



OPEN

Higher-order interactions shape microbial interactions as microbial community complexity increases

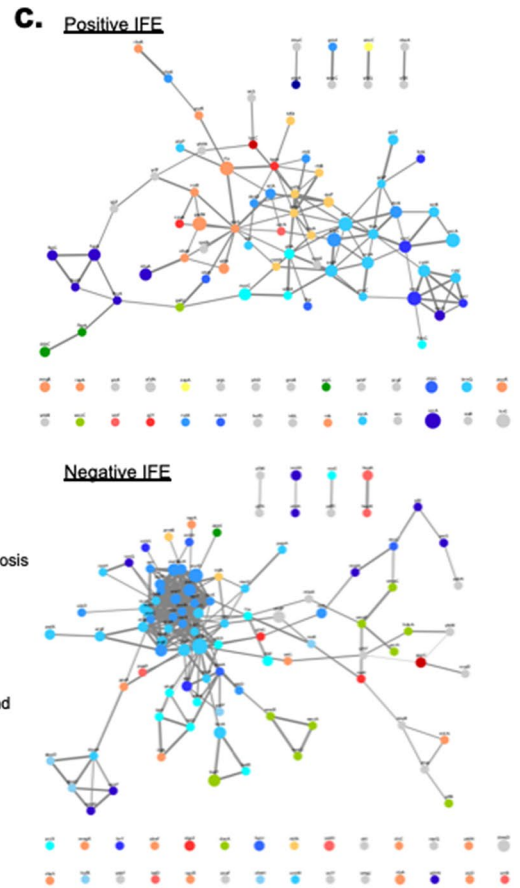
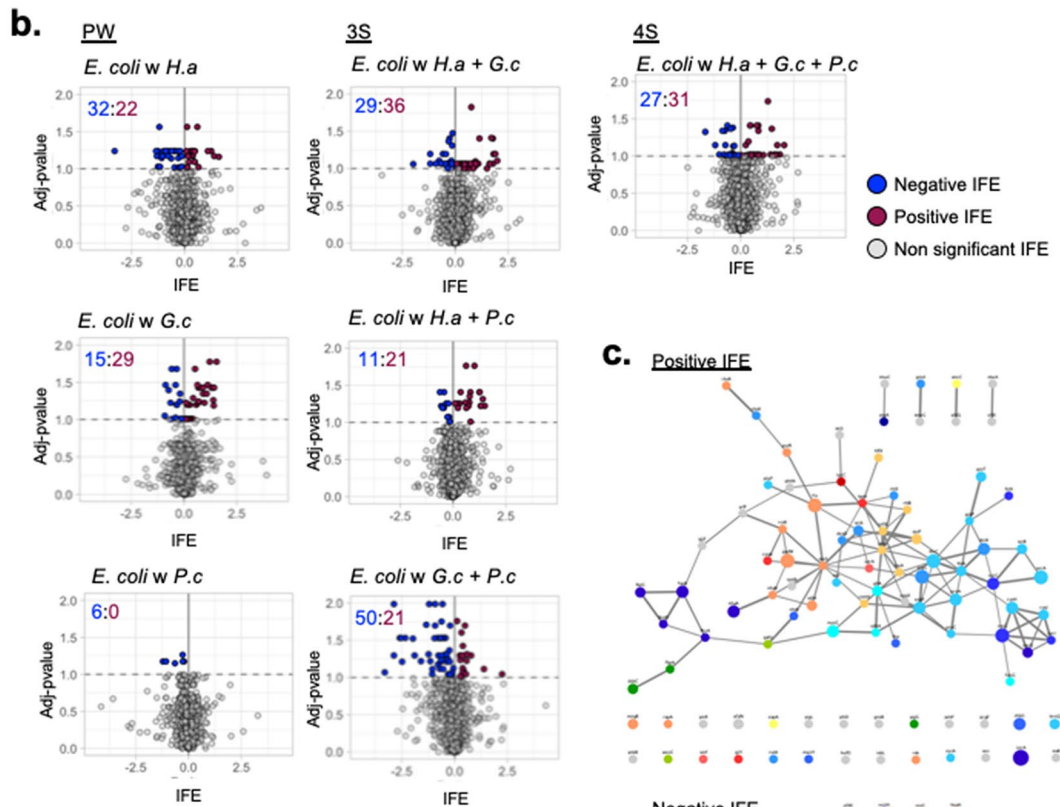
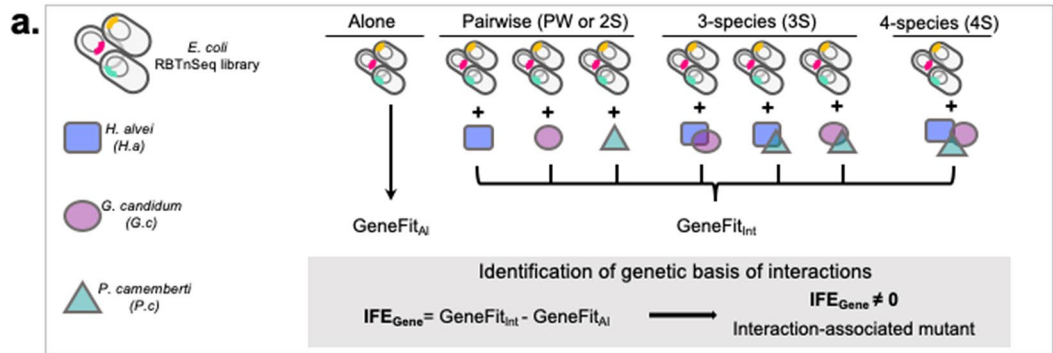
Manon A. Morin^{1✉}, Anneliese J. Morrison^{2,3}, Michael J. Harms^{2,3} & Rachel J. Dutton^{1✉}

