. The *Wolbachia* B supergroup, ST-41, has been detected from a phorid fly (morph0285) in our dataset. The same ST-41 has been found in calyptrate flies (Stahlhut et al., 2010) as well as from several other lepidopterans (Ilinsky & Kosterin, 2017; Narita et al., 2011; Russell et al., 2009; Salunke et al., 2012). This is not unexpected given the diversity of *Wolbachia* infections. However, what is unexpected is the location of the hosts with ST-41 ranges from North America, Africa, Russia, South and South-Eastern Asia all the way to Japan. Unfortunately, it is difficult to conjecture about the reasons behind such a huge range, as corroborating community-wide data is lacking.

The above two instances testify to the utility of a MLST based approach to understand *Wolbachia* diversity and spread across global arthropod communities. Moreover, these cases also highlight the importance of collecting community-wide data to understand the probable chain of transfer of these bacteria. Such data can also illuminate similar connections for the spread of *Arsenophonus* and *Cardinium* if employed with multi-locus data (Jousselin, Cœur d'Acier, Vanlerberghe-Masutti, & Duron, 2013; Stouthamer, Kelly, Mann, Schmitz-Esser, & Hunter, 2019).

A major goal of endosymbiont research is to explain the tempo and mode of their spread across arthropod communities across the world. We contend that evaluating endosymbiont diversity within specific ecological communities is the key to understand this spread. Such studies would give us specific examples of bacterial strains that are better at spreading as well as uncover specific ecological roles of arthropod hosts which are more amenable to horizontal transfer of their resident endosymbionts. As data from such studies accumulate higher level patterns will emerge which can then be empirically tested.

Acknowledgements:

We would like to thank IISER Mohali for providing the experimental facilities and funding for this work. We acknowledge all EVOGEN lab members for their comments on manuscript. MG thank University Grant Commission (UGC) for fellowship (ID: 325974). AG is grateful to ICAR for necessary facilities. RK was funded by post-doctoral fellowship from IISER-Mohali. We would like to thank Professor Raman Rajagopal, Department to Zoology, University of Delhi, for providing us with *Cardinium* and *Arsenophonus* positive controls and Yogesh Bangal for help in field sampling.

References:

Agrawal, A. A., Ackerly, D. D., Adler, F., Arnold, A. E., Cáceres, C., Doak, D. F., . . . Mooney, K. A. (2007). Filling key gaps in population and community ecology. *Frontiers in Ecology and the Environment, 5* (3), 145-152.

Ahmed, M. Z., Breinholt, J. W., & Kawahara, A. Y. (2016). Evidence for common horizontal transmission of *Wolbachia* among butterflies and moths. *BMC evolutionary biology, 16* (1), 1-16.

Ahmed, M. Z., Li, S.-J., Xue, X., Yin, X.-J., Ren, S.-X., Jiggins, F. M., . . . Qiu, B.-L. (2015). The intracellular bacterium *Wolbachia* uses parasitoid wasps as phoretic vectors for efficient horizontal transmission. *PLoS Pathog, 11* (2), e1004672.

Bailly-Bechet, M., Martins-Simões, P., Szöllősi, G. J., Mialdea, G., Sagot, M.-F., & Charlat, S. (2017). How long does *Wolbachia* remain on board? *Molecular Biology and Evolution, 34* (5), 1183-1193.

Baldo, L., Ayoub, N. A., Hayashi, C. Y., Russell, J. A., Stahlhut, J. K., & Werren, J. H. (2008). Insight into the routes of *Wolbachia* invasion: high levels of horizontal transfer in the spider genus *Agelenopsis* revealed by *Wolbachia* strain and mitochondrial DNA diversity. *Molecular Ecology, 17* (2), 557-569.

Baldo, L., Dunning Hotopp, J. C., Jolley, K. A., Bordenstein, S. R., Biber, S. A., Choudhury, R. R., . . .

Werren, J. H. (2006). Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis* .*Appl Environ Microbiol, 72* (11), 7098-7110.

Baudry, E., Bartos, J., Emerson, K., Whitworth, T., & Werren, J. H. (2003). *Wolbachia* and genetic variability in the birdnest blowfly*Protocalliphora sialia* . *Mol Ecol, 12* (7), 1843-1854.

