environmental microbiology

Environmental Microbiology (2022) **24**(8), 3463–3485 doi:10.1111/1462-2920.15720

Genetic variation is associated with differences in facilitative and competitive interactions in the *Rhizobium leguminosarum* species complex

Bryden Fields,[1] **Emma K. Moffat**,[1,2] **Ellie Harrison**,[2] **Stig U. Andersen**,[3] **J. Peter W. Young**[1] **and Ville-Petri Friman**[1*]

[1]*Department of Biology, University of York, York, UK.*

[2]*Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK.*

[3]*Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark.*

Summary

Competitive and facilitative interactions influence bacterial community composition, diversity and functioning. However, the role of genetic diversity for determining interactions between coexisting strains of the same, or closely related, species remains poorly understood. Here, we investigated the type (facilitative/inhibitory) and potential underlying mechanisms of pairwise interactions between 24 genetically diverse bacterial strains belonging to three genospecies (gsA,C,E) of the *Rhizobium leguminosarum* species complex. Interactions were determined indirectly, based on secreted compounds in cell-free supernatants, and directly, as growth inhibition in cocultures. We found supernatants mediated both facilitative and inhibitory interactions that varied greatly between strains and genospecies. Overall, gsE strains indirectly suppressed growth of gsA strains, while their own growth was facilitated by other genospecies' supernatants. Similar genospecies-level patterns were observed in direct competition, where gsA showed the highest susceptibility and gsE the highest inhibition capacity. At the genetic level, increased gsA susceptibility was associated with a non-random distribution of quorum sensing and secondary metabolite genes across genospecies. Together, our results suggest that genetic variation is associated with facilitative and competitive interactions, which could be important ecological mechanisms explaining *R. leguminosarum* diversity.

Received 17 March, 2021; revised 8 July, 2021; accepted 10 August, 2021. *For correspondence. Department of Biology, University of York, Wentworth Way, York, YO10 5DD, UK.
E-mail ville.friman@york.ac.uk

ORIGINALITY-SIGNIFICANCE STATEMENT

Understanding the link between genetic and phenotypic variation is central for evolutionary ecology. While functional differences between different bacterial species have been widely characterized, the role of genetic diversity in determining functional differences between different strains of the same, and closely related, species is less well understood. Here we show that genetic variation within the *R. leguminosarum* species complex is associated with differences facilitative and competitive interactions, which are more consistent at the genospecies compared to strain level. These interaction differences could be further statistically associated with variation in metabolic capacity and the presence of several genes associated with direct inhibition and cell-to-cell quorum sensing signalling. Together, our results demonstrate a clear link between genetic and phenotypic variation in *R. leguminosarum*, suggesting that competition and facilitation could be important ecological drivers behind *R. leguminosarum* diversification.

Introduction

Bacteria predominantly exist in complex, diverse communities where interacting species and genotypes coexist side by side despite competition for the same resources and niches (Stubbendieck and Straight, 2016; Gorter *et al*., 2020). Bacterial interactions, thus, range from facilitative to competitive, affecting diversification, spatial distribution and abundance of different species and generating functional attributes at the community level that might be unobservable when only individual species are considered (Gorter *et al*., 2020). The effects of bacterial interactions can further cascade through trophic networks, affecting pathogen or symbiont abundance and their ability to interact with their eukaryotic hosts (Van Der Heijden *et al*., 2008; Jousset *et al*., 2011; Becker *et al*., 2012; Peters *et al*., 2012; Braga *et al*., 2016; Lee *et al*., 2016; Huang *et al*., 2017). For example, increased antagonistic interactions within communities of *Pseudomonas fluorescens*, a plant mutualist with the ability to inhibit plant pathogenic bacteria, led to a reduction in bacterial density and loss of plant protection from pathogens (Becker *et al*., 2012). In addition, facilitative interactions have been observed between two symbionts of the marine oligochaete, *Olavius algarvensis*, generating an endosymbiotic syntrophic sulphur cycle which aids the

hosts in the removal of undesirable fermentation end products, providing energy and increased protein yields for both symbionts (Dubiller *et al*., 2001). Previous studies have predominantly investigated bacterial interactions at the species level, focusing on species richness and community composition using, for example, 16S rRNA amplicon sequencing (Langille *et al*., 2013; Birtel *et al*., 2015; Burke and Darling, 2016; Fuks *et al*., 2018). However, bacterial species are highly diverse and this intraspecies genetic variation is often adaptive and linked with variation in bacterial fitness (Levin, 1988; Rainey and Travisano, 1993). As a result, acknowledging both intra- and interspecies bacterial diversity is vital for understanding microbial community dynamics from assembly to functioning (Jousset *et al*., 2011; Lee *et al*., 2016; Stubbendieck and Straight, 2016; Gorter *et al*., 2020).

Inter- and intraspecific competition between bacterial strains is mediated via two main mechanisms. First, strains can compete indirectly in an exploitative manner via shared resources, where more competitive strains reduce the abundance of less competitive strains, having a positive effect on their own frequency within the population (Hibbing *et al*., 2010; Checcucci *et al*., 2017). Genetically similar strains are thought to exert stronger competition for shared resources due to higher niche overlap of metabolic capabilities (Griffin *et al*., 2004). Such indirect interactions have mostly been studied in pairwise cultures, where the growth of the focal strain of interest is assayed under exposure to the supernatant of the other microbes (Plazinski and Rolfe, 1985; Bladergroen *et al*., 2003; Yadav *et al*., 2005; Abd-Alla *et al*., 2014; Homa *et al*., 2019). Bacteria can also interact via interference (direct) competition, whereby strains actively prevent one another's growth in cocultures (Ghoul and Mitri, 2016; Checcucci *et al*., 2017). For example, bacterial species can carry temperate phages that could potentially suppress the growth of sensitive strains (Schwinghamer and Brockwell, 1978; Burns *et al*., 2015). In addition, broad-spectrum antibiotics are well known to prevent the growth of other competing species and consequently alter community composition (Slattery *et al*., 2001; Hibbing *et al*., 2010; Tyc *et al*., 2014). Some strains produce bacteriocins, which are growth-inhibitory agents active only against closely related strains (Hirsch, 1979; Riley and Gordon, 1999; Gardner *et al*., 2004; Portella *et al*., 2009; Dobson *et al*., 2012). Similarly to broad-spectrum antibiotics, bacteriocins can strongly inhibit the growth of sensitive strains, change strain community composition, and additionally influence competition for host colonization (Schwinghamer and Brockwell, 1978; Hirsch, 1979; Triplett and Barta, 1987; Wilson *et al*., 1998; Oresnik *et al*., 1999; Riley and Gordon, 1999; Gardner *et al*., 2004; Gillor *et al*., 2009). Therefore, competition within and between species is an important determinant of microbial community functioning.

While competition is generally the most common interaction between strains (Foster and Bell, 2012), bacteria can also interact in a facilitative manner whereby interactions between organisms benefit at least one, and do not harm any, of the interacting individuals (Bruno *et al*., 2003; Elias and Banin, 2012; Zee and Bever, 2014). Facilitative interactions occurring between closely related individuals within a community include cooperative interactions that can be favoured by kin selection (Strassmann *et al*., 2011; Zee and Bever, 2014); these are often mediated by production of public goods, such as iron-scavenging siderophores, providing benefits for genetically related cooperating individuals (Lamont *et al*., 2002; Harrison *et al*., 2008; Joshi *et al*., 2008; Kramer *et al*., 2019). Public goods production is often mediated through secretion of quorum sensing (QS) chemical signalling molecules that increase in concentration in a cell density-dependent manner leading to altered gene expression and coordinated population-level behaviour (Miller and Bassler, 2001; Wisniewski-Dyé and Downie, 2002; Hmelo, 2017; Girard *et al*., 2019). Facilitative, mutualistic or unidirectional, interactions can also occur between more distantly related species within a community, and these are often mediated via syntrophic interactions where waste products from one strain can be metabolized by another strain (Rosenzweig *et al*., 1994; Bruno *et al*., 2003; Morris *et al*., 2013; Harcombe *et al*., 2018; Silva *et al*., 2019; Koch *et al*., 2020). Facilitative interactions can be maintained in communities due to spatial structure of the environment, which increases the likelihood that genetically similar individuals, or mutually dependent species, are located close to each other (Zee and Bever, 2014; Dolinšek *et al*., 2016; Nadell *et al*., 2016). Such spatial structuring has also been shown to be important for species coexistence and early colonization of host-associated environments (Larkin and Martiny, 2017; Speare *et al*., 2018). While interspecific interactions tend to be more often competitive than facilitative (Foster and Bell, 2012), such patterns are relatively unexplored between genotypes of the same species.

To study the role of genetic diversity for competitive and facilitative interactions, we focused on *Rhizobium leguminosarum*, which is a bacterial symbiont capable of providing accessible nitrogen to legumes in return for carbon (Lupwayi *et al*., 2006; Mishra *et al*., 2013). Previous work has shown that natural rhizobium populations in the soil are often very diverse (Kumar *et al*., 2015) and strains previously categorized as *R. leguminosarum* form a species complex of at least 18 genospecies that are separated by average nucleotide values below 96% (Kumar *et al*., 2015; Cavassim *et al*., 2019; Young

et al., 2021). Up to five of these genospecies have been documented to live in sympatry in rhizosphere (Kumar *et al*., 2015; Boivin *et al*., 2020; Cavassim *et al*., 2020), and they include host-specific symbiovars that nodulate clover (*Rhizobium leguminosarum* symbiovar *trifolii*; *Rlt*) and vetch (*Rhizobium leguminosarum* symbiovar *viciae*). Multiple factors including plant species and varieties (Miranda-Sánchez *et al*., 2016; Vuong *et al*., 2017; Clúa *et al*., 2018), soil abiotic properties (Rice *et al*., 1977; Harrison *et al*., 1989; Xiong *et al*., 2017; Liu *et al*., 2019), agricultural management practices (Kiers *et al*., 2002; Shu *et al*., 2012; Weese *et al*., 2015), and microbial competition (Schwinghamer and Brockwell, 1978; Wilson *et al*., 1998; Hibbing *et al*., 2010) could drive and maintain rhizobial diversity, providing different ecological opportunities. Here, we studied variation in facilitative and competitive interactions between 24 *Rlt* strains belonging to three distinct genospecies (gsA, gsC and gsE) (Kumar *et al*., 2015; Cavassim *et al*., 2020) isolated from organic (all genospecies) and conventional farms (gsC only), using *in vitro* model systems. To achieve this, indirect facilitative and inhibitory interactions were characterized by comparing the growth of strains in their own supernatant compared to their growth in a different strain's supernatant. In addition, inhibition zones produced in direct contact on soft agar plates were quantified as evidence of direct growth repression by specific strains. Our main aim was to explore if the observed genetic variation at the strain (genotype) and genospecies levels is associated with competitive and facilitative interactions and identifying underlying differences in metabolic capacity and genes, which could potentially explain this variation.

Results

Indirect facilitative and inhibitory interactions were observed at both strain and genospecies levels

To assess pairwise indirect interactions between *Rlt* strains, supernatant assays were undertaken in 576 pairwise combinations. Strains were chosen from three genetically distinct genospecies (gsA, gsC, gsE) isolated from organic farms (referred to as OA, OC and OE respectively), and gsC from conventionally managed plots (CC) (Fig. 1A and B). Strain growth in the supernatants of other *Rlt* strains was compared to growth when the strain was grown in its own supernatant, resulting in a relative growth index (RGI) (Fig. 1C). RGI measurements were derived from 62 h growth time point near strains' carrying capacity. This index correlated with strains' maximum growth rate ($R = 0.9$) and area under the curve ($R = 0.89$) metrics, being a suitable proxy for bacterial growth and potential competitive ability (Fig. S1). While both facilitative and inhibitory interactions were identified (Fig. 2A), interactions were predominantly neutral (i.e., an RGI of 1; mean RGI: 1.006; Fig. S2A). However, some strain inoculant and supernatant combinations showed extreme facilitative or inhibitory interactions: in 55 combinations (9.55%) strains grew worse (RGI < 0.75) and in another 55 combinations strains grew better (RGI > 1.25) in other strains' supernatants, indicative of inhibitory and facilitative interactions respectively.

Inoculant and supernatant interactions were influenced by individual strain differences (inoculant strain effect: $X^2_{19,18} = 354.82$, $p < 0.001$; supernatant strain effect: $X^2_{19,18} = 246.01$, $p < 0.001$). For example, strain SM168A (OE) showed a strong strain identity effect when acting as the inoculant and supernatant. The strain grew better on average in other supernatant treatments than in its own [Fig. 3; $RGI_{inoculant} = 1.381$, 95% confidence interval (CI) = 1.316–1.446], and its supernatant highly suppressed the growth of other strains on average (Fig. 3; $RGI_{supernatant} = 0.883$, 95% CI = 0.858–0.908). Conversely, strain SM154C (OA) grew significantly worse in other supernatant treatments on average (Fig. 3; $RGI_{inoculant} = 0.612$, 95% CI = 0.580–0.644), and its supernatant facilitated growth of other strains (Fig. 3; $RGI_{supernatant} = 1.269$, 95% CI = 1.223–1.315). To confirm these individual strain effects were not the result of differences in strain growth rate causing a quicker depletion of resources, we calculated the average maximum growth rates for SM168 and SM154C strains growth in the control 100% TY treatment. We found that SM154C had a faster growth rate than SM168 (SM154C: 0.0055 OD_{600nm}/h, SM168A: 0.0049 OD_{600nm}/h), and consequently, this pattern is unlikely a result of resources depletion differences between the strains. In addition, we compared the relationship between strains' inhibitory and facilitative capacity. Overall, a negative correlation was observed between the supportiveness of an *Rlt* strain's supernatant and its growth as an inoculant (Fig. 3; linear regression: $coeff_{RGIsup} = -0.8387$, $p < 0.001$), indicative of a negative relationship between the capability to be facilitated by other strains and strains' inhibitory capacity.