Non-pairwise interactions, or higher-order interactions (HOIs), in microbial communities have been described as significant drivers of emergent features in microbiomes. Yet, the re-organization of microbial interactions between pairwise cultures and larger communities remains largely unexplored from a molecular perspective but is central to our understanding and further manipulation of microbial communities. Here, we used a bottom-up approach to investigate microbial interaction mechanisms from pairwise cultures up to 4-species communities from a simple microbiome (*Hafnia alvei*, *Geotrichum candidum*, *Penicillium camemberti* and *Escherichia coli*). Specifically, we characterized the interaction landscape for each species combination involving *E. coli* by identifying *E. coli*'s interaction-associated mutants using an RB-TnSeq-based interaction assay. We observed a deep reorganization of the interaction-associated mutants, with very few 2-species interactions conserved all the way up to a 4-species community and the emergence of multiple HOIs. We further used a quantitative genetics strategy to decipher how 2-species interactions were quantitatively conserved in higher community compositions. Epistasis-based analysis revealed that, of the interactions that are conserved at all levels of complexity, 82% follow an additive pattern. Altogether, we demonstrate the complex architecture of microbial interactions even within a simple microbiome, and provide a mechanistic and molecular explanation of HOIs.

Microbiomes are multidimensional systems containing up to thousands of interacting species. In part due to this complexity, a common strategy to investigate their biology has been to use bottom-up or reductionist approaches. Using in vitro microbial communities or monocultures, the features of microbial communities are measured in low-dimensional settings (or pairwise settings), with the underlying objective to understand whether high-dimension phenotypes (phenotype for communities of 3 species and more) can be predicted from lower-level observations. The lack of predictability of microbial community features from monoculture observations has demonstrated the critical importance of microbial interactions in shaping microbial communities. For instance, community assembly^{1,2}, community function³, community resistance to invasion⁴ or its effects on its host⁵ are different from the simple combination of individual species effects. Some of these studies have also shown that, while 2-species, or pairwise culture, observations could partially predict what happens in higher communities^{1,2,6}, context-specific phenotypes emerge as the microbial community becomes more complex, limiting the use of pairwise information to obtain a descriptive picture of larger communities. For instance, the function (amyolytic activity) of a soil-associated microbial community of 6 species has been shown to differ from the simple linear combination (*i.e.* the addition) of the observed function of pairwise cultures³. Similarly, 13 to 43% of measured *Drosophila melanogaster* traits (lifespan, reproduction and development) in flies carrying a microbiome of 5 species were not predictable from the associated traits from flies inoculated with pairwise combinations of the same 5 species⁵. In the zebrafish gut microbiome, while strong negative pairwise interactions were measured, the assembly of a more complex microbiome of 5 species strikingly diverged from any assembly prediction based on pairwise interactions, and all 5 species actually co-occur in the final microbiome⁷.