Bing, X. L., Xia, W. Q., Gui, J. D., Yan, G. H., Wang, X. W., & Liu, S. S. (2014). Diversity and evolution of the *Wolbachia* endosymbionts of *Bemisia* (Hemiptera: Aleyrodidae) whiteflies. *Ecol Evol, 4* (13), 2714-2737.

Brown, A. N., & Lloyd, V. K. (2015). Evidence for horizontal transfer of *Wolbachia* by a *Drosophila* mite. *Exp Appl Acarol, 66* (3), 301-311.

Brown, J. J., Mihaljevic, J. R., Des Marteaux, L., & Hrček, J. (2020). Metacommunity theory for transmission of heritable symbionts within insect communities. *Ecology and evolution, 10* (3), 1703-1721.

Bruen, T. C., Philippe, H., & Bryant, D. (2006). A simple and robust statistical test for detecting the presence of recombination.*Genetics, 172* (4), 2665-2681.

Bryant, D., & Moulton, V. (2004). Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol, 21* (2), 255-265.

Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of statistics* , 265-270.

Chao, A., Hwang, W.-H., Chen, Y., & Kuo, C. (2000). Estimating the number of shared species in two communities. *Statistica sinica* , 227-246.

Colwell, R. K. (2013). EstimateS: Statistical estimation of species richness and shared species from samples. Version 9 and earlier. User's Guide and application. *Published at: http://purl. oclc. org/estimates* .

Dewayne Shoemaker, D., Machado, C. A., Molbo, D., Werren, J. H., Windsor, D. M., & Herre, E. A. (2002). The distribution of*Wolbachia* in fig wasps: correlations with host phylogeny, ecology and population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 269* (1506), 2257-2267.

Didelot, X., & Falush, D. (2007). Inference of bacterial microevolution using multilocus sequence data. *Genetics, 175* (3), 1251-1266.

Duron, O., Bouchon, D., Boutin, S., Bellamy, L., Zhou, L., Engelstädter, J., & Hurst, G. D. (2008). The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC biology, 6* (1), 27.

Ellegaard, K. M., Klasson, L., Naslund, K., Bourtzis, K., & Andersson, S. G. (2013). Comparative genomics of *Wolbachia* and the bacterial species concept. *PLoS Genet, 9* (4), e1003381.

Gebiola, M., Giorgini, M., Kelly, S. E., Doremus, M. R., Ferree, P. M., & Hunter, M. S. (2017). Cytological analysis of cytoplasmic incompatibility induced by *Cardinium* suggests convergent evolution with its distant cousin *Wolbachia* . *Proc Biol Sci, 284* (1862), 20171433.

Gehrer, L., & Vorburger, C. (2012). Parasitoids as vectors of facultative bacterial endosymbionts in aphids. *Biology letters, 8* (4), 613-615.

Gerth, M., Röthe, J., & Bleidorn, C. (2013). Tracing horizontal*Wolbachia* movements among bees (Anthophila): a combined approach using multilocus sequence typing data and host phylogeny.*Molecular Ecology, 22* (24), 6149-6162.

Gonçalves, M. F., Pereira, J. A., & Liu, T. (2012). Abundance and diversity of soil arthropods in the olive grove ecosystem. *Journal of Insect Science, 12* (1).

Gonella, E., Pajoro, M., Marzorati, M., Crotti, E., Mandrioli, M., Pontini, M., . . . Alma, A. (2015). Plant-mediated interspecific horizontal transmission of an intracellular symbiont in insects.*Sci Rep, 5* , 15811.

Haine, E. R., Pickup, N. J., & Cook, J. M. (2005). Horizontal transmission of *Wolbachia* in a *Drosophila* community.*Ecological Entomology, 30* (4), 464-472.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series* , *41* , 95-98.

Hebert, P. D., Cywinska, A., Ball, S. L., & Dewaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270* (1512), 313-321.

Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A., & Werren, J. H. (2008). How many species are infected with*Wolbachia* ?–a statistical analysis of current data. *FEMS microbiology letters, 281* (2), 215-220.

Hoy, M. A., & Jeyaprakash, A. (2005). Microbial diversity in the predatory mite *Metaseiulus occidentalis* (Acari: Phytoseiidae) and its prey, *Tetranychus urticae* (Acari: Tetranychidae).*Biological Control, 32* (3), 427-441.