Interactions were significantly influenced by genospecies (inoculant*supernatant genospecies effect: $X^2_{19,9} = 102.5$, $p < 0.001$; Fig. S3A; Table S1). OE strains were the most facilitated in the supernatants of other strains (Fig. S3A), and on average their supernatants consistently suppressed OA strains' growth (compared to OA inoculants in OA supernatants as the model intercept reference level; $coeff_{supOE}$: estimate = −0.382, $t = -6.791$, $p < 0.001$). Parametric bootstrapping of 95% confidence intervals further confirmed that the growth inhibition of OA inoculants in OE supernatants was a reliable effect (Table S10). In addition, OE inoculants displayed facilitated growth in OA supernatants, and this

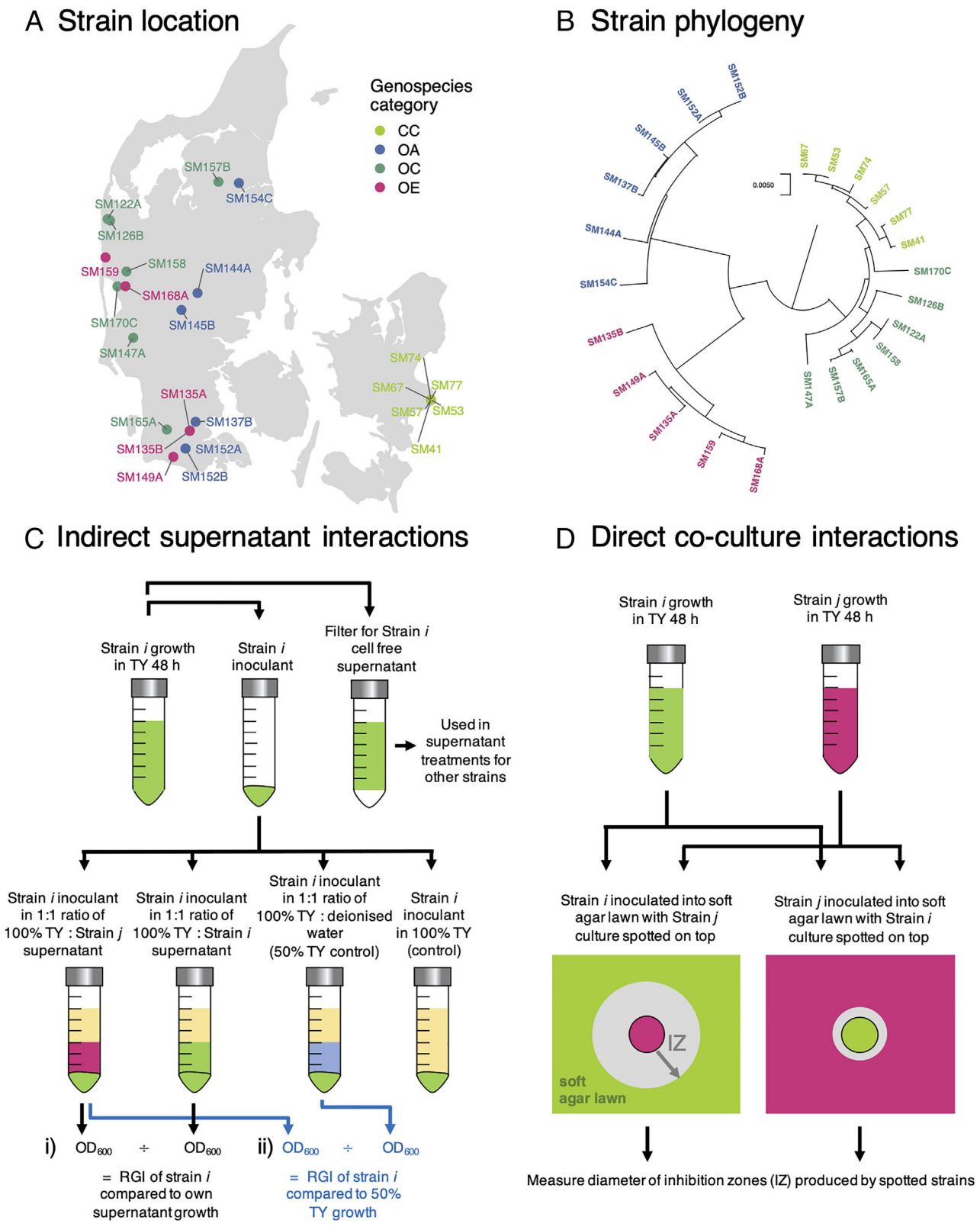

Fig 1. Legend on next page.

effect was also confirmed to be a reliable effect ($coeff_{inocOE}$: estimate = 0.209, t = 3.08, p < 0.01; Table S10). To ensure genospecies effects were not influenced by strongly interacting individual strains, SM168A and SM154C were omitted, and the model was re-calculated. Despite strain exclusion, interactions remained similar and statistically significant, indicative of true genospecies-level interactions (Table S2; Table S10). It was also found that OA strains grew to lower densities in OE supernatants than in 50% TY control treatments (Fig. S5). This suggests that OA growth inhibition was not purely due to nutrient resource depletion but also associated with inhibition through secretions. The effect of farming practice origin (conventional and organic isolates) on strain interactions was also tested with genospecies C strains as previous studies have identified that diversity and metabolic capability of bacterial populations can differ depending on agricultural management practice (Kennedy and Smith, 1995; Palmer and Young, 2000; Depret *et al*., 2004; Schmidt *et al*., 2019). No significant effects were found (Table S3; Table S4; Fig. S3A) and gsC strains isolated from both organic (OC) and conventional (CC) farms grew to similar densities on average in the supernatants of other strains compared to their own (Fig. S4B; Table S3).

Furthermore, the type of interaction between strains (facilitative, inhibitory, or neutral) was significantly associated with the genetic relatedness between strains, and the specific genospecies pairs ($F_{536,543}$ = 5.878, p < 0.001; Table S5; Fig. S2B). For example, when OA inoculant strains interacted with supernatants from different genospecies groups (OC, OE or CC), they showed a greater growth inhibition with increasing genetic dissimilarity (OA – other gs group: slope = 293.6, p < 0.0001; Table S6). OC inoculant strains also showed a similar association when interacting with supernatants from other genospecies, displaying greater growth inhibition with increased genetic dissimilarity, whereas CC inoculant strains showed no association with genetic relatedness (OC – other gs group: slope = 1.443, p < 0.01; CC – other gs group: slope = −0.718, p > 0.05; Table S6). The effect of OE inoculants in other genospecies' supernatants was not significantly associated with the genetic relatedness (OE – other gs group: slope = −63.1, p > 0.05; Table S6), and this was likely because OE displayed the relatively strongest inhibitory ability across all tested strains. Furthermore, these trends were maintained even when SM168A and SM154C were removed (Table S6). In contrast to genospecies level interactions, interactions within genospecies were not strongly associated with the genetic relatedness (OA–OA: slope = 0.7, p > 0.05; OC–OC: slope = −1.8, p > 0.05; OE-OE: slope = −0.8, p > 0.05; CC–CC: slope = 0.4, p > 0.05; Table S6; Fig. S2B). Together, our results demonstrate that both facilitative and inhibitory indirect interactions were common at both strain and genospecies level (specifically OA and OE). In addition, the effects of relatedness were more important for competitive interactions at the genospecies level.

Direct inhibitory interactions were observed at both strain and genospecies levels

To assess whether strain interactions could be detected in direct interference competition, strains were co-cultured by growing one strain as a soft agar lawn, and by spotting liquid cultures of other strains on top and observing whether spotted strains induced zones of inhibition (Fig. 1D). Inhibition zones of varying sizes were observed in 92 out of 576 possible strain combinations (15.97%) after 72 h of growth, while most of the interactions were neutral (Fig. 2B).

Similar to the supernatant growth assays, significant interactions were observed at both strain and genospecies levels. Individual strain differences were observed for both the spotted and the lawn strains (spot strain effect: $X^2_{19,18}$ = 9.59, p < 0.01; lawn strain effect: $X^2_{19,18}$ = 759.68, p < 0.001). The OA strain SM145B, SM154C and SM144A were the only strains susceptible to inhibition eliciting inhibition zones of more than 5 mm by other strains (Fig. 2B). On the other hand, one OA strain (SM137B) was resistant to inhibition by all other strains. Only one OE strain (SM149) was susceptible to inhibition by other strains (Fig. 2B). Furthermore, a single CC strain (SM53) showed susceptibility to all OE strains (Fig. 2B). We additionally identified that stains that were more genetically similar (ANI) were more likely to produce similar inhibition zone profiles on each soft agar strain lawn (Fig. S6; linear regression: R^2 = 0.1272,

Fig 1. Strain collection (A) metadata, (B) phylogenetic relatedness, and experimental designs for the (C) indirect interaction supernatant assay and (D) direct interaction spot plating assay. *Rhizobium leguminosarum* symbiovar *trifolii* (*Rlt*) strains were isolated from white clover (*Trifolium repens*) root nodules collected from agricultural managements across Denmark. Isolates were categorized by genospecies classification and environmental treatment (organic management genospecies A = OA; organic management genospecies C = OC, organic management genospecies E = OE, and conventional management genospecies C = CC). A phylogeny of strains was based on the concatenation of 305 core genes using the neighbour-joining method from Cavassim *et al*. (2020). (i) calculates the relative growth index (RGI) of strain *i* in the strain *j* supernatant compared to growth in its own supernatant, and (ii) calculates the RGI of strain *i* in the strain *j* supernatant compared to growth in the 50% Tryptone Yeast broth control treatment.

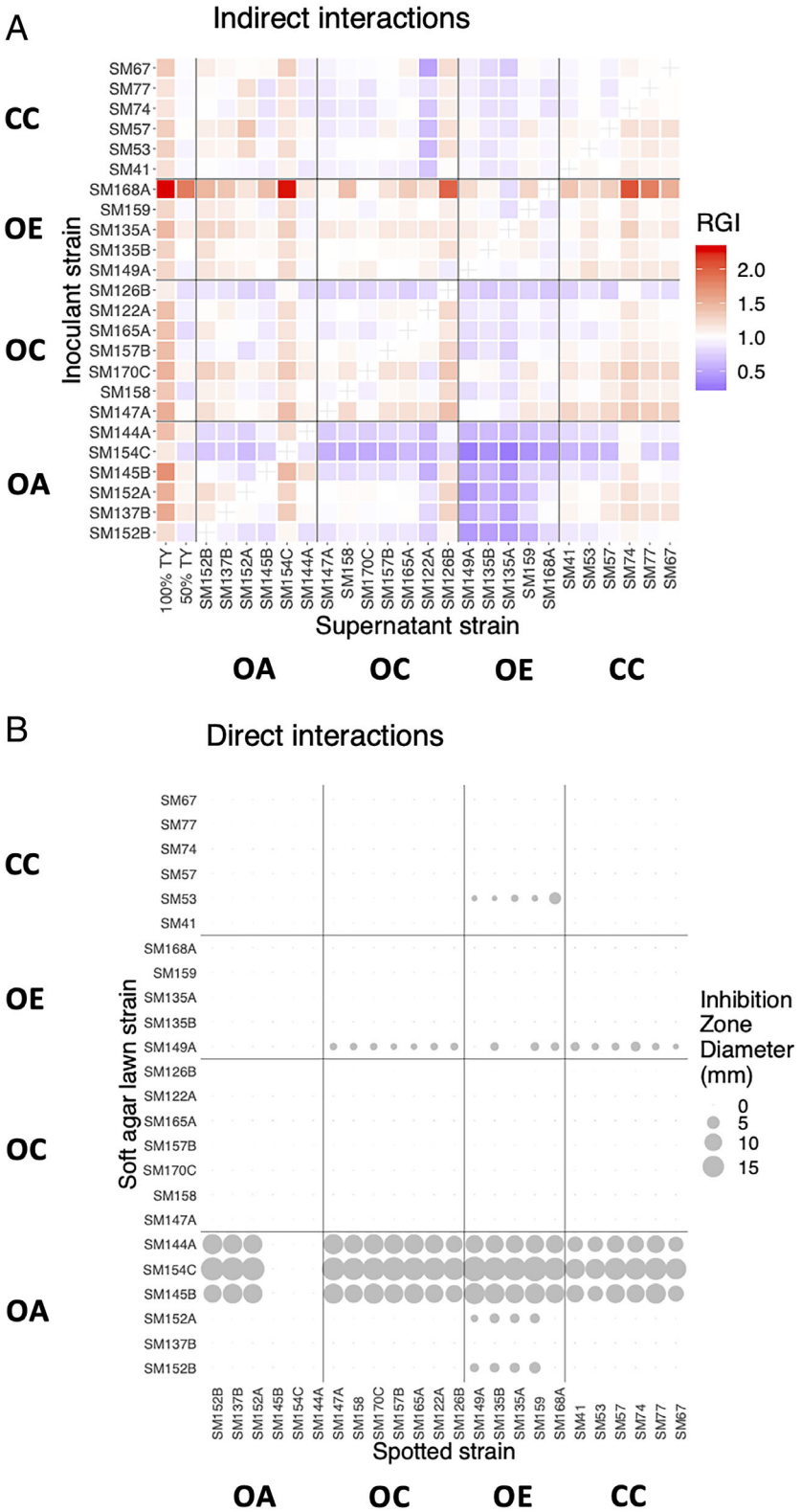

Fig 2. Evaluation of indirect and direct pairwise interactions of 24 *Rlt* strains. **(**A) All strains were grown in Tryptone Yeast broth depleted by other strains (supernatant). Indirect strain interactions were measured by calculating the relative growth indices (RGIs) of strains inoculated into each other's supernatants ($n = 5$). High RGI values denote for high inoculant strain growth. (B) The mean diameter of inhibition zones (mm) produced by strains spotted onto soft agar lawns of other strains. The size of the circles indicates the diameter of the inhibition zones.

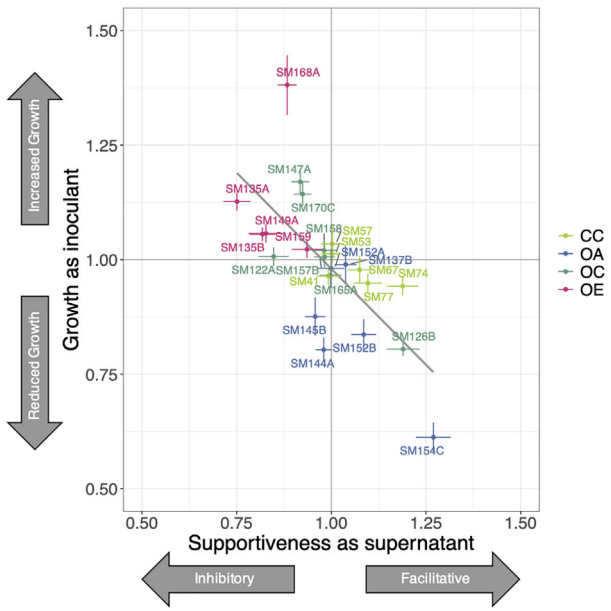

Fig 3. The supportiveness of *Rlt* strain supernatants for growth of other strains correlate negatively with their own growth as inoculants. Growth of inoculant was calculated by averaging the RGI of strain *i* when grown in all other supernatant treatments (excluding control Tryptone Yeast broth treatments). Supportiveness of supernatant was calculated by averaging the RGI of all other strains grown in the supernatant of strain *i*. Strains are grouped by genospecies and environmental treatment. Grey line displays the regression line fit by linear model, and error bars display 95% confidence intervals. RGI calculation is provided in the methods.

$\text{coeff}_{ANI} = -74.555$, $p < 0.001$). Overall, strains predominantly displayed stronger strain identity effects in direct interactions compared to indirect interactions. However, strains with the strongest identity effects were all from the same genospecies (OA).

At the genospecies level, significant interactions were also observed (spot * lawn genospecies effect: $X^2_{19,9} = 95.933$, $p < 0.001$; Fig. S3B; Table S7). OA strains were the most susceptible to inhibition by other strains (Fig. S3B), while the OE strains were the most inhibitory on OA agar lawns (compared to OA spots on OA lawns as the model reference level; coeff_{spotOE}: estimate = 4.203, $t = 9.80$, $p < 0.001$; Table S10). Similarly, OE strains were able to inhibit at least one strain in each genospecies group, except for OC strains, which were resistant to inhibition by all other strains (Fig. 3B; Fig. S4C and D). OA strains' susceptibility was largely driven by the three highly susceptible strains (SM145B, SM154C and SM144A). Similar to the supernatant assays, farming practice had no effect on strain susceptibility or inhibitory capacity within genospecies C strains (-Table S8; Table S9; Fig. S3B). One exception to this was that OC strains were able to produce on average significantly slightly larger inhibition zones than CC strains on OA soft agar lawns (Table S8). Moreover, the size of the inhibition zone was positively correlated with the strains' RGI in the supernatant assays, indicative of a positive relationship between direct inhibition and ability to be facilitated by the supernatant of others (Fig. S7). Moreover, the slope of this association was stronger when the strains were more inhibitory on average (inhibition zones with a diameter more than 5 mm (Fig. S7; linear regression: $R^2 = 0.344$, $\text{coeff}_{RGI} = 8.541$, $p < 0.001$) compared to less than 5 mm (Fig. S7; linear regression: $R^2 = 0.1919$, $\text{coeff}_{RGI} = 2.751$, $p < 0.05$). Together, these results show that direct pairwise interactions were qualitatively similar to indirect supernatant assays and mainly mediated by relatively strong interactions between the relatively most susceptible OA and the most inhibitory OE genospecies.