Reorganization of the interaction profile with varying community complexity and the presence of higher-order interactions (HOIs) are the most likely explanation for the lack of predictability of complex community phenotypes from pairwise observations. In ecology, HOIs are described as key to stabilizing communities and promoting biodiversity in biological ecosystems^{8–10}. They are commonly described as the modification of pairwise interactions when another species is introduced, or any interaction that cannot be described by a pairwise model^{11,12}. For instance, while the bacterium *Escherichia coli* can successfully invade cultures of *Chlamydomonas*

¹School of Biological Science, University of California San Diego, San Diego 92093, USA. ²Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR, USA. ³Institute of Molecular Biology, University of Oregon, Eugene, OR, USA. ✉email: manon.morin@ymail.com; rjdutton@ucsd.edu



COG Categories

- Energy production and conversion
- Amino acid metabolism and transport
- Nucleotide metabolism and transport
- Carbohydrate metabolism and transport
- Coenzyme metabolism
- Inorganic ion transport and metabolism
- Secondary structure
- Cell wall/membrane/envelop biogenesis
- Cell motility
- Post-translational modification, protein turnover, chaperone
- Function unknown
- Cell cycle control and mitosis
- Translation
- Transcription
- Replication and repair
- Signal transduction
- Intracellular trafficking and secretion

◀ **Figure 1.** Changes of *E. coli*'s genes associated with interaction-associated mutants in 2-species, 3-species and 4-species cultures. **(a)** Experimental design for the identification of interaction-associated mutants in 7 interactive conditions from the Brie community. The *E. coli* RB-TnSeq Keio_ML9 (Wetmore et al. 2015) is either grown alone or in 2, 3 or 4 species cultures to calculate *E. coli* gene fitness in each condition (in triplicate). Interaction fitness effect (IFE) is calculated for each gene in each interactive culture as the difference of the gene fitness in the interactive condition and in growth alone. IFE that are significantly different from 0 (two-sided t-test, Benjamini–Hochberg correction for multiple comparisons) highlight interaction-associated mutants in an interactive condition. **(b)** Volcanoplots of IFEs calculated for each interactive condition. Adjusted p-values lower than 0.1 highlight significant IFEs. Negative IFEs (blue) identify negative interactions and positive IFE (red) identify positive interactions. Numbers on each plot indicate the number of negative (blue) or positive (red) IFEs. **(c)** Functional analysis of the interaction-associated genes (significant IFEs). Genes of interaction-associated mutants have been separated into two groups: negative IFE and positive IFE. For each group, we represent the STRING network of the genes associated with interaction-associated mutants (Nodes). Edges connecting the genes represent both functional and physical protein association and the thickness of the edges indicates the strength of data support (minimum required interaction score: 0.4—medium confidence). Nodes are colored based on their COG annotation and the size of each node is proportional to the number of interactive conditions in which that given gene has been found associated with a significant IFE. Higher resolution of the networks with apparent gene names are found in Supplementary Figs. 2, 3.

reinhardtii as well as cultures of *Tetrahymena thermophila*, it is unable to invade a coculture of the alga and the ciliate. This is due to *C. reinhardtii* inhibition of *E. coli* aggregation specifically in the presence of *T. thermophila*, which renders the bacterium vulnerable to predation¹³. Changes in microbial interactions and HOIs are undoubtedly a significant feature of microbiome ecology and functioning. Thus, investigating HOIs and deciphering how microbial interactions are rearranged when microbial systems increase in complexity are essential to our understanding of microbiomes and our ability to manipulate them.