Huigens, M., De Almeida, R., Boons, P., Luck, R., & Stouthamer, R. (2004). Natural interspecific and intraspecific horizontal transfer of parthenogenesis–inducing *Wolbachia* in *Trichogramma* wasps.*Proceedings of the Royal Society of London. Series B: Biological Sciences, 271* (1538), 509-515.

Huson, D. H., Kloepper, T., & Bryant, D. (2008). SplitsTree 4.0-Computation of phylogenetic trees and networks.*Bioinformatics, 14* , 68-73.

Ilinsky, Y., & Kosterin, O. E. (2017). Molecular diversity of*Wolbachia* in Lepidoptera: Prevalent allelic content and high recombination of MLST genes. *Mol Phylogenet Evol, 109* , 164-179.

Jiggins, F. M., von Der Schulenburg, J. H., Hurst, G. D., & Majerus, M. E. (2001). Recombination confounds interpretations of *Wolbachia*evolution. *Proc Biol Sci, 268* (1474), 1423-1427.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: a better web interface.*Nucleic Acids Res, 36* , W5-9.

Jolley, K. A., Bray, J. E., & Maiden, M. C. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST. org website and their applications. *Wellcome open research, 3* .

Jousselin, E., Cœur d'Acier, A., Vanlerberghe-Masutti, F., & Duron, O. (2013). Evolution and diversity of *Arsenophonus* endosymbionts in aphids. *Molecular Ecology, 22* (1), 260-270.

Khan, Z., Saxena, R., & Rueda, B. (1988). Responses of rice-infesting and grass-infesting populations of *Nilaparvata lugens* (Homoptera: Delphacidae) to rice plants and Leersia grass and to their steam distillate extracts. *Journal of economic entomology, 81* (4), 1080-1088.

Kittayapong, P., Jamnongluk, W., Thipaksorn, A., Milne, J., & Sindhusake, C. (2003). *Wolbachia* infection complexity among insects in the tropical rice-field community. *Molecular Ecology, 12* (4), 1049-1060.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.*Mol Biol Evol, 33* (7), 1870-1874.

Landmann, F. (2019). The *Wolbachia* Endosymbionts. *Microbiol Spectr, 7* (2), 139-153.

Le Clec'h, W., Chevalier, F. D., Genty, L., Bertaux, J., Bouchon, D., & Sicard, M. (2013). Cannibalism and predation as paths for horizontal passage of *Wolbachia* between terrestrial isopods. *PloS one, 8* (4), e60232.

Legendre, P., & Legendre, L. F. (2012). *Numerical ecology* (Vol. 24). Philadelphia: Elsevier.

Li, S.J., Ahmed, M. Z., Lv, N., Shi, P.Q., Wang, X.M., Huang, J.L., & Qiu, B.L. (2017). Plant mediated horizontal transmission of*Wolbachia* between whiteflies. *The ISME journal, 11* (4), 1019-1028.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics, 25* (11), 1451-1452.

Maiden, M. C., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., . . . Caugant, D. A. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences, 95* (6), 3140-3145.

Malloch, G., & Fenton, B. (2005). Super-infections of *Wolbachia*in byturid beetles and evidence for genetic transfer between A and B super-groups of *Wolbachia* . *Molecular Ecology, 14* (2), 627-637.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol, 1* (1), vev003.

Mascarenhas, R. O., Prezotto, L. F., Perondini, A. L., Marino, C. L., & Selivon, D. (2016). *Wolbachia* in guilds of *Anastrepha*fruit flies (Tephritidae) and parasitoid wasps (Braconidae). *Genet Mol Biol, 39* (4), 600-610.

Morrow, J. L., Frommer, M., Shearman, D., & Riegler, M. (2014). Tropical tephritid fruit fly community with high incidence of shared *Wolbachia* strains as platform for horizontal transmission of endosymbionts. *Environmental Microbiology, 16* (12), 3622-3637.

Mouton, L., Thierry, M., Henri, H., Baudin, R., Gnankine, O., Reynaud, B., . . . Delatte, H. (2012). Evidence of diversity and recombination in*Arsenophonus* symbionts of the Bemisia tabacispecies complex.*BMC microbiology, 12* (S1), S10.