The most inhibitory OE strains displayed the greatest metabolic capacity compared to other genospecies

To compare the overall metabolic capacity of different genospecies the growth of all strains was determined in 31 single substrates important for the growth in microbial communities using Ecoplates (Biolog Hayward, CA, USA, Fig. S8). On average, OE strains could metabolize a significantly higher number of single substrates than OA strains (Kruskal–Wallis $X^2 = 7.940$, df = 3, p-value <0.05; Dunn's post hoc $p < 0.05$; no significant difference with OC or CC strains). Furthermore, OE strains displayed a higher metabolic capacity than CC strains in terms of higher Average Well Colour Development (AWCD; Fig. 4A; Kruskal–Wallis: $X^2 = 9.3349$, df = 3, $p < 0.05$; Dunn's post hoc: adjusted $p < 0.05$). No significant difference was observed between the overall metabolic capacity of gsC strains isolated from organic (OC) or conventional (CC) sites. The overall metabolic capacity correlated positively with strains' ability to grow on other strains' supernatant, suggesting the potential for cross-feeding (Fig. 4B; PGLS regression: $\text{coeff}_{RGI} = 0.162538$, $t = 3.0574$, $p < 0.01$, Lambda ML = 0.161). We further analysed the metabolic capacities of all strains using principal components analysis (PCA) by averaging single-substrate data over six resource type groups: amines, amino-acids, carbohydrates, carboxylic acids, complex carbon sources and phosphate carbon sources (Table S11). The first two principal components explained 76% of the total variance. While most of the genospecies groups did not differ from each other, OE strains were found to significantly differ from CC strains (Fig. 4C; Table S12; PERMANOVA: $F_{3,22} = 3.8293$, $p < 0.01$; post hoc adjusted $p < 0.05$). The separation of OE strains along the PC1 was associated with overall higher amino acid, carbohydrate and phosphate carbon metabolism (Fig. 4D), and

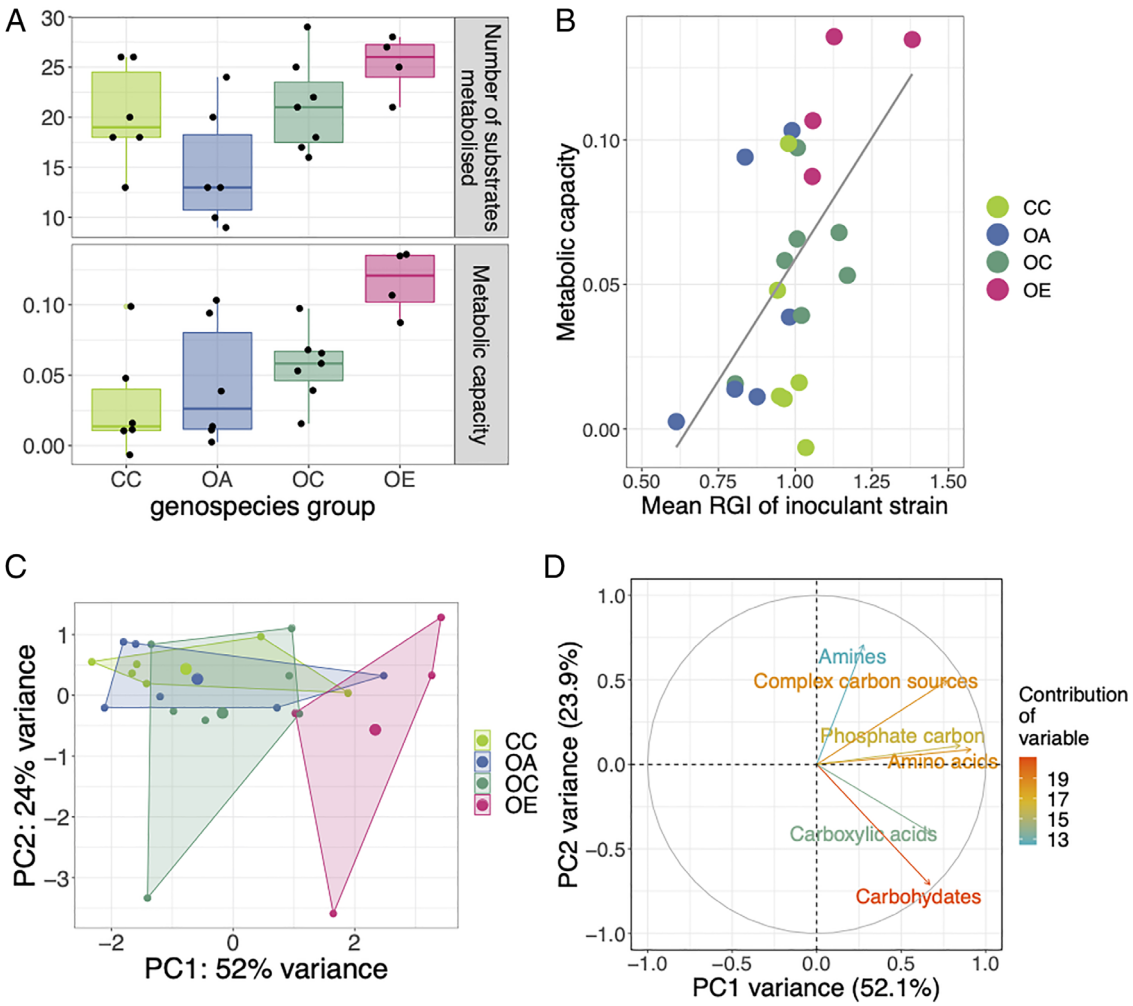

Fig 4. Metabolic differences between *Rlt* genospecies groups. (A) Number of substrates metabolized, and metabolic capacity (average well colour development) was calculated across 31 single substrate growth treatments for 23 *Rlt* strains grouped by their genospecies and environmental treatment. (B) The average RGI of an inoculant strain in all other supernatant treatments correlated with the strain's ability to metabolize the 31 carbon substrates (correlation coefficient $R = 0.55$, $p > 0.01$, regression line shows linear model fit). (C) Principal components analysis for metabolic capacity of 31 single substrate treatments averaged across six resource type groups. Points show individual strains coloured by genospecies group. (D) Contribution of different metabolite groups explaining the observed strain-level variance of the first two principal components. Resource type groups are coloured by their percentage contribution of the total variance for principal components 1 and 2. Individual substrates within each resource type group can be found in Table S11.

glycyl-L-glutamic acid (amino acid group) and Tween 40 (complex carbon sources) contributed most to the variance of PC1 and PC2 respectively (Fig. S9). Interestingly, metabolic similarity (calculated by Euclidean distance of 31 single substrate treatment OD values) did not strongly correlate with the genetic similarity of strains measured by average nucleotide identity (Mantel R statistic $= -0.275$, $p > 0.05$), which suggests that relatedness was not linked with strains' ability to compete for these resources. Together, these results suggest that the most inhibitory genospecies OE also had the highest metabolic capacity.

Using comparative genomics to identify potential genetic mechanisms underlying interactions between strains and genospecies

To explore the underlying mechanisms behind indirect and direct interactions, the presence of genes associated with secondary metabolite clusters, known rhizobial bacteriocins and prophages were identified using comparative genomics (see Experimental Procedures). In addition, the presence of genes linked to quorum-sensing pathways was analysed, as QS systems can be linked to both facilitative and inhibitory interactions between strains.

Several secondary metabolite biosynthesis gene clusters were identified in some or all strains using antiSMASH (Fig. 5A) (Blin *et al*., 2019), including Type III polyketide synthases (T3PKS), bacteriocin, terpene, arylopolyene, ectoine, homoserine lactones (hserlactones), proteusin, pheganomycin-style protein ligase-containing cluster (fused) and non-ribosomal peptide synthetase (NRPS) clusters (Table S13). Clusters associated with homoserine lactones were the most abundant secondary metabolite biosynthesis gene clusters in all strains. Bacteriocin and proteusin gene clusters were not found in CC strains but were present in all other genospecies groups. In addition, all OA strains lacked NRPS clusters, whereas at least one NRPS cluster could be identified in OE and CC strains, and clusters appeared sporadically in OC strains. Only all OE strains contained an NRPS gene cluster with 100% antiSMASH similarity to the gene cluster encoding the synthesis of the siderophore, Vicibactin (Fig. 5A), which is used to sequester iron in the rhizosphere (Heemstra *et al*., 2009; Wright *et al*., 2013). However, by BLASTn search, *fhuBCD* genes encoding Vicibactin siderophore uptake proteins (Stevens *et al*., 1999) were found in all strains (Fig. 5A), with OE strains showing the highest percentage identities (Supplementary File 2). We further tested whether Vicibactin gene cluster presence was associated

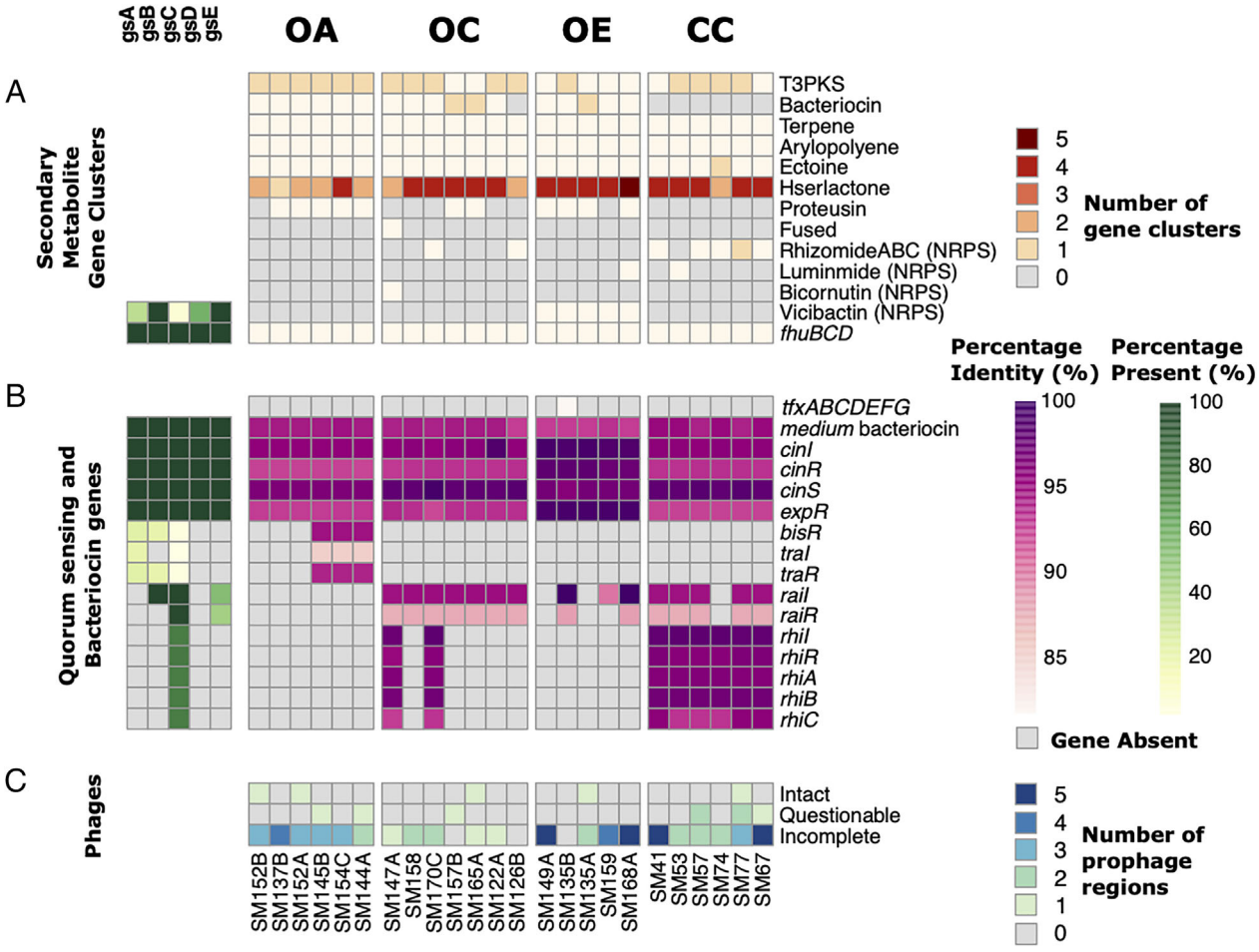

Fig 5. Percentage identity of *Rhizobium leguminosarum* bacteriocins and quorum sensing associated genes, and the number of secondary metabolite gene clusters and phages, found in *Rlt* strains. (A) Secondary metabolite gene cluster colours correspond to the number of clusters identified for each type of gene cluster. Secondary metabolite clusters identified to have 100% similarity to known clusters were reported as individually in the table (e.g., Vicibactin). *fhuBCD* genes were identified using BLASTn and reported as present if all three genes were identified with >90% percentage identity and matching the full length of the query sequences. In addition, from a dataset of 196 *Rlt* strains (Cavassim *et al*., 2020), the percentage of strains in each genospecies containing Vicibactin synthase (*vbsS*) and *fhuBCD* (percentage present) is displayed on the left, with colours corresponding to increasing percentage of strains with the gene in each genospecies. (B) Quorum sensing and bacteriocin genes colours correspond to increasing percentage identity of identified genomic regions to reference sequences. Grey boxes highlight genes that were not present in a specific genome. In addition, the incidence of these genes in 196 *Rlt* strains is shown on the left. Accession numbers for quorum sensing associated gene reference sequences can be found in Table S14. (C) Phage colours correspond to the number of prophage regions of intact, questionable or incomplete quality identified in each strain genome. Strains are grouped according to their genospecies and environmental treatment (OA, OC, OE, CC).

with strains' direct inhibitory interactions (Fig. 2B); however, we identified no significant association between Vicibactin gene presence and direct inhibitory ability (-Table S16), which was likely due to the differences between genospecies groups being confounded by phylogenetic signal correction in the regression model. Several gsC strains also contained an NRPS with 100% antiSMASH similarity to the rhizomideABC synthase cluster in *Paraburkholderia* spp., which encodes lipopeptides suggested to have weak inhibitory activity against other bacterial species such as *Staphylococcus aureus* and *Bacillus subtilis* (Wang *et al*., 2018).

The presence of known bacteriocins was screened using BLASTn. The *medium* bacteriocin was identified in all 24 strains with a percentage sequence identity varying between 92.76% and 95.17% to the *medium* bacteriocin reference sequence (Table S14; Table S15), but all hits matched only to the last 10% of the query sequence, suggesting it is unlikely these strains have active medium bacteriocins. OA strains were found to have a slightly higher percentage sequence similarity to the reference sequence compared to OE strains (Fig. 5B; Table S15). In addition, bacteriocin trifolitoxin genes (*tfx*) were only found in OE strain SM135B. However, only the putative immunity protein *tfxG* was identified, which suggests that SM135B could be immune to the effects of trifolitoxin but unlikely to be able to produce the toxin.