In this work, we investigate how interactions are reorganized when the complexity of a microbial community increases from two to four species. We use a simplified cheese rind microbiome composed of the two gamma-proteobacteria *Escherichia coli* and *Hafnia alvei*, the yeast *Geotrichum candidum* and the filamentous fungus *Penicillium camemberti*. Our previous work comparing the pairwise interaction patterns of this model microbiome to the interaction pattern in the full community highlighted the prevalence of HOIs, including both the lack of conservation of pairwise interactions and the emergence of non-pairwise interactions¹⁴. However, we couldn't precisely resolve the origin of these HOIs or specific rules underlying the emergence of these HOIs. For instance, we couldn't determine whether 2-species interactions are no longer happening after the introduction of a specific species or the introduction of all the other species. Similarly, we couldn't conclude whether community-specific interactions are actually specific to the community or whether they arise at an intermediate level and are maintained in the whole community. Relying on the ability to deconstruct and reconstruct this model system and on the RB-TnSeq-based interaction assay we have previously optimized to compare gene fitness values across multiple conditions¹⁵, we aim to identify quantitative changes in gene fitness values for a pooled RB-TnSeq library of *E. coli* in different interactive conditions to identify the genetic basis of interactions at every level of community complexity. The pooled library of *E. coli*'s insertion mutants allows us to probe the environment created by the presence of other species compared to growth alone and thus to infer potential resulting interactions.

Comparing interaction-associated mutants across conditions, we can identify pairwise interaction-associated mutants that are also observed at higher levels of complexity as well as interaction-mutants associated with HOIs. Here, interaction-associated mutants associated with HOIs are defined as any interaction-associated mutant never found in any 2-species condition that emerges in the 3 or 4 species cultures as well as any 2-species interaction-associated mutant that is no longer observed in the 3 or 4 species cultures. Analysis of the genes associated with HOIs-mutants and their functions allowed us to further characterize the deep reorganization of the interaction landscape and highlights that it is mostly associated with the reprogramming of metabolic interactions and the introduction of a fungal partner. In the last part of this work, we then focus on the genetic basis of 2-species interaction-associated mutants that are maintained in higher levels of complexity to elucidate quantitative principles behind interaction conservation. We use an epistasis and quantitative genomics approach^{16,17} to understand whether interactions that are maintained follow a linear, or additive, pattern. For each interaction-associated gene in each interactive condition, we measured an interaction fitness effect that corresponds to the gene fitness change between the interactive condition and growth alone. The evaluation of interaction fitness effects as quantitative traits allows us to investigate whether maintained 2-species interactions follow an additive behavior in more complex conditions and to define another form of HOIs as cases in which there is a lack of additivity in interactions that are conserved from simpler to more complex community composition (i.e., non-additivity of 2-species interaction fitness effects in 3 and 4-species conditions). This form of HOIs is consistent with a more quantitative definition of HOIs, highly similar to the definition of epistasis in population genetics¹⁸, that identify HOIs (or epistasis) as any deviation, for a given quantitative trait, from the prediction of a linear model where only pairwise interactions are included. Carrying out this analysis, we observe that 82% of the conserved interactions follow an additive pattern of conservation from 2-species to 4-species, and that 18% of the conserved 2-species interaction is associated with non-linear models of conservation.

Overall, our work provides a unique illustration of the highly complex reorganization of interaction mechanisms when microbial community complexity changes. This provides a mechanistic explanation of HOIs in microbial communities that is essential for the further global understanding of microbial communities.

Results

Sets of interaction-associated mutants change across interactive conditions. To investigate how microbial interactions are reorganized in a microbial community with increasing complexity, we reconstructed in vitro a modified bloomy rind cheese-associated microbiome on Cheese Curd Agar plates (CCA plates) as described in our previous work¹⁴. Growth as a biofilm on agar plates models the surface-associated growth of these communities, and allows inclusion of the filamentous fungus, *P. camemberti*, which grows poorly in shaken liquid culture. The original community is composed of the gamma-proteobacterium *H. alvei*, the yeast *G. candidum* and the mold *P. camemberti*. Using a barcoded transposon library of the model bacterium *E. coli* as a probe to identify interactions, we investigated microbial interactions in 2-species cultures (*E. coli* + 1 community member), in 3-species cultures (*E. coli* + 2 community members) and in 4-species cultures (or whole community: *E. coli* + 3 community members) (Fig. 1a).