Nakamura, Y., Kawai, S., Yukuhiro, F., Ito, S., Gotoh, T., Kisimoto, R., . . . Noda, H. (2009). Prevalence of *Cardinium* bacteria in planthoppers and spider mites and taxonomic revision of "*Candidatus Cardinium hertigii* " based on detection of a new *Cardinium* group from biting midges. *Applied Environmental Microbiology, 75* (21), 6757-6763.

Narita, S., Kageyama, D., Hiroki, M., Sanpei, T., Hashimoto, S., Kamitoh, T., & Kato, Y. (2011). *Wolbachia* -induced feminisation newly found in *Eurema hecabe* , a sibling species of *Eurema mandarina* (Lepidoptera: Pieridae). *Ecological Entomology, 36* (3), 309-317.

Olson, D. M. (1991). A comparison of the efficacy of litter sifting and pitfall traps for sampling leaf litter ants (Hymenoptera, Formicidae) in a tropical wet forest, Costa Rica. *Biotropica* , 166-172.

Qu, L.-Y., Lou, Y.-H., Fan, H.-W., Ye, Y.-X., Huang, H.-J., Hu, M.-Q., . . . Zhang, C.-X. (2013). Two endosymbiotic bacteria, *Wolbachia*and *Arsenophonus* , in the brown planthopper *Nilaparvata lugens* . *Symbiosis, 61* (1), 47-53.

Querner, P., & Bruckner, A. (2010). Combining pitfall traps and soil samples to collect Collembola for site scale biodiversity assessments.*Applied Soil Ecology, 45* (3), 293-297.

Rambaut, A. (2009). FigTree. Tree figure drawing tool.*http://tree. bio. ed. ac. uk/software/figtree/* .

Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). *Molecular ecology notes, 7* (3), 355-364.

Raychoudhury, R., Baldo, L., Oliveira, D. C., & Werren, J. H. (2009). Modes of acquisition of *Wolbachia* : horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex.*Evolution, 63* (1), 165-183.

Rhoades, P., Griswold, T., Waits, L., Bosque-Pérez, N. A., Kennedy, C. M., & Eigenbrode, S. D. (2017). Sampling technique affects detection of habitat factors influencing wild bee communities. *Journal of Insect Conservation, 21* (4), 703-714.

Riley, J. R., Smith, A., & Reynolds, D. (2003). The feasibility of using vertical-looking radar to monitor the migration of brown planthopper and other insect pests of rice in China. *Insect Science, 10* (1), 1-19.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Hohna, S., . . . Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space.

Syst Biol, 61 (3), 539-542.

Rowley, D. L., Coddington, J. A., Gates, M. W., Norrbom, A. L., Ochoa, R. A., Vandenberg, N. J., & Greenstone, M. H. (2007). Vouchering DNA-barcoded specimens: test of a nondestructive extraction protocol for terrestrial arthropods. *Molecular ecology notes, 7* (6), 915-924.

Russell, J. A., Funaro, C. F., Giraldo, Y. M., Goldman-Huertas, B., Suh, D., Kronauer, D. J., . . . Pierce, N. E. (2012). A veritable menagerie of heritable bacteria from ants, butterflies, and beyond: broad molecular surveys and a systematic review. *PloS one, 7* (12), e51027.

Russell, J. A., Goldman-Huertas, B., Moreau, C. S., Baldo, L., Stahlhut, J. K., Werren, J. H., & Pierce, N. E. (2009). Specialization and geographic isolation among *Wolbachia* symbionts from ants and lycaenid butterflies. *Evolution: International Journal of Organic Evolution, 63* (3), 624-640.

Sabu, T. K., & Shiju, R. T. (2010). Efficacy of pitfall trapping, Winkler and Berlese extraction methods for measuring ground-dwelling arthropods in moistdeciduous forests in the Western Ghats. *Journal of Insect Science, 10* (1).

Salunke, B. K., Salunkhe, R. C., Dhotre, D. P., Walujkar, S. A., Khandagale, A. B., Chaudhari, R., . . . Shouche, Y. S. (2012). Determination of *Wolbachia* diversity in butterflies from Western Ghats, India, by a multigene approach. *Appl Environ Microbiol, 78* (12), 4458-4467.

Semiatizki, A., Weiss, B., Bagim, S., Rohkin-Shalom, S., Kaltenpoth, M., & Chiel, E. (2020). Effects, interactions, and localization of*Rickettsia* and *Wolbachia* in the house fly parasitoid,*Spalangia endius* . *Microb Ecol, 80* (3), 718-728.