Given that antiSMASH identified many unknown homoserine lactone clusters, we separately searched for known *R. leguminosarum* quorum-sensing *N*-acyl homoserine lactone (AHL) synthase genes (*cinI*, *raiI*, *rhiI*, and *traI*) and their related regulatory genes (*cinR*, *raiR*, *rhiR*, and *traR*) using BLASTn (Fig. 5B; Table S14; Table S15). The *cinI* and *cinR* genes, which encode the *small* bacteriocin AHL (3OH-$C_{14:1}$-HSL) synthase and its transcriptional regulator, were found in all 24 strains with high percentage identity to the reference sequence (Fig. 5B). The *raiI*/*raiR* quorum-sensing system was found to be the second most common system and present in 58% of strains, but absent from the OA strains. Genes *expR* and *cinS*, required for *raiR* expression, were present in all 24 strains (Fig. 5B). The *rhiI*/*rhiR* system, and the rhizosphere-expressed *rhiABC* operon (of undetermined function) that it regulates, was found exclusively in genospecies C strains (Cubo *et al*., 1992; Rodelas *et al*., 1999). Furthermore, three OA strains (SM144A, SM154C and SM145B), that showed high susceptibility to direct inhibition (Fig. 2B), harboured *traI*, *traR and bisR* genes which were only found together in these strains (Fig. 5B). We additionally found that the presence of the *tra* quorum-sensing system in OA strains was significantly correlated with the inhibition zone sizes produced from direct inhibitory strain interactions on OA soft agar lawns (Table S16). Together, our findings suggest that the presence of different quorum-sensing signals potentially influenced the variation in observed competitive strain interactions.

An intact prophage region was detected in two OA strains, along with one strain each from the OC, OE and CC groups (Fig. 5C). In addition, multiple prophage regions were identified as questionable or incomplete within each genospecies group, with CC strain genomes containing the largest number of questionable and incomplete prophage regions in total (Table S17). However, the overall number of prophage regions did not correlate with strains' ability to suppress other strains (Table S17), and hence, was unlikely to be linked to observed competitive interactions.

To evaluate whether gene distributions in the 24 strains were representative of their wider genospecies diversity, we further used BLASTn to search for quorum-sensing-, bacteriocin- and Vicibactin-associated genes in 196 *Rlt* strains from Cavassim *et al*. (2020) (which include the 24 strains used in this study), as these genes displayed a degree of genospecies specificity within our 24 strains and suitable reference sequences could be identified to carry out the analysis (Supplementary File 2). Concurring with the analysis of the 24 strains, evaluation of 196 strains identified a separation in the distribution of quorum-sensing genes between genospecies. The search for *raiI*/*raiR* genes across the 196 strains confirmed that, as in the 24-strain dataset, these genes were found in the majority of gsC strains and approximately half of gsE strains, and were absent from gsA, gsB and gsD (Fig. 5B; Supplementary File 2). Similarly, analysis of the 196 strains confirmed that the *rhiI*/*rhiR*/*rhiABC* genes were only present in gsC strains (Fig. 5B; Supplementary File 2). While only three OA strains carried the *bisR*/*traI*/*traR* genes in the subset of 24 strains, across the 196 strains these genes were present in a small proportion of gsA, gsB and gsC strains but absent or incomplete from gsD and gsE strains (Fig. 5B; Supplementary File 2). We also assessed the distribution of the secondary metabolite gene, *vbsS*, which encodes the Vicibactin siderophore NRPS. All gsB and gsE strains were found to contain *vbsS*, whereas less than 10% of gsC strains, and less than 40% of gsA strains, contained this gene (Fig. 5A; Supplementary File 2). Furthermore, siderophore uptake genes, *fhuBCD* (Stevens *et al*., 1999), were found in all strains, with gsE strains displaying the relatively highest mean percentage identity, and gsA showing one of the lowest mean percentage identities (Fig. 5A; Supplementary File 2). Overall, we identified a statistically significant, non-random distribution of quorum-sensing and secondary metabolite genes across *R. leguminosarum* genospecies (Table S18), indicative of a divergence of genes linked to facilitative and competitive interactions.

Discussion

Here, we investigated the indirect and direct interactions of *Rlt* strains. While most of the interactions were neutral, highly significant facilitative and competitive interactions were found at both strain and genospecies levels. Especially clear indirect and direct interactions were observed between the highly inhibitory OE and the most susceptible OA genospecies. Overall, strains that produced more-inhibitory supernatants tended to grow better in other strain supernatants, while strains that had particularly facilitative supernatants, grew worse in the supernatants of the other strains. Based on correlation analyses, we found that interactions were potentially associated with differences in strains' metabolic capacity and the presence of multiple different homoserine lactone quorum-sensing biosynthesis gene clusters. In addition, a non-random distribution of different quorum-sensing systems across genospecies was identified, while OE strains harboured Vicibactin siderophore synthesis gene with the highest sequence similarity with the reference gene, which could have provided an advantage in iron competition. This work highlights the importance of genetic diversity and relatedness for explaining phenotypic differences in facilitative and competitive interactions between coexisting *R. leguminosarum* genotypes and genospecies. Such variation could affect both the composition and functioning of rhizosphere microbial communities and potentially the likelihood of symbiotic establishment by the relatively more frequent genotypes (Schwinghamer and Brockwell, 1978; Oresnik *et al*., 1999; Denison and Kiers, 2004; Blanco *et al*., 2010).

Indirect inhibitory and facilitative supernatant interactions were significantly influenced by genospecies effects and predominantly driven by the inhibition of OA strains by OE strains. This is in line with previous studies that have associated *Rhizobium* signalling molecules found in cell-free supernatant with growth inhibition and potential for symbiotic establishment (Bladergroen *et al*., 2003; Sanchez-Contreras *et al*., 2007; Checcucci *et al*., 2017). However, the significance of these interactions has not previously been investigated at the genospecies level. While previous studies have found that bacterial population diversities can differ depending on agricultural management (Palmer and Young, 2000; Depret *et al*., 2004; Schmidt *et al*., 2019), no significant effects on indirect competitive interactions were found between genospecies C isolates from organic or conventional management practices. Furthermore, we are unable to be certain that any genetic differences between the organic and conventional isolates are necessarily due to the differences in farming management regimes due to the multitude of confounding variables associated with environmental origins. Overall, the strength of the competitive and facilitative interactions varied greatly across all pairwise strain combinations and depended on interactions between specific strains (Fig. 2).

Interaction differences between genospecies observed in this study are interesting, as genospecies-exclusive phenotypes have not been previously reported (Kumar *et al*., 2015). One explanation for these patterns could be genetic relatedness, which could affect the strength and balance of competitive and facilitative interactions. In support for this, we found that indirect competitive interactions between genospecies were more strongly associated with genetic relatedness compared to within-genospecies strain interactions (Fig. S2B; Tables S5–S6). Previous research has suggested that genetically related bacterial strains will exert relatively stronger competition towards each other through resource competition, because they are more likely to share similar metabolic pathways, exhibiting overlap in their resource niche (Griffin *et al*., 2004; Foster and Bell, 2012; Russel *et al*., 2017). However, we found no correlation between metabolic similarity and genetic similarity of core genes, and despite some genospecies level patterns, even closely related individual genotypes varied considerably in how they interacted (Fig. 2A). Other studies have also shown no association between genetic relatedness and competitiveness of bacterial interactions (Foster and Bell, 2012). Furthermore, it is possible that competitive interactions were largely driven by variation in accessory genome content, that varies extensively between and within genospecies and often includes specialized adaptive functions like metabolism-related genes (Crossman *et al*., 2008; Mazur *et al*., 2013; Ormeño-Orrillo and Martínez-Romero, 2013; Kumar *et al*., 2015; Cavassim *et al*., 2020).

While a large variation in metabolic capacity was identified across strains, no substrate was found to be exclusively metabolized by a single genospecies (Fig. S8), which supports previous findings (Kumar *et al*., 2015). Strains showed the greatest phenotypic variation in amino acid metabolism and general carbohydrate metabolism, such as sugars, as reported previously (Wielbo *et al*., 2010), but also in phosphate carbon sources, which has not been observed before (Fig. 4C and D). OE strains were able to metabolize the greatest number of single substrates and displayed the highest metabolic capacity across 31 single substrates measured. This result supports the hypothesis that OE strains can efficiently deplete nutrients in the complex TY medium, acting as generalists and subsequently produce supernatants that are poor at supporting the growth of other strains. Correspondingly, OA strains showed lower metabolic capacity and metabolized fewer substrates in general. As a result, OA strains likely left more nutrients

unutilized in the supernatant, which could explain the relatively high facilitation of other strains. These results could be explained by resource specialization for OA strains, which has been observed at both intra- and interspecies levels in bacterial communities (Evans *et al*., 2020; Nuccio *et al*., 2020; Schreiber and Ackermann, 2020). High metabolic versatility is likely to be advantageous for bacterial survival in the soil, helping the strains to metabolize more diverse plant-secreted compounds and nutrients in the bulk and rhizosphere soils (Wielbo *et al*., 2010; Mazur *et al*., 2013). In the case of rhizobia, cross-feeding could also improve strains' chances to establish symbiosis with legumes, if it increases their survival and relative abundance in diverse microbial communities (Bruno *et al*., 2003; Silva *et al*., 2019).

While clear direct inhibition was observed at both genospecies and genotype levels, it was not clearly linked with genetic similarity. OE strains were the most capable of producing inhibition zones and could inhibit all but one of the OA strains. Genospecies C strains seemed to be the most resistant to inhibition overall. Furthermore, the strength of inhibition was found to vary at the level of individual strains, and we observed two discrete inhibition zone sizes, suggesting two potentially distinct mechanisms of inhibition (Fig. 2B; Fig. S7). Direct inhibitory rhizobial interactions have been observed previously, driven by the production of bacteriocins and activity of quorum-sensing associated mechanisms (Hirsch, 1979; Joseph *et al*., 1983; Schripsema *et al*., 1996; Oresnik *et al*., 1999; Lithgow *et al*., 2000; Wilkinson *et al*., 2002; McAnulla *et al*., 2007; Joshi *et al*., 2008). We suggest that the large inhibition zone sizes are caused by AHLs, and the smaller inhibition zones likely produced by larger bacteriocins or other higher molecular weight molecules. For example, *R. leguminosarum* strains have been shown to produce three types of bacteriocins called *small*, *medium* and *large* due to their molecular weights and diffusion properties (Hirsch, 1979; Sanchez-Contreras *et al*., 2007). *Small* was later found to be a quorum-sensing AHL linked with co-transcriptional growth inhibition, and *medium* was found to be an RTX-like protein (Hirsch, 1979; Schripsema *et al*., 1996; Oresnik *et al*., 1999; Lithgow *et al*., 2000). The greater susceptibility of three OA strains (SM144A, SM154C, SM145B) to inhibition by other strains, and their larger inhibition zone diameters (Fig. 2B), could be explained by the presence of quorum-sensing *traI*, *traR* and *bisR* genes (Fig. 5B), which confer greater sensitivity to 3OH-$C_{14:1}$-HSL (*small* bacteriocin), leading to growth inhibition (Hirsch, 1979; van Brussel *et al*., 1985; Gray *et al*., 1996; Schripsema *et al*., 1996; Wilkinson *et al*., 2002; Wisniewski-Dyé and Downie, 2002; McAnulla *et al*., 2007). The sensitive OA strains do not inhibit each other, despite having the *cin* genes, and this is consistent with the *cin* genes being downregulated in strains containing the *bisR*/*traI*/*traR* pathway in *R. leguminosarum* (McAnulla *et al*., 2007).

Furthermore, *Rhizobium* strains produce the greatest diversity of QS molecules known among soil bacteria, suggesting a wide range of potential interactions between strains (Cha *et al*., 1998; Wisniewski-Dyé and Downie, 2002). All the 24 strains used in this study contained the 'master regulator' of downstream rhizobial AHL quorum-sensing pathways: *cinI*/*R* (Lithgow *et al*., 2000). However, large variation in the presence of downstream inducer and regulator quorum sensing genes, such as *rhiI*/*rhiR*, *rail*/*raiR* and *traI*/*traR*, also likely contributed to the differences in facilitative and inhibitory interactions (Lithgow *et al*., 2000; Wisniewski-Dyé and Downie, 2002; Sanchez-Contreras *et al*., 2007). For example, both inducer and regulator genes for the *rhi* and *rai* quorum sensing pathways were found together only in *Rlt* gsC strains (CC and OC) (Fig. 5B), but have previously been identified in *R. etli* and *R. leguminosarum* symbiovar *viciae* and *phaseoli* strains (Wisniewski-Dyé *et al*., 2002; Edwards *et al*., 2009; Downie, 2010). These two pathways display some level of redundancy as they can produce some of the same AHLs and could influence nodulation efficiency and symbiotic capacity (Cubo *et al*., 1992; Rodelas *et al*., 1999; Wisniewski-Dyé *et al*., 2002; Wisniewski-Dyé and Downie, 2002). The redundancy of quorum-sensing systems in rhizobia could also potentially provide some resilience to inhibitory quorum-sensing mechanisms imposed by other rhizosphere bacteria. In addition, analysis of 196 *Rlt* strains confirmed a non-random pattern of quorum-sensing genes across different genospecies, such as the presence of *rhi* quorum-sensing genes only in gsC strains, suggesting genospecies may utilize different combinations of competitive interaction mechanisms in the rhizosphere (Supplementary File 2). Together, these findings suggest that sensitivity to direct inhibition by other strains could have been mediated by both quorum-sensing pathways and bacteriocin production (Hirsch, 1979; Wilkinson *et al*., 2002; Wisniewski-Dyé and Downie, 2002). As these pathways also regulate root adhesion, symbiosis plasmid transfer and growth in rhizospheres of nodulated legumes, rhizobial competition could also indirectly affect the symbiotic capacity of strains through pleiotropy (Wilkinson *et al*., 2002; Wisniewski-Dyé and Downie, 2002; Danino *et al*., 2003; He *et al*., 2003; McAnulla *et al*., 2007).

Comparative genomics was further used to explore the production of secondary metabolites. All five OE strains were found to contain one Vicibactin siderophore NRPS gene cluster, whereas NRPS clusters were completely absent in OA strain genomes (Table S13). The *fhuBCD* genes, encoding a Vicibactin siderophore uptake system,

were found in all strains but with a higher percentage identity in OE strains compared to OA strains which could be indicative of potential sequence differences which may influence binding-specificity towards Vicibactin (Fig. 5A; Supplementary File 2). In addition, analysis of 196 *Rlt* strains confirmed that the Vicibactin synthase gene, *vbsS*, was present in all gsB and gsE strains, and was also found in some gsA strains (Fig. 5A; Supplementary File 2). Vicibactin can be used by *R. leguminosarum* strains to sequester iron from rhizosphere environments and is associated with productive symbioses, as iron is vital for successful nitrogenase function (Heemstra *et al*., 2009; Geetha and Joshi, 2013; Wright *et al*., 2013). The sequestering of iron from the growth environment can provides siderophores with antimicrobial-like qualities, as it prevents the growth of other strains with less effective siderophores or lacking the cell-specific siderophore receptors (Kramer *et al*., 2019). Despite the resource complexity of TY broth, it is also an iron-limited medium, and therefore, Vicibactin production by OE strains could have provided a competitive advantage compared to other strains.

It will be important to validate these observed strain interactions in natural and rhizosphere soil communities in the future. While our results suggest that OE strains could be highly competitive, their frequency is relatively low in white clover nodule populations across Europe, while the gsC strains are often the most abundant (Kumar *et al*., 2015; Cavassim *et al*., 2020; Moeskjaer *et al*., 2020). Relatively higher frequencies of gsE strains have however been observed in pea and faba bean nodules (Boivin *et al*., 2020), which suggest that their abundance could also depend on the specific legume host. Overall, the observed link between rhizobial genotypes and phenotypes, could potentially be explained by local adaptation and niche specialization, enabling strains to coexist in sympatry within the competitive rhizosphere environment (Hunt *et al*., 2008; Oakley *et al*., 2010; Koeppel *et al*., 2013). In support for this idea, genotypic differences between strains have been found to influence competitive interactions, further affecting colonization, spatial distribution and functionality of microbial populations (Becker *et al*., 2012; Braga *et al*., 2016; Speare *et al*., 2018; Koch *et al*., 2020; Wang *et al*., 2021). Such hypotheses could be tested in the future by quantifying strain interactions in the presence of plants to directly test how selection by strain competition and host-mediated selection drive rhizobial evolution. Moreover, recent advancements in sequencing technologies for strain genotype identification, such as the MAUI-seq high throughput amplicon sequencing method (Fields *et al*., 2020) or strain ID tagging (Mendoza-Suárez *et al*., 2020), could be used to discriminate different genotypes in rhizosphere microbial communities. Another crucial area for future investigation would be to induce targeted mutations in the quorum-sensing pathway genes that are putatively involved in the direct interactions observed within this study, which could confirm whether the specific quorum-sensing mechanisms are influencing these direct strain interactions. In a broader perspective, our results suggest that competitive and facilitative interactions could potentially be important ecological mechanisms driving and maintaining *R. leguminosarum* diversity. In addition to species richness, intra-species diversity could thus also play an important role for the functioning of microbial communities.

Experimental procedures

Rhizobial strains

Twenty-four *Rhizobium leguminosarum* symbiovar *trifolii* (*Rlt*) strains isolated from organic farms and a conventional plant-breeding trial across Denmark were selected from the NCHAIN *Rlt* isolate collection (Cavassim *et al*., 2020). Strains were genetically characterized based on core gene phylogeny as genospecies (gs) A, C and E (Kumar *et al*., 2015) and further labelled into four categories based on environmental source: organic gsA (OA, $n = 6$), organic gsC (OC, $n = 7$), organic gsE (OE, $n = 5$) and conventional gsC (CC, $n = 6$) (Table S19). A whole-genome average nucleotide identity (ANI) value greater than 95% is commonly accepted to equate to a DNA–DNA hybridization value of 70%, and therefore, strains with <95% ANI values are often considered to belong to distinct species (Goris *et al*., 2007). We used a slightly different measure of ANI based on 441,287 shared single nucleotide polymorphisms in 6529 genes present in at least 100 strains (Cavassim *et al*., 2020). Within-genospecies ANI averaged 97.8% and ranged between 96.8%–99.9%. Between-genospecies ANI averaged 90.3% and ranged between 90.2%–90.4%. A species phylogeny of the 24 strains was generated from Cavassim *et al*. (2020)s 196-strain phylogeny, which was based on the concatenation of 305 core genes using the neighbour-joining method. The tree from Cavassim *et al*. (2020) was pruned for the 24 strains with the drop. tip() function from R phytools package and replotted with the plotTree() function. Strains were routinely cultured on Tryptone Yeast (TY) agar or liquid medium during the experiments.

Determining indirect rhizobial interactions using TY supernatant assay

Potential facilitative and inhibitory pairwise strain interactions were determined indirectly based on each strain's

growth in the supernatant of every other strain. Strains were revived from frozen glycerol stocks in 40 ml of Tryptone Yeast broth (TY broth: 5 g tryptone, 2.5 g yeast extract, 1.47 g $CaCl_2$ per litre volume) for 48 h (28°C, 180 rpm) and 600 μl from each 40 ml culture was saved for later use as an inoculum. The remaining culture was centrifuged (10 min, 4000 rpm) and the supernatant was filtered through a 0.2 μm syringe filter. An equal volume of fresh 100% TY broth was added to produce a 50:50 supernatant-broth mixture (supernatant treatment) for each strain. Supernatant treatments thus accounted for the resource consumption and secreted metabolites by one strain, both of which could affect the growth of other target strains grown in said supernatant treatments. In addition, strains were grown in 100% TY and 50% TY (1:1 of 100% TY and deionized water) broth control treatments. The 50% TY control treatment was used to determine a strain's minimum expected growth from a supernatant treatment if the supernatant invoked no inhibitory effects on growth. This is because the amount of added TY broth is the same in 50% control and supernatant treatments. A 100% TY control treatment was used to ensure strains grew well in a complex nutrient medium and any observed reductions in growth were either due to lower nutrient broth concentrations (50% TY control) or inhibitory metabolites within the supernatant treatments (Fig. 2A).

To start the growth assays, 200 μl of each 1:1 supernatant-broth mixture were added to 96-well plates with five replicates per strain. Supernatant treatments were inoculated with the initial inocula (~0.2 μl) using a sterilized microplate pin replicator (Boekel). One well of each supernatant treatment was inoculated with water as a no-growth control. Strains were grown at 28°C and OD_{600} measurements (Tecan Infinite spectrophotometer) were taken at 0 h and after 24, 39, 48 and 62 h after strain inoculation as a proxy of growth. A total of 624 inoculant-supernatant combinations were analysed, including 100% and 50% TY treatments. Relative growth indices (RGIs) were calculated for all strains in all supernatant treatments after 62 h growth when strains were approaching carrying capacity (Fig. S10). RGIs were calculated as the ratio of the growth of strain *i* in strain *j*'s cell-free supernatant in relation to growth in its own supernatant for each combination (Fig. 1C; Fig. 2A; Fig. 3; Fig. 4B; Figs S2-S4; Fig. S7).

Therefore, a value of 1 indicates strain *i* grows equally well in the supernatant of strain *j* as in its own supernatant. If a strain grows to a higher optical density in another strain's supernatant than its own, it receives a score of >1 (i.e., displays facilitated growth from strain *j* supernatant). The converse is assumed if strain *i* grows to a lower optical density in strain *j* supernatant, than when grown in its own. An RGI more than 1 suggests that nutrient resources are left behind by a strain, or that metabolites excreted by a strain can be used for additional growth by another strain. An RGI less than 1 could suggest that the majority of nutrient resources available have been consumed by a strain; therefore, offering no additional nutrients than a 50% TY control treatment. Alternatively, it could suggest that a strain has secreted inhibitory metabolites that prevent growth of other strains. Therefore, to only evaluate the effect of resource consumption on rhizobial growth, additional RGIs were calculated by comparing a strains' growth in each treatment compared to when grown in the 50% TY control treatment (Fig. 1C; Fig. S5).

Determining direct rhizobial interactions in co-cultures using spot assays on TY agar plates

Direct inhibitory pairwise interactions (interference competition) between rhizobium strains were determined by spotting a liquid culture of each strain on a bacterial lawn of every other strain (Fig. 1D). Level of inhibition was determined as the inhibition halo diameter of the bacterial lawn around the spotted bacterial colony. For the assay, strains were grown in 5 ml TY broth for 48 h at 28°C. Optical density of cultures showed strains had grown to 0.055–0.09 OD_{600}. OA strains had significantly higher initial culture spot inoculum ODs compared to other genospecies groups (Kruskal–Wallis: chi-squared = 20.036, df = 3, $p < 0.001$). However, a simple linear regression confirmed that no significant association was observed between the size of the inhibition zone around the culture spot and the optical density of the inoculum used for the assay (Fig. S11; $coeff_{inoculumOD}$: −6.47, $p > 0.05$).

Four-hundred microlitres of culture was then mixed with 40 ml of soft TY agar (7.5% grams of agar per 1 l volume) and plated in square petri-dishes and left to cool. Two plates of soft agar were made for each strain and 12 strains were spotted on each plate. In addition, uninoculated 100% TY soft agar plates were used as a control. Two microlitres of each rhizobium culture was spotted onto the plates. Also, a control uninoculated TY broth spot was placed in the corner of each plate as a control to ensure no inhibition was observed from the sterile broth alone. Plates were incubated at 28°C and imaged with digital camera at 24, 48 and 72 h.

The inhibition zone diameters and culture spot diameters were compared to identify if the growth of the spotted strain correlated with the level of inhibition. A very weak positive correlation was identified (Pearson's correlation *R* statistic = 0.34, $p < 0.001$) and a simple linear regression found inhibition zone diameters increased by 2.187 mm on average for every 1 mm increase in culture spot diameter ($coeff_{spotdiameter}$: 2.187, $p < 0.001$). However, this relationship was heavily biassed by a few

strains producing small inhibition zones (Fig. S12A). When these samples were removed, the resulting correlation weakened (Pearson's correlation R statistic = 0.17, $p < 0.01$; Fig. S12B) and a simple linear regression identified that association with inhibition zone diameter reduced to 1.288 mm on average for every 1 mm increase in spot diameter ($coeff_{spotdiameter}$: 1.288, $p < 0.01$). To control for the growth of the spotted bacterium, the diameter of the spotted bacterial culture was subtracted from the diameter of the inhibition zone in all analyses. Inhibition zone diameters were calculated by subtracting the Feret diameter of the culture spot from the Feret diameter of the inhibition zone using ImageJ (v.1.52 k). All strain lawn and spot combinations were measured in three technical replicates (with the exception of strains SM137B and SM122A where two replicates were used, and strain SM152B where one spotting replicate was used). All replicates were used to calculate average mean inhibition zones for all pairwise strain combinations.

Characterizing the metabolic capacity of rhizobial genotypes

To determine differences in metabolic capacity, all strains were grown on 31 single substrates using EcoPlates (Biolog Hayward, CA, USA). The 31 single substrates were defined into the following resource type groups: amines, amino acids, carbohydrates, carboxylic acids, complex carbons, and phosphate carbon and water as a control (Table S11). All amino acids are assumed to be available in TY broth as subcomponents, and although other single substrates may not likely be present within TY broth, they provided additional understanding of a strain's metabolic potential. One replicate was generated for each strain, therefore, strains were grouped by genospecies, and substrates were grouped into above-mentioned resource type groups (Table S11), to enable statistical analysis.

Before measurements, all strains were grown in 10 ml TY broth for 48 h (28°C, 180 rpm), centrifuged to form a pellet and re-suspended into 10 ml PBS buffer, and incubated for 2 h at room temperature. One-hundred microlitres of bacterial suspension was added to each of the 31 Ecoplate carbon sources and water control wells. Reduction of tetrazolium dye within each carbon source well occurs when microbes can metabolize the resource and subsequently respire. Plates were incubated at 28°C and OD_{590} measurements of the developed dye coloration were taken at 72 h. ODs were normalized by subtracting the control water well OD from the substrate well ODs for each strain (Fig. S8). Strains generating normalized OD values greater than 0 for a substrate well were considered able to metabolize that particular substrate (Fig. S8). To explore whether a strain had a broad (resource generalist) or narrow (resource specialist) resource use range, the total number of substrates metabolized was counted and further used to calculate the mean number of substrates metabolized per genospecies group. In addition, the average well colour development (AWCD) of each strain was calculated to reflect an overall metabolic capability of each genospecies. AWCD was calculated using 72 h OD_{590} measurements of each well (Garland and Mills, 1991; Garland, 2006): AWCD = $[\Sigma(S-C)]/n$, where S is the substrate well OD_{590} value, C is the control well OD_{590} value and n is the number of substrates (i.e., 31 for AWCD across all substrate treatments). AWCD values were also calculated for substrates grouped into six resource-type groups (Table S11). Strain SM159 (OE) was removed from the analyses due to abnormally high OD values likely due to contamination.

Comparative genomic analyses of candidate gene clusters linked with strain interactions

To investigate potential genetic differences underlying rhizobial interactions, we used comparative genomic analyses to explore the presence of genes linked with bacteriocins, quorum-sensing, secondary metabolism and prophages using various pipelines. To compare the presence of different bacteriocins and quorum-sensing systems, we used BLASTn to search full genome assemblies of the 24 rhizobial strains for genes (listed in Table S14) known to be associated with these pathways in *Rhizobium leguminosarum* (Schripsema *et al*., 1996; Wisniewski-Dyé and Downie, 2002; Gonzalez and Marketon, 2003; McAnulla *et al*., 2007). BLASTn outputs were filtered to show only hits with an e-value less than 0.1. Genes were only considered present if hits spanned the full length of the query sequence. If no full-length matches could be identified for a gene in any strain, then, partial matches were evaluated. To compare differences in the presence of potentially important secondary metabolism clusters, such as antimicrobial compounds and siderophores, secondary metabolite biosynthesis gene clusters were searched using antiSMASH 5.0 with default settings (Blin *et al*., 2019). All identified secondary metabolite clusters containing two or more genes were included in the analysis, even if the cluster had no similarity to specific known clusters. Clusters that antiSMASH identified to have 100% similarity to known secondary metabolite clusters were displayed separately in the heatmap (e.g., Vicibactin). *fhuBCD* genes, which encode the Vicibactin siderophore uptake system, were identified using BLASTn and reported as present if the subject sequences had >90% percentage identity and matched the full length of the three query gene

sequences (Table S14). To compare the number of potential prophage regions across genospecies groups, putative prophage regions were searched using PHASTER and identified as either 'intact', 'incomplete' or 'questionable' based on default parameter settings (Arndt *et al*., 2016).

To evaluate whether the gene distributions in the 24 strains were representative of their distribution across genospecies in a larger strain dataset, we further searched with BLASTn for quorum-sensing, bacteriocin and secondary metabolite associated genes in 196 *Rlt* strains from Cavassim *et al*. (2020), which include the 24 strains used in this study (Supplementary File 2). BLASTn outputs were filtered to show only results with an e-value less than 0.1. BLAST hits were only included in the analysis if they spanned the full length of the query sequence. If for a gene no full-length matches could be identified in any strain, then partial matches were evaluated. Mean percentage identity and range for each gene were calculated using R (Supplementary File 2). All query sequence GenBank accession numbers can be found in Table S14.