Sintupachee, S., Milne, J., Poonchaisri, S., Baimai, V., & Kittayapong, P. (2006). Closely related *Wolbachia* strains within the pumpkin arthropod community and the potential for horizontal transmission via the plant. *Microbial ecology, 51* (3), 294-301.

Smith, E. P., & van Belle, G. (1984). Nonparametric estimation of species richness. *Biometrics* , 119-129.

Stahlhut, J. K., Desjardins, C. A., Clark, M. E., Baldo, L., Russell, J. A., Werren, J. H., & Jaenike, J. (2010). The mushroom habitat as an ecological arena for global exchange of *Wolbachia* .*Molecular Ecology, 19* (9), 1940-1952.

Stireman, J. O., & Singer, M. S. (2003). What determines host range in parasitoids? An analysis of a tachinid parasitoid community.*Oecologia, 135* (4), 629-638.

Stouthamer, C. M., Kelly, S. E., Mann, E., Schmitz-Esser, S., & Hunter, M. S. (2019). Development of a multi-locus sequence typing system helps reveal the evolution of *Cardinium hertigii* , a reproductive manipulator symbiont of insects. *BMC Microbiol, 19* (1), 266.

Vavre, F., Fleury, F., Lepetit, D., Fouillet, P., & Bouletreau, M. (1999). Phylogenetic evidence for horizontal transmission of*Wolbachia* in host-parasitoid associations. *Mol Biol Evol, 16* (12), 1711-1723.

Vos, M., & Didelot, X. (2009). A comparison of homologous recombination rates in bacteria and archaea. *ISME J, 3* (2), 199-208.

Weinert, L. A., Araujo-Jnr, E. V., Ahmed, M. Z., & Welch, J. J. (2015). The incidence of bacterial endosymbionts in terrestrial arthropods.*Proc Biol Sci, 282* (1807), 20150249.

Weller, M., & Bossart, J. (2017). Insect community diversity tracks degradation and recovery of a wastewater assimilation marsh in southeast Louisiana. *Wetlands, 37* (4), 661-673.

Werren, J. H., Baldo, L., & Clark, M. E. (2008). *Wolbachia* : master manipulators of invertebrate biology. *Nat Rev Microbiol, 6* (10), 741-751.

Werren, J. H., & Bartos, J. D. (2001). Recombination in*Wolbachia* . *Curr Biol, 11* (6), 431-435.

Werren, J. H., Windsor, D., & Guo, L. R. (1995). Distribution of *Wolbachia* among neotropical arthropods. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 262* (1364), 197-204.

Werren, J. H., Zhang, W., & Guo, L. R. (1995). Evolution and phylogeny of *Wolbachia* : reproductive parasites of arthropods. *Proc Biol Sci, 261* (1360), 55-63.

Yahara, K., Didelot, X., Jolley, K. A., Kobayashi, I., Maiden, M. C., Sheppard, S. K., & Falush, D. (2016). The Landscape of Realized Homologous Recombination in Pathogenic Bacteria. *Mol Biol Evol, 33* (2), 456-471.

Zchori-Fein, E., & Perlman, S. J. (2004). Distribution of the bacterial symbiont *Cardinium* in arthropods. *Mol Ecol, 13* (7), 2009-2016.

Zele, F., Santos, I., Olivieri, I., Weill, M., Duron, O., & Magalhaes, S. (2018). Endosymbiont diversity and prevalence in herbivorous spider mite populations in South-Western Europe. *FEMS microbiology ecology, 94* (4), fiy015.

Zhang, K. J., Han, X., & Hong, X. Y. (2013). Various infection status and molecular evidence for horizontal transmission and recombination of *Wolbachia* and *Cardinium* among rice planthoppers and related species. *Insect Science, 20* (3), 329-344.

Zhao, D. X., Chen, D. S., Ge, C., Gotoh, T., & Hong, X. Y. (2013). Multiple infections with *Cardinium* and two strains of *Wolbachia* in the spider mite *Tetranychus phaselus* Ehara: revealing new forces driving the spread of *Wolbachia* . *PloS one, 8* (1), e54964.