Statistical analysis

All analyses were conducted in R (R Core Team, 2017). Multiple statistical approaches were used to analyse the data, including mixed-effect models, likelihood ratio (LR) tests and parametric bootstrapping of 95% confidence intervals. Supernatant interactions where a strain was grown in its own supernatant (therefore generating an RGI of 1) were excluded from the analysis, to avoid biassing data distributions. To analyse the effects of genospecies and farm treatment group on direct and indirect interactions between strains, maximum likelihood (ML) mixed-effect models were produced for both supernatant assays, and spot plating assays, with lme4 R package (v.1.1-21). For supernatant assays, inoculant genospecies group (OA, OC, OE, CC) and supernatant genospecies group were included as fixed effects, while individual inoculant strain IDs and supernatant strain IDs were categorized as crossed random effects. For spot plating assays, liquid culture spot genospecies group (OA, OC, OE, CC) and soft agar lawn genospecies group were included as fixed effects, and similarly, individual spotted strain IDs and soft agar lawn strain IDs were classed as crossed random effects. LmerTest generated *t*-values, degrees of freedom and *p*-values for fixed effect parameters in the models. To analyse the effects of individual strains on direct and indirect interactions between strains, the model containing both random effects was compared by likelihood ratio (LR) test with a reduced model with either inoculant strain ID or supernatant/soft agar strain ID random effect removed, using the anova() R function. To identify genospecies group differences for each variable, genospecies groups were ranked by average value, and the genospecies group with the lowest value set as the intercept. Therefore, for all variables in all models the intercept was set to the OA genospecies group. First, the full models were generated whereby fixed effects included an interaction. The significance of the fixed-effect interaction was tested by the likelihood ratio test using anova() by comparing the full-interaction model with a reduced model with no interaction. If chi-squared *p*-values were < 0.05, model fits were determined as significantly different. In addition, the reliability of the fixed effects was determined by parametric bootstrapping of fixed effects as 95% confidence intervals in the final models (bootMer and boot.ci with 1000 bootstrap replicates). Fixed-effect parameters with 95% confidence intervals that spanned zero were considered nonsignificant. In order to test whether some strains influenced the observed interaction effects, specific strains were removed, and the models were rerun to confirm fixed-effect parameters remained significant. To further determine if OC and CC strains displayed significant differences in inhibitory activity when acting as the inoculant or supernatant/soft agar strain, the estimated marginal means of interactions from the mixed-effect model were compared using emmeans package in R with Tukey adjusted p-value correction applied.

To identify the overall trend across supernatant interactions, Pearson's correlation coefficient and simple linear regression [lm() in R] were used. To investigate whether genetic similarity of strains determined indirect interactions, a multiple linear regression model was produced observing how RGI was influenced by a three-way interaction between pairwise average nucleotide identity (ANI), inoculant genospecies, and genospecies interaction type (i.e., strains are the same or different genospecies). LR test was undertaken to test the difference between the full model with three-way interaction against the reduced model with ANI interaction removed. Emmeans R package was further utilized to test for the significance of interaction slopes from zero. Furthermore, a Euclidean distance value for each pairwise strain comparison was calculated using the mean inhibition zone diameters produced by spotted strains on each soft agar lawn strain to represent the similarity of their inhibitory capacity profiles across lawn treatments. The similarity of inhibitory capacity profiles was then correlated with pairwise strain ANI values using a simple linear regression in R (Fig. S6).

Non-parametric Kruskal–Wallis test was used to compare metabolic capacities as a measure of AWCD (across 31 single substrate treatments) and to determine whether genospecies groups could metabolize a significantly different number of single substrates. Dunn's post-

hoc test was used to identify direct differences between groups from non-parametric Kruskal–Wallis tests with Bonferroni adjusted *p*-values. We used phylogenetic generalized least squares (PGLS) regressions from the R caper package to determine the correlation between strain's RGI in supernatant treatments and metabolic capacity (as a measure of AWCD across 31 single substrate treatments, while accounting for phylogenetic non-independence between the data points. The phylogeny generated for Fig. 1B was utilized to estimate the phylogenetic impact on the covariance among residuals and consequently correct for phylogenetic relatedness. The ML estimate of the lambda branch length transformation used to fit the model was set to 0.161 and likelihood ratio tests indicate that the ML estimate was not significantly different from 0 ($p = 0.2999$), but was significantly different from 1 ($p < 0.001$). Principal component analysis (PCA) was calculated with R prcomp using singular value decomposition to calculate principal components for explaining: (i) metabolic capacity (AWCD) of *Rlt* strains across six resource type groups; and (ii) the strain metabolic capacity across single substrate treatments (i.e., not grouped). PERMANOVA using adonis() in the R vegan package was used to test for significance of PCA clustering and significant pairwise genospecies interactions were identified with post hoc testing using pairwise. Adonis() and Bonferroni *p*-value correction.

To determine if there was a correlation between the presence of the Vicibactin (Fig. 5A) and *tra* quorum-sensing genes (Fig. 5B) in spotted or soft agar lawn strains, and inhibition zone sizes observed in direct interactions (Fig. 2B), we generated logistical regression models using Markov Chain Monte Carlo Generalized Linear Mixed Models (family = 'categorical') corrected for phylogenetic relatedness between strains using the MCMCglmm R package. We corrected for phylogenetic relatedness of strains either acting as the spotted strain or soft agar lawn strain using the phylogeny created previously (Fig. 1B), depending on whether gene presence was being evaluated in spotted or lawn strains. Models were run with a total of 5,050,000 iterations, of which 1000 iterations were sampled based on 50,000 burnings and with a thinning interval of 5000. Due to the response variable (gene presence) being binary, we used a fixed residual variance of one, in the prior modifications. Furthermore, to evaluate the significance of the non-random distribution of gene presence across genospecies, Fisher's Exact Test for Count Data were carried out on the gene presence data for the 196 strains in R (Table S18; Supplementary File 2).

All R scripts used for figure generation and statistical analyses are available in the GitHub repository: http://github.com/brydenfields/rhizobium_interactions.

Acknowledgements

The authors would like to thank members of the Friman lab group at the University of York for their comments on the manuscript, and M. Izabel A. Cavassim for generation of the phylogenetic information which was utilized for genomic analyses. This work was funded by Innovation Fund Denmark as part of the NCHAIN research consortium (4105-00007A, SUA), and initially conceptualized as part of research funded by British Ecological Society small grant (SR16/1349, EH). VPF is funded by the Royal Society (RSG \R1\180213 and CHL\R1\180031) and jointly by a grant from UKRI, Defra, and the Scottish Government, under the Strategic Priorities Fund Plant Bacterial Diseases programme (BB/T010606/1) at the University of York.

Author contributions

Conceptualization: EH, VPF; Methodology: BF, EKM, EH, JPWY, VPF; Validation: BF, EKM; Formal analysis: BF; Investigation: BF, EKM; Resources: EH, SUA, JPWY, VPF; Data curation: BF, EKM; Writing - original draft: BF, VPF; Writing - review and editing: BF, EH, SUA, JPWY, VPF; Visualization: BF; Supervision: EH, SUA, JPWY, VPF; Project administration: EH, VPF; Funding acquisition: EH, SUA, VPF.

References

Abd-Alla, M.H., El-Enany, A.-W.E., Nafady, N.A., Khalaf, D. M., and Morsy, F.M. (2014) Synergistic interaction of *Rhizobium leguminosarum* bv. Viciae and arbuscular mycorrhizal fungi as a plant growth promoting biofertilizers for faba bean (*Vicia faba* L.) in alkaline soil. *Microbiol Res* **169**: 49–58.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., and Wishart, D.S. (2016) PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* **44**: W16–W21.

Becker, J., Eisenhauer, N., Scheu, S., and Jousset, A. (2012) Increasing antagonistic interactions cause bacterial communities to collapse at high diversity. *Ecol Lett* **15**: 468–474.

Birtel, J., Walser, J.C., Pichon, S., Bürgmann, H., and Matthews, B. (2015) Estimating bacterial diversity for ecological studies: methods, metrics, and assumptions. *PLoS One* **10**: e0125356.

Bladergroen, M.R., Badelt, K., and Spaink, H.P. (2003) Infection-blocking genes of a symbiotic *rhizobium leguminosarum* strain that are involved in temperature-dependent protein secretion. *Mol Plant Microbe Interact* **16**: 53–64.

Blanco, A.R., Sicardi, M., and Frioni, L. (2010) Competition for nodule occupancy between introduced and native strains of *Rhizobium leguminosarum* biovar trifolii. *Biol Fertil Soils* **46**: 419–425.

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S.Y., *et al*. (2019) antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* **47**: W81–W87.

Boivin, S., Ait Lahmidi, N., Sherlock, D., Bonhomme, M., Dijon, D., Heulin-Gotty, K., *et al*. (2020) Host-specific competitiveness to form nodules in *rhizobium leguminosarum* symbiovar viciae. *New Phytol* **226**: 555–568.

Braga, R.M., Dourado, M.N., and Araújo, W.L. (2016) Microbial interactions: ecology in a molecular perspective. *Brazilian J Microbiol* **47**: 86–98.

Bruno, J.F., Stachowicz, J.J., and Bertness, M.D. (2003) Inclusion of facilitation into ecological theory. *Trends Ecol Evol* **18**: 119–125.

van Brussel, A.A., Zaat, S.A., Wijffelman, C.A., Pees, E., and Lugtenberg, B.J. (1985) Bacteriocin small of fast-growing rhizobia is chloroform soluble and is not required for effective nodulation. *J Bacteriol* **162**: 1079–1082.

Burke, C.M., and Darling, A.E. (2016) A method for high precision sequencing of near full-length 16S rRNA genes on an Illumina MiSeq. *PeerJ* **2016**: e2492.

Burns, N., James, C.E., and Harrison, E. (2015) Polylysogeny magnifies competitiveness of a bacterial pathogen in vivo. *Evol Appl* **8**: 346–351.

Cavassim, M.I.A., Moeskjaer, S., Moslemi, C., Fields, B., Bachmann, A., Vilhjálmsson, B., *et al*. (2019) The genomic architecture of introgression among sibling species of bacteria. *bioRxiv*. https://doi.org/10.1101/526707

Cavassim, M.I.A., Moeskjaer, S., Moslemi, C., Fields, B., Bachmann, A., Vilhjálmsson, B.J., *et al*. (2020) Symbiosis genes show a unique pattern of introgression and selection within a *rhizobium leguminosarum* species complex. *Microb Genomics* **6**(4): e000351.

Cha, C., Gao, P., Chen, Y.-C., Shaw, P.D., and Farrand, S. K. (1998) Production of acyl-Homoserine lactone quorum-sensing signals by gram-negative plant-associated bacteria.

Checcucci, A., DiCenzo, G.C., Bazzicalupo, M., and Mengoni, A. (2017) Trade, diplomacy, and warfare: the quest for elite rhizobia inoculant strains. *Front Microbiol* **8**: 2207.

Clúa, J., Roda, C., Zanetti, M.E., and Blanco, F.A. (2018) Compatibility between legumes and rhizobia for the establishment of a successful nitrogen-fixing Symbiosis. *Genes (Basel)* **9**: 125.

Crossman, L.C., Castillo-Ramírez, S., McAnnula, C., Lozano, L., Vernikos, G.S., Acosta, J.L., *et al*. (2008) A common genomic framework for a diverse assembly of plasmids in the symbiotic nitrogen fixing bacteria. *PLoS One* **3**(7): e2567. https://doi.org/10.1371/annotation/4a58a9bd-7531-41d0-b101-53039a8.

Cubo, M.T., Economou, A., Murphy, G., Johnston, A.W.B., and Downie, J.A. (1992) Molecular characterization and regulation of the rhizosphere-expressed genes rhiABCR that can influence nodulation by *rhizobium leguminosarum* biovar viciae. *J Bacteriol* **174**: 4026–4035.

Danino, V.E., Wilkinson, A., Edwards, A., and Downie, J.A. (2003) Recipient-induced transfer of the symbiotic plasmid pRL1JI in *rhizobium leguminosarum* bv. Viciae is regulated by a quorum-sensing relay. *Mol Microbiol* **50**: 511–525.

Denison, R.F., and Kiers, E.T. (2004) Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. *FEMS Microbiol Lett* **237**: 187–193.

Depret, G., Houot, S., Allard, M.-R., Breuil, M.-C., Nouaïm, R., and Laguerre, G. (2004) Long-term effects of crop management on *Rhizobium leguminosarum* biovar viciae populations. *FEMS Microbiol Ecol* **51**: 87–97.

Dobson, A., Cotter, P.D., Paul Ross, R., and Hill, C. (2012) Bacteriocin production: a probiotic trait? *Appl Environ Microbiol* **78**: 1–6.

Dolinšek, J., Goldschmidt, F., and Johnson, D.R. (2016) Synthetic microbial ecology and the dynamic interplay between microbial genotypes. *FEMS Microbiol Rev* **40**: 961–979.

Downie, J.A. (2010) The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. *FEMS Microbiol Rev* **34**: 150–170.

Dubiller, N., Mülders, C., Ferdelman, T., De Beer, D., Pernthaler, A., Klein, M., *et al*. (2001) Endosymbiotic sulphate-reducing and sulphide-oxidizing bacteria in an oligochaete worm. *Nature* **411**: 298–302.

Edwards, A., Frederix, M., Wisniewski-Dyé, F., Jones, J., Zorreguieta, A., and Allan Downie, J. (2009) The cin and rai quorum-sensing regulatory systems in *Rhizobium leguminosarum* are coordinated by ExpR and CinS, a small regulatory protein coexpressed with CinI. *J Bacteriol* **191**: 3059–3067.

Elias, S., and Banin, E. (2012) Multi-species biofilms: living with friendly neighbors. *FEMS Microbiol Rev* **36**: 990–1004.

Evans, C.R., Kempes, C.P., Price-Whelan, A., and Dietrich, L.E.P. (2020) Metabolic heterogeneity and cross-feeding in bacterial multicellular systems. *Trends Microbiol* **28**: 732–743.

Fields, B., Moeskjær, S., Friman, V., Andersen, S.U., and Young, J.P.W. (2020) MAUI-seq: Metabarcoding using amplicons with unique molecular identifiers to improve error correction. *Mol Ecol Resour* **00**: 1–18.

Foster, K.R., and Bell, T. (2012) Competition, not cooperation, dominates interactions among Culturable microbial species. *Curr Biol* **22**: 1845–1850.

Fuks, G., Elgart, M., Amir, A., Zeisel, A., Turnbaugh, P.J., Soen, Y., and Shental, N. (2018) Combining 16S rRNA gene variable regions enables high-resolution microbial community profiling. *Microbiome* **6**: 17.

Gardner, A., West, S.A., and Buckling, A. (2004) Bacteriocins, spite and virulence. *Proc R Soc B Biol Sci* **271**: 1529–1535.

Garland, J.L. (2006) Analysis and interpretation of community-level physiological profiles in microbial ecology. *FEMS Microbiol Ecol* **24**: 289–300.

Garland, J.L., and Mills, A.L. (1991) Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. *Appl Environ Microbiol* **57**: 2351–2359.

Geetha, S.J., and Joshi, S.J. (2013) Engineering Rhizobial bioinoculants: a strategy to improve iron nutrition. *Sci World J* **2013**: 315890.

Ghoul, M., and Mitri, S. (2016) The ecology and evolution of microbial competition. *Trends Microbiol* **24**: 833–845.

Gillor, O., Giladi, I., and Riley, M.A. (2009) Persistence of colicinogenic *Escherichia coli* in the mouse gastrointestinal tract. *BMC Microbiol* **9**: 165.

Girard, L., BLanchEt, E., StiEn, D., Baudart, J., Suzuki, M., and LaMi, R. (2019) Evidence of a large diversity of N-acyl-Homoserine lactones in symbiotic *Vibrio fischeri* strains associated with the squid *Euprymna scolopes*. *Microbes Env* **34**: 99–103.

Gonzalez, J.E., and Marketon, M.M. (2003) Quorum sensing in nitrogen-fixing rhizobia. *Microbiol Mol Biol Rev* **67**: 574–592.

Goris, J., Konstantinidis, K., Klappenbach, J., Coenye, T., Vandamme, P., and Tiedje, J. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81–91.