Data Accessibility:

All sequences obtained from this study were deposited in NCBI. Accession numbers are: for *CO1* data- MN447522- MN447531, MN520845- MN521147, MN901899; *Wolbachia gatB* - MN594583- MN594618;*coxA* - MN594619- MN594654; *hcpA* - MN594655- MN594690;*ftsZ* - MN594691- MN594726; *fbpA* - MN594727- MN594762;*Cardinium* 16*S* - MN594564- MN594574; *Arsenophonus*16*S* - MN594575- MN594582 (Table S3). *Wolbachia* MLST data was also deposited on PubMLST database having ID ST541-544, 547-548, 550, 552-560, 562-575 (Table 1). Morphospecies images along with their corresponding *CO1* gene sequences were also deposited in BOLD database having process ID SAEVG001-20: SAEVG314-20.

Author Contributions:

MG and RR designed the study. RK designed and performed sampling of soil arthropods and sorting of morphospecies. AG identified insect samples. MG did the screening, barcoding, sequencing and analysis of the data. MG and RR wrote the paper.

Figure 1 : A) Rarefaction curve of morphospecies found (in black) showing species richness in the soil arthropod community. Colored lines represent expected number of morphospecies. B) The distribution of three endosymbionts screened across different host arthropod orders.

Figure 2 : Phylogenetic analysis of (A) *Wolbachia,* (B) *Cardinium* and (C) *Arsenophonus* found, shown with some known sequences for better resolution. *Wolbachia* phylogenetic tree was constructed using MLST data in ClonalFrame with 50% majority rule consensus. *Cardinium* and *Arsenophonus* phylogeny was made in MEGA7 using 16*S rRNA* gene fragment. Dotted nodes represent bootstrap value >50. *Wolbachia* infections are shown as STs whereas *Cardinium* and *Arsenophonus* are labelled with host taxa that they infected. Infections obtained in this study are in BOLD. *Brugia malayi, Amoebophilus asiaticus* and*Proteus mirabilis* were taken as outgroup for *Wolbachia* ,*Cardinium* and *Arsenophonus* phylogenetic analysis, respectively.

Figure 3 : Association between infected host (left) and endosymbiont (right) phylogeny with (A) *Wolbachia* A supergroup, (B) *Wolbachia* B supergroup, (C) *Cardinium* and (D)*Arsenophonus* Host phylogeny was constructed by MrBayes using partial *CO1* sequences. Black dots represent clade credibility value >70. Phylogenetic relationship between different *Wolbachia* strains was interpreted through ClonalFrame.*Cardinium*

and *Arsenophonus* 16*S rDNA* phylogeny was constructed using MEGA7. Correlation between the two phylogenies suggest phylogenetic incongruence and extensive horizontal transfer across host taxa.

Figure 4 : Maximum likelihood phylogenetic trees of *coxA*(left), concatenated MLST dataset (centre) and *gatB* (right) gene made using MEGA7. Black dots represent bootstrap value >50.*Wolbachia* ST-N2 and ST-560 clustered with B supergroup in concatenated MLST phylogenetic tree whereas in *coxA* phylogenetic tree, these strains clustered with A supergroup indicating recombination between *Wolbachia* supergroup A and B. Similarly, ST-552 clustered with F supergroup in concatenated MLST tree, but the individual*gatB* gene tree shows it to be from A supergroup, indicating recombination between A and F supergroup. These three cases (ST-N2, ST-560, and ST-552) represents between supergroup recombination of gene or gene segment.

Table 1 : Allele profiles of MLST genes for 36 unique host-*Wolbachia* combinations. (Bold numbers represent new alleles and STs).

Table 2 : Recombination events detected in the *Wolbachia*MLST sequences. Putative breakpoints indicate concatenated sequences of MLST genes in the order (*gatB-coxA-hcpA-ftsZ-fbpA*). p-value was kept at < 0.01.

Hosted file

`Main Figures_MG.pdf` available at https://authorea.com/users/377606/articles/494252-are-ecological-communities-the-seat-of-endosymbiont-horizontal-transfer-and-diversification-a-case-study-with-soil-arthropod-community

Hosted file

`Main Tables_MG.pdf` available at https://authorea.com/users/377606/articles/494252-are-ecological-communities-the-seat-of-endosymbiont-horizontal-transfer-and-diversification-a-case-study-with-soil-arthropod-community