Gorter, F.A., Manhart, M., and Ackermann, M. (2020) Understanding the evolution of interspecies interactions in microbial communities. *Phil Trans R Soc B* **375**: 20190256.

Gray, K.M., Pearson, J.P., Downie, J.A., Boboye, B.E., Greenberg, E.P., Allan Downie, J., *et al*. (1996) Cell-to-cell signaling in the symbiotic nitrogen-fixing bacterium *Rhizobium leguminosarum*: autoinduction of a stationary phase and rhizosphere-expressed genes. *J Bacteriol* **178**: 372–376.

Griffin, A.S., West, S.A., and Buckling, A. (2004) Cooperation and competition in pathogenic bacteria. *Nature* **430**: 1024–1027.

Harcombe, W.R., Chacón, J.M., Adamowicz, E.M., Chubiz, L.M., and Marx, C.J. (2018) Evolution of bidirectional costly mutualism from byproduct consumption. *Proc Natl Acad Sci U S A* **115**: 12000–12004.

Harrison, F., Paul, J., Massey, R.C., and Buckling, A. (2008) Interspecific competition and siderophore-mediated cooperation in *Pseudomonas aeruginosa*. *ISME J* **2**: 49–55.

Harrison, S.P., Jones, D.G., and Young, J.P.W. (1989) Rhizobium population genetics: genetic variation within and between populations from diverse locations. *Microbiology* **135**: 1061–1069.

He, X., Chang, W., Pierce, D.L., Seib, L.O., Wagner, J., and Fuqua, C. (2003) Quorum sensing in rhizobium sp. strain NGR234 regulates conjugal transfer (tra) gene expression and influences growth rate. *J Bacteriol* **185**: 809–822.

Heemstra, J.R., Walsh, C.T., and Sattely, E.S. (2009) Enzymatic tailoring of ornithine in the biosynthesis of the rhizobium cyclic Trihydroxamate Siderophore Vicibactin. *J Am Chem Soc* **131**: 15317–15329.

Hibbing, M.E., Fuqua, C., Parsek, M.R., and Peterson, S.B. (2010) Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* **8**: 15–25.

Hirsch, P.R. (1979) Plasmid-determined Bacteriocin production by *rhizobium leguminosarum*. *Microbiology* **113**: 219–228.

Hmelo, L.R. (2017) Quorum sensing in marine microbial environments. *Ann Rev Mar Sci* **9**: 257–281.

Homa, M., Sándor, A., Tóth, E., Szebenyi, C., Nagy, G., Vágvölgyi, C., and Papp, T. (2019) In vitro interactions of *Pseudomonas aeruginosa* with scedosporium species frequently associated with cystic fibrosis. *Front Microbiol* **10**: 441.

Huang, W., Traulsen, A., Werner, B., Hiltunen, T., and Becks, L. (2017) Dynamical trade-offs arise from antagonistic coevolution and decrease intraspecific diversity. *Nat Commun* **8**: 2059.

Hunt, D.E., David, L.A., Gevers, D., Preheim, S.P., Alm, E. J., and Polz, M.F. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science (80-)* **320**: 1081–1085.

Joseph, M.V., Desai, J.D., and Desai, A.J. (1983) Production of antimicrobial and Bacteriocin-like substances by *rhizobium trifolii*. *Appl Environ Microbiol* **45**: 532–535.

Joshi, F.R., Kholiya, S.P., Archana, G., and Desai, A.J. (2008) Siderophore cross-utilization amongst nodule isolates of the cowpea miscellany group and its effect on plant growth in the presence of antagonistic organisms. *Microbiol Res* **163**: 564–570.

Jousset, A., Schmid, B., Scheu, S., and Eisenhauer, N. (2011) Genotypic richness and dissimilarity opposingly affect ecosystem functioning. *Ecol Lett* **14**: 537–545.

Kennedy, A.C., and Smith, K.L. (1995) Soil microbial diversity and the sustainability of agricultural soils. *Plant and Soil* **170**: 75–86.

Kiers, E.T., West, S.A., and Denison, R.F. (2002) Mediating mutualisms: farm management practices and evolutionary changes in symbiont co-operation. *J Appl Ecol* **39**: 745–754.

Koch, H., Germscheid, N., Freese, H.M., Noriega-Ortega, B., Lücking, D., Berger, M., *et al*. (2020) Genomic, metabolic and phenotypic variability shapes ecological differentiation and intraspecies interactions of *Alteromonas macleodii*. *Sci Rep* **10**: 809.

Koeppel, A.F., Wertheim, J.O., Barone, L., Gentile, N., Krizanc, D., and Cohan, F.M. (2013) Speedy speciation in a bacterial microcosm: new species can arise as frequently as adaptations within a species. *ISME J* **7**: 1080–1091.

Kramer, J., Özkaya, Ö., and Kümmerli, R. (2019) Bacterial siderophores in community and host interactions. *Nat Rev Microbiol* **18**: 152–163.

Kumar, N., Lad, G., Giuntini, E., Kaye, M.E., Udomwong, P., Jannah Shamsani, N., *et al*. (2015) Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*. *Open Biol* **5**: 140133.

Lamont, I.L., Beare, P.A., Ochsner, U., Vasil, A.I., and Vasil, M.L. (2002) Siderophore-mediated signaling regulates virulence factor production in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **99**: 7072–7077.

Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., and Reyes, J.A. (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* **31**: 814–821.

Larkin, A.A., and Martiny, A.C. (2017) Microdiversity shapes the traits, niche space, and biogeography of microbial taxa. *Environ Microbiol Rep* **9**: 55–70.

Lee, K.W.K., Hoong Yam, J.K., Mukherjee, M., Periasamy, S., Steinberg, P.D., Kjelleberg, S., and Rice, S.A. (2016) Interspecific diversity reduces and functionally substitutes for intraspecific variation in biofilm communities. *ISME J* **10**: 846–857.

Levin, B.R. (1988) Frequency-dependent selection in bacterial populations. *Philos Trans R Soc London B, Biol Sci* **319**: 459–472.

Lithgow, J.K., Wilkinson, A., Hardman, A., Rodelas, B., Wisniewski-Dye, F., Williams, P., and Downie, J.A. (2000) The regulatory locus cinRI in *rhizobium leguminosarum* controls a network of quorum-sensing loci. *Mol Microbiol* **37**: 81–97.

Liu, F., Hewezi, T., Lebeis, S.L., Pantalone, V., Grewal, P.S., and Staton, M.E. (2019) Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly. *BMC Microbiol* **19**: 201.

Lupwayi, N.Z., Clayton, G.W., and Rice, W.A. (2006) Rhizobial inoculants for legume crops. *J Crop Improv* **15**: 289–321.

Mazur, A., Stasiak, G., Wielbo, J., Koper, P., Kubik-Komar, A., and Skorupska, A. (2013) Phenotype profiling of *rhizobium leguminosarum* bv. Trifolii clover nodule isolates reveal their both versatile and specialized metabolic capabilities. *Arch Microbiol* **195**: 255–267.

McAnulla, C., Edwards, A., Sanchez-Contreras, M., Sawers1, R.G., Allan, J., Correspondence, D., and Downie, J.A. (2007) Quorum-sensing-regulated transcriptional initiation of plasmid transfer and replication genes in *Rhizobium leguminosarum* biovar viciae. *Microbiology* **153**: 2074–2082.

Mendoza-Suárez, M.A., Geddes, B.A., Sánchez-Cañizares, C., Ramírez-González, R.H., Kirchhelle, C., Jorrin, B., and Poole, P.S. (2020) Optimizing Rhizobium-legume symbioses by simultaneous measurement of rhizobial competitiveness and N2 fixation in nodules. *Proc Natl Acad Sci U S A* **117**: 9822–9831.

Miller, M.B., and Bassler, B.L. (2001) Quorum sensing in bacteria. **55**: 165–199.

Miranda-Sánchez, F., Rivera, J., and Vinuesa, P. (2016) Diversity patterns of Rhizobiaceae communities inhabiting soils, root surfaces and nodules reveal a strong selection of rhizobial partners by legumes. *Environ Microbiol* **18**: 2375–2391.

Mishra, D.J., Rajvir, S., Mishra, U.K., and Kumar, S.S. (2013) Role of bio-fertilizer in organic agriculture: a review. *Res J Recent Sci* **2**: 39–41.

Moeskjaer, S., Tausen, M., Andersen, S.U., Peter, J., and Young, W. (2020) Amplicons and isolates: rhizobium diversity in fields under conventional and organic management. *bioRxiv* 2020.09.22.307934.

Morris, B.E.L., Henneberger, R., Huber, H., and Moissl-Eichinger, C. (2013) Microbial syntrophy: interaction for the common good. *FEMS Microbiol Rev* **37**: 384–406.

Nadell, C.D., Drescher, K., and Foster, K.R. (2016) Spatial structure, cooperation and competition in biofilms. *Nat Rev Microbiol* **14**: 589–600.

Nuccio, E.E., Starr, E., Karaoz, U., Brodie, E.L., Zhou, J., Tringe, S.G., *et al*. (2020) Niche differentiation is spatially and temporally regulated in the rhizosphere. *ISME J* **14**: 999–1014.

Oakley, B.B., Carbonero, F., Van Der Gast, C.J., Hawkins, R.J., and Purdy, K.J. (2010) Evolutionary divergence and biogeography of sympatric niche-differentiated bacterial populations. *ISME J* **4**: 488–497.

Oresnik, I.J., Twelker, S., and Hynes, M.F. (1999) Cloning and characterization of a *Rhizobium leguminosarum* gene encoding a Bacteriocin with similarities to RTX toxins. *Appl Environ Microbiol* **65**: 2833–2840.

Ormeño-Orrillo, E., and Martínez-Romero, E. (2013) Phenotypic tests in rhizobium species description: an opinion and (a sympatric speciation) hypothesis. *Syst Appl Microbiol* **36**: 145–147.

Palmer, K.M., and Young, J.P.W. (2000) Higher diversity of *Rhizobium leguminosarum* biovar viciae populations in arable soils than in grass soils. *Appl Environ Microbiol* **66**: 2445–2450.

Peters, B.M., Jabra-Rizk, M.A., O'May, G.A., William Costerton, J., and Shirtliff, M.E. (2012) Polymicrobial interactions: impact on pathogenesis and human disease. *Clin Microbiol Rev* **25**: 193–213.

Plazinski, J., and Rolfe, B.G. (1985) Interaction of azospirillum and Rhizobium strains leading to inhibition of nodulation. *Appl Environ Microbiol* **49**: 990–993.

Portella, A.C.F., Karp, S., Scheidt, G.N., Woiciechwski, A.L., Parada, J.L., and Soccol, C.R. (2009) Modelling antagonic effect of lactic acid eacteria supernatants on some pathogenic bacteria. *Braz Arch Biol Technol* **52**: 29–36.

R Core Team. (2017) *R: A Language and Environment for Statistical Computing [Internet]*. Vienna, Austria, p. 2014.

Rainey, P.B., and Travisano, M. (1993) Adaptive radiation in a heterogeneous environment. *Stratigr Trans R Soc Edinb Earth Sci* **84**: 177–188.

Rice, W.A., Penney, D.C., and Nyborg, M. (1977) Effects of soil acidity on rhizobia numbers, nodulation and nitrogen fixation by alfalfa and red clover. *Can J Soil Sci* **57**: 197–203.

Riley, M.A., and Gordon, D.M. (1999) The ecological role of bacteriocins in bacterial competition. *Trends Microbiol* **7**: 129–133.

Rodelas, B., Lithgow, J.K., Wisniewski-Dye, F., Hardman, A., Wilkinson, A., Economou, A., *et al*. (1999) Analysis of quorum-sensing-dependent control of rhizosphere-expressed (rhi) genes in *rhizobium leguminosarum* bv. Viciae. *J Bacteriol* **181**: 3816–3823.

Rosenzweig, R.F., Sharp, R.R., Treves, D.S., and Adams, J. (1994) Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* **137**: 903–917.

Russel, J., Røder, H.L., Madsen, J.S., Burmølle, M., and Sørensen, S.J. (2017) Antagonism correlates with metabolic similarity in diverse bacteria. *PNAS* **114**: 10684–10688.

Sanchez-Contreras, M., Bauer, W.D., Gao, M., Robinson, J. B., Downie, J.A., and Allan Downie, J. (2007) Quorum-sensing regulation in rhizobia and its role in symbiotic interactions with legumes. *Philos Trans R Soc Lond B Biol Sci* **362**: 1149–1163.

Schmidt, J.E., Vannette, R.L., Igwe, A., Blundell, R., Casteel, C.L., and Gaudin, A.C.M. (2019) Effects of agricultural management on rhizosphere microbial structure and function in processing tomato plants. *Appl Environ Microbiol* **85**: e01064–e01019.

Schreiber, F., and Ackermann, M. (2020) Environmental drivers of metabolic heterogeneity in clonal microbial populations. *Curr Opin Biotechnol* **62**: 202–211.

Schripsema, J., De Rudder, K.E.E., Van Vliet, T.B., Lankhorst, P.P., De Vroom, E., Kijne, J.W., and Van Brussel, A.A.N. (1996) Bacteriocin small of *Rhizobium leguminosarum* belongs to the class of N-acyl-L-

Homoserine lactone molecules, known as autoinducers and as quorum sensing co-transcription factors. *J Bacteriol* **178**: 366–371.
Schwinghamer, E.A.A., and Brockwell, J. (1978) Competitive advantage of bacteroicin- and phage-producing strains of *Rhizobium trifolii* in mixed culture. *Soil Biol Biochem* **10**: 383–387.
Shu, W., Pablo, G.P., Jun, Y., and Danfeng, H. (2012) Abundance and diversity of nitrogen-fixing bacteria in rhizosphere and bulk paddy soil under different duration of organic management. *World J Microbiol Biotechnol* **28**: 493–503.
Silva, V.M.A., Martins, C.M., Cavalcante, F.G., Alves Ramos, K., Lopes da Silva, L., Rodrigues de Menezes, F. G., *et al*. (2019) Cross-feeding among soil bacterial populations: selection and characterization of potential bio-inoculants. *J Agric Sci* **11**: 23–34.
Slattery, M., Rajbhandari, I., and Wesson, K. (2001) Competition-mediated antibiotic induction in the marine bacterium *Streptomyces tenjiimariensis*. *Microb Ecol* **41**: 90–96.
Speare, L., Cecere, A.G., Guckes, K.R., Smith, S., Wollenberg, M.S., Mandel, M.J., *et al*. (2018) Bacterial symbionts use a type VI secretion system to eliminate competitors in their natural host. *PNAS* **115**: E8528–E8537.
Stevens, J.B., Carter, R.A., Hussain, H., Carson, K.C., Dilworth, M.J., and Johnston, A.W.B. (1999) The fhu genes of *Rhizobium leguminosarum*, specifying siderophore uptake proteins: FhuDCB are adjacent to a pseudogene version of fhuA. *Microbiology* **145**: 593–601.
Strassmann, J.E., Gilbert, O.M., and Queller, D.C. (2011) Kin discrimination and cooperation in microbes. *Annu Rev Microbiol* **65**: 349–367.
Stubbendieck, R.M., and Straight, P.D. (2016) Multifaceted interfaces of bacterial competition. *J Bacteriol* **198**: 2145–2155.
Triplett, E.W., and Barta, T.M. (1987) Trifolitoxin production and nodulation are necessary for the expression of superior nodulation competitiveness by *Rhizobium leguminosarum* bv. Trifolii strain T24 on clover. *Plant Physiol* **85**: 335–342.
Tyc, O., van den Berg, M., Gerards, S., van Veen, J.A., Raaijmakers, J.M., de Boer, W., and Garbeva, P. (2014) Impact of interspecific interactions on antimicrobial activity among soil bacteria. *Front Microbiol* **5**: 567.
Van Der Heijden, M.G.A., Bardgett, R.D., and Van Straalen, N.M. (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* **11**: 296–310.
Vuong, H.B., Thrall, P.H., and Barrett, L.G. (2017) Host species and environmental variation can influence rhizobial community composition. *J Ecol* **105**: 540–548.
Wang, J., Mei, X., Wei, Z., Raza, W., & Shen, Q. (2021) Effect of bacterial intra-species community interactions on the production and activity of volatile organic compounds. *Soil Ecol Lett* **3**(1): 32–41.
Wang, X., Zhou, H., Chen, H., Jing, X., Zheng, W., Li, R., *et al*. (2018) Discovery of recombinases enables genome mining of cryptic biosynthetic gene clusters in Burkholderiales species. *Proc Natl Acad Sci U S A* **115**: E4255–E4263.
Weese, D.J., Heath, K.D., Dentinger, B.T.M., and Lau, J.A. (2015) Long-term nitrogen addition causes the evolution of less-cooperative mutualists. *Evolution (N Y)* **69**: 631–642.
Wielbo, J., Marek-Kozaczuk, M., Mazur, A., Kubik-Komar, A., and Skorupska, A. (2010) Genetic and metabolic divergence within a *rhizobium leguminosarum* bv. Trifolii population recovered from clover nodules. *Appl Environ Microbiol* **76**: 4593–4600.
Wilkinson, A., Danino, V., Wisniewski-Dyé, F., Lithgow, J.K., and Downie, J.A. (2002) N-acyl-homoserine lactone inhibition of rhizobial growth is mediated by two quorum-sensing genes that regulate plasmid transfer. *J Bacteriol* **184**: 4510–4519.
Wilson, R.A., Handley, B.A., and Beringer, J.E. (1998) Bacteriocin production and resistance in a field population of *rhizobium leguminosarum* biovar viciae. *Soil Biol Biochem* **30**: 413–417.
Wisniewski-Dyé, F., and Downie, J.A. (2002) Quorum-sensing in Rhizobium. *Antonie Van Leeuwenhoek* **81**: 397–407.
Wisniewski-Dyé, F., Jones, J., Chhabra, S.R., and Downie, J.A. (2002) raiIR genes are part of a quorum-sensing network controlled by cinI and cinR in *rhizobium leguminosarum*. *J Bacteriol* **184**: 1597–1606.
Wright, W., Little, J., Liu, F., and Chakraborty, R. (2013) Isolation and structural identification of the trihydroxamate siderophore vicibactin and its degradative products from *rhizobium leguminosarum* ATCC 14479 bv. Trifolii. *Biometals* **26**: 271–283.
Xiong, H.Y., Zhang, X.X., Guo, H.J., Ji, Y.Y., Li, Y., Wang, X.L., *et al*. (2017) The epidemicity of facultative microsymbionts in faba bean rhizosphere soils. *Soil Biol Biochem* **115**: 243–252.
Yadav, V., Gupta, J., Mandhan, R., Chhillar, A.K., Dabur, R., Singh, D.D., and Sharma, G.L. (2005) Investigations on anti-aspergillus properties of bacterial products. *Lett Appl Microbiol* **41**: 309–314.
Young, J.P.W., Moeskjaer, S., Afonin, A., Rahi, P., Maluk, M., James, E.K., *et al*. (2021) Defining the *Rhizobium leguminosarum* species complex. *Genes 2021* **12**: 111.
Zee, P.C., and Bever, J.D. (2014) Joint evolution of kin recognition and cooperation in spatially structured rhizobium populations. *PLoS One* **9**: e95141.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Correlations of relative growth index (RGI) compared to Area under the Curve index (AUC index) and Maximum Growth Rate index. AUC index was calculated by dividing the inoculant strain i's mean AUC value when grown in their own supernatant by their mean AUC value when grown in strain j's supernatant. Maximum Growth Rate index was calculated by dividing the inoculant strain *i*'s mean Maximum Growth Rate value when grown in their own supernatant by their mean Maximum Growth Rate value when grown

in strain *j*'s supernatant. Correlation was calculated by Pearson's correlation coefficient with R cor() function.

Figure S2. (A) Distribution of the relative growth index of all strains grown in each other's supernatants. Relative growth index is calculated as described in the methods. Bin size = 0.01 (B) Average nucleotide identity (ANI) of interacting strains correlated to relative growth index (RGI) of strain grown in other strain's supernatant. Inoculant strains interacting with a strain from another genospecies group are indicated as 'between' genospecies interaction, whereas strains of the same genospecies group interacting are indicated as 'within' genospecies interaction. Lines of best fit are calculated via linear model. C) The between genospecies group panel in b presented with the x-axis reduced to view the inoculant genospecies groups at lower ANI values.

Figure S3. The mean growth of Rlt strains in different supernatants averaged over genospecies, and average genospecies inoculant inhibition zone formation on different soft agar treatments. (A) Mean RGIs of genospecies groups (OA, OC, OE, CC) inoculated (e.g., I-OA) into the supernatants of other genospecies groups (e.g., Sup-OA). (A) Mean inhibition zone diameter (mm) of genospecies groups when liquid cultures (e.g., Spot-OA) are spotted onto soft agar lawns of other genospecies groups strains (e.g., Lawn-OA). Error bars display 95% confidence intervals. Strain combinations were grouped by genospecies inoculant group and genospecies supernatant/soft agar group.

Figure S4. (A) Mean RGIs of each genospecies group inoculants in all other strain supernatants. (B) Mean RGIs of all genospecies group inoculants in each genospecies group supernatant. (C) Mean inhibition zone diameter around each genospecies group strains spotted on soft agar lawns of all other strains. (D) Mean inhibition zone diameter of spotted strains on each genospecies soft agar lawn. Error bars display 95% confidence intervals. Rhizobia strain combinations were grouped by genospecies inoculant group and genospecies supernatant/soft agar group.

Figure S5. Relative growth indices (RGIs) of Rhizobium leguminosarum symbiovar trifolii strains inoculated into each other's supernatants (1:1 mixture of supernatant and Tryptone Yeast (TY) both) compared to their growth in 50% TY. OA = organic genospecies A, OC = organic genospecies C, OE = organic genospecies E, and CC = conventional genospecies C.

Figure S6. Average Nucleotide Identity between spotted and soft agar lawn strain correlated to Euclidean distance of strain's inhibitory capacity profile when acting as the spotted strain. Inhibitory capacity profile was generated by calculating the Euclidean distance between strains based on the mean inhibition zone sizes produced when acting as the spotted strain on each soft agar lawn strain. Regression lines are fit by linear model.

Figure S7. Growth in supernatant correlated to size of inhibition zone on soft agar. If strains are able to produce an inhibition zone, it is more likely that they have a larger inhibition zone on a strain's agar lawn if they grew well in that same strain's supernatant. Mean Relative Growth Index of strain *i* as inoculant in the supernatant of strain *j* (n=5) correlate to the mean inhibition zone diameter (mm) produced by strain *i* on soft agar lawns of strain *j* (n<=3). Regression lines are fit by linear model.

Figure S8. Metabolic capacity of 23 Rhizobium leguminosarum symbiovar trifolii strains on 31 single substrate growth treatments. Ability to metabolize substrates was determined using Biolog Ecoplates and measuring OD590 nm of tetrazolium dye in each well. OD590 nm values were normalized by subtracting control well OD (water) from the substrate well OD after 72 h growth. Values of 0.00 OD590 nm or less were identified as no observable substrate metabolism (red), values of more than 0.00 OD590 nm were considered to have putative capacity to metabolize the substrate (blue).

Figure S9. Principal Components analysis for metabolic utilization of each Ecoplate substrate by 23 Rlt strains. (A) Each point represents a Rhizobium strain and is coloured and grouped by genospecies group. PC1 accounted for 38.9% of the variance, and PC2 23.6%. (B) The association of the variables with the first two principal components, which are coloured by percentage contribution of each specific variable to the first two principal components. OA = organic genospecies A, OC = organic genospecies C, OE = organic genospecies E, and CC = conventional genospecies C.

Figure S10. Growth curves of 24 *Rhizobium leguminosarum* symbiovar *trifolii* strains grown in each other's supernatants in pairwise combinations. Data were grouped by genospecies and farm treatment categories. Strains were additionally grown in 100% and 50% Tryptone Yeast (TY) broth as controls. Optical density (OD_{600}) of strains was measured for 62 h, and values were normalized by subtracting the 0 h time point optical density. Error bars represent one standard error of the mean. Headings indicate the source of the supernatant, colours indicate the growing strains. OA = organic genospecies A (n = 6, blue), OC = organic genospecies C (n =7, dark green), OE = organic genospecies E (n = 4, pink), and CC = conventional genospecies C (n = 6, light green).

Figure S11. Optical density of strain inocula used for liquid culture spotting compared to inhibition zone diameter produced from culture spot on soft lawns of all other strains. Inhibition zone diameter was calculated by deducting the Feret diameter of the culture spot from the Feret diameter of the inhibition zone. Points are coloured by the genospecies grouping of the inoculum strain.

Figure S12. (A) Correlation between the diameter of the inhibition zone (mm) and the diameter of the liquid culture spot (mm) after 72 h growth. (B) Correlation between the diameter of the inhibition zone (mm) and the diameter of the liquid culture spot (mm) after 72 h growth after removal of liquid culture spots with a diameter of less than 1 mm.

Table S1. Linear mixed effects models for supernatant indirect inhibition assay. Model formula: Mean Relative Growth Index ~ inoculant group *supernatant group + (1|inoculant strain) + (1|supernatant strain).

Table S2. Linear mixed effects models for supernatant indirect inhibition assay, with strains SM154C and SM168A removed. Model formula: Mean Relative Growth Index ~ inoculant group * supernatant group + (1|inoculant strain) + (1|supernatant strain).

Table S3. Estimated marginal means of genospecies supernatant effects on genospecies inoculant growth. Estimates are calculated based on the full model.

Table S5. Linear regression model for Supernatant Relative Growth Index against genetic relatedness of interacting strains (Average Nucleotide Identity). Model formula: Mean Relative Growth Index ~ Average Nucleotide Identity * inoculant group * genospecies interaction.

Table S6. Estimates of slopes of the ANI covariate trend for each level of inoculant genospecies and genospecies interaction type (i.e. between genospecies or within genospecies) for indirect supernatant interactions. Estimates are calculated based on the full model. These slopes can be viewed in **Figure S1B**.

Table S7. Linear mixed effects models for spot plating direct inhibition assay. Model formula: Mean inhibition zone diameter ~ spot group * lawn group + (1|spot strain) + (1|lawn strain).

Table S8. Estimated marginal means of genospecies soft agar treatment effect on inhibition zone formation by different genospecies spotted inoculants. Estimates are calculated based on the full model

Table S9. Estimated marginal means of genospecies spotted inoculant inhibition zone formation on different genospecies soft agar lawn treatments. Estimates are calculated based on the full model.

Table S10. Parametric bootstrapping of fixed effects 95% confidence intervals for indirect supernatant (Table S1, Table S2) and direct spot plating (Table S7, Table S8) models.

Table S11. Ecoplate single substrate growth treatments grouped into 6 resource type groups.

Table S12. PERMANOVA of strains' metabolic capacity for 6 resource type groups based on average well colour development of 31 single substrate growth treatments. Genospecies groups correspond to OA, OC, OE and CC.

Table S13. The number of secondary metabolite biosynthesis gene clusters identified in 24 *Rhizobium leguminosarum* symbiovar trifolii strains. antiSMASH was used to identify the gene clusters, and putative clusters were only counted if gene cluster regions contained at least 2 identifiable metabolite biosynthesis related genes. Clusters with 100% antiSMASH similarity to other known clusters were reported as individual columns in the table (e.g., Vicibactin). fhuBCD genes were identified using BLASTn and reported as present if the gene was identified within a strain with > 90% percentage identity and matching the full length of the query sequences. Strains are divided into 4 genospecies/environmental origin groups; OA = organic genospecies A, OC = organic genospecies C, OE = organic genospecies E, and CC = conventional genospecies C. T3PKS = Type III polyketide synthases, Hserlactone = Homoserine lactone cluster, Fused = Pheganomycin-style protein ligase-containing cluster, NRPS = Non-ribosomal peptide synthetase cluster.

Table S14. GenBank accessions of known quorum sensing, bacteriocin and secondary metabolite associated gene sequences. QS refers for quorum sensing; NRPS refers to non-ribosomal peptide synthase.

Table S15. Percentage identity of *Rhizobium leguminosarum* bacteriocins and quorum sensing genes found in 24 Rlt strains to reference sequences

Table S16. Markov Chain Monte Carlo Generalized Linear Mixed Model Logistical Regression with phylogenetic correction evaluating the association between tra quorum sensing system presence in either the spotted strain or soft agar lawn strain and inhibition zone size of strain combinations in direct interaction experiments. G-structure accounts for random effect variance, R-structure accounts for residual variance, and fixed effects are also labelled. Number of sampled iterations = 1000, Number of Burning iterations = 50000, Thinning = 5000, Number of total iterations = 5050000. DIC = Deviance Information Criterion. Priors were set with the following R code: prior1 <- list(R = list(V = 1, fix = 1), G = list(G1 = list(V = 1, nu = 1, alpha.mu = 0, alpha.V = 1000), G2 = list(V = 1, nu = 1, alpha.mu = 0, alpha.V = 1000))). Model formulas can be found in the R scripts deposited on the github page: http://github.com/brydenfields/rhizobium_interactions.

Table S17. The number of putative prophage regions identified in 24 Rhizobium leguminosarum symbiovar trifolii strains. Prophage regions were identified using PHASTER, which assigns a completeness score to identified regions based on the proportion of phage-related genes within the region; 'intact', 'questionable' or 'incomplete'. The strain can be divided into 4 genospecies/environmental origin groups; OA = organic genospecies A, OC = organic genospecies C, OE = organic genospecies E, and CC = conventional genospecies C.

Table S18. Fisher's Exact Test for Count Data carried out on each individual gene to test whether the distribution of gene presence across genospecies A–E is significantly different from a normal/random distribution in the 196 strain dataset from Cavassim et al. (2020).

Table S19. Twenty-four Rhizobium leguminosarum symbiovar trifolii strains isolated from Trifolium repens nodules across Danish farm sites. OA = organic genospecies A, OC = organic genospecies C, OE = organic genospecies E, and CC = conventional genospecies C.

Supplementary File 2 Table 1. Percentage identity of Rhizobium leguminosarum bacteriocins, quorum sensing and secondary metabolite associated genes found in 196 Rlt strains (Cavassim et al., 2020) including the 24 strains used in this study. Mean and range calculations were only based on subject sequences which span the full length of the query sequence. If no full-length sequence matches were found, and only partial matches were found in all strains, then the partial sequence match percentage identies were used. BLASTn outputs were filtered to only show hits with an evalue less than 0.1. *multiple gene copies present in strains, and therefore percentage identities for all gene copies were included in mean and range calculations. For each gene, the number of strains within multiple gene copies within each genospecies are as follows: raiI: gsC = 9 strains, rhiR: gsC = 2 strains.