Quantification of species' final CFUs after 3 days of growth highlighted consistent growth for *H. alvei* and *G. candidum* independent of the culture condition and slightly reduced growth for *E. coli* in interactive conditions compared to growth alone (Dunnett's test against growth alone; adjusted-*p* value $\leq 5\%$) except for the 2-species growth with *P. camemberti* (Supplementary Fig. 1). Although we were unable to quantify spores of *P. camemberti* after three days, growth of *P. camemberti* was visually evident in all of the expected samples. Quantitative analysis of *E. coli*'s library final growth using an epistatic model highlighted that the growth of *E. coli* in the 3-species and 4-species condition can be predicted from the corresponding 2-species growths (Supplementary Fig. 1).

Previously, we developed an assay and a pipeline to identify microbial genes associated with interactions by adapting the original RB-TnSeq approach¹⁹ to allow for consistent implementation of biological replicates as well as for direct quantitative comparison of fitness values between different culture conditions¹⁵. More specifically, the original RB-TnSeq assay relies on the use of a dense pooled library of randomly barcoded transposon mutants of a given microorganism (RB-TnSeq library)¹⁹ containing multiple insertion mutants for each gene as well as intergenic insertion mutants. Measuring the variation of the abundance of each transposon mutant before and after growth, the pipeline allows the calculation of a fitness value for each insertion-mutant as well as a fitness value for each gene corresponding to the average of the insertion-mutants' fitness of the associated genes across biological replicates. A negative fitness indicates that disruption of this gene decreases growth of the mutant relative to a wild type strain, whereas a positive fitness value indicates increased growth in the studied condition. Then, we infer the interactions based on the effects of insertion-mutants between interactive growth and growth alone. In other words, we measure and compare gene fitness across the different studied conditions. Any significant change in fitness values identifies an interaction-associated mutant. The subsequent analysis of interactions, including the inference of the interaction mechanisms and the comparison of interactions across the different interactive conditions, is mainly based on the nature of the disrupted genes by the transposon and their characterized function. Also, by measuring interactions as the difference of fitness value of a given gene between growth with other species and growth alone, we consider that interactions between insertion-mutants of the RB-TnSeq library are controlled and included in our calculation. Then, any interaction-associated mutant predominantly identifies inter-species interactions.

In this work, we used the *E. coli* RB-TnSeq Keio_ML9 library¹⁹ and grew it for 3 days alone or in the seven different interactive conditions studied here (Fig. 1a). This library contains 152,018 pooled insertion mutants with an average of 16 individual insertion mutants per gene and many intergenic insertion mutants. For each interactive condition, we calculated the Interaction Fitness Effect (IFE) associated with 3699 *E. coli* genes as the difference between the gene fitness in the studied interactive condition and the gene fitness in growth alone (Supplementary Data 1). Negative IFE occurs when gene fitness decreases in the interactive condition, and positive IFE occurs when gene fitness improves in the interactive condition. We then tested for all the IFEs that are significantly different from 0 (adjusted *p*-value ≤ 0.1 ; two-sided *t*-test and Benjamini–Hochberg correction for multiple comparison²⁰) to screen for interactions and to identify, in each condition, the insertion-mutants that are associated with inter-species interactions. Here, we identified between 6 (with *P. camemberti*) and 71 (with *H. alvei* + *P. camemberti*) significant IFEs per condition (Fig. 1b). Both negative IFEs and positive IFEs were found in each interactive condition except for the 2-species culture with *P. camemberti*, where only negative interactions were identified. A total of 330 significant IFEs associated with 218 unique genes were identified (as the same gene can be associated with a significant IFE in multiple conditions) including 125 genes associated with negative IFE and 120 genes associated with positive IFE (Supplementary Figs. 2, 3). Altogether, we didn't notice any strong correlation between the number and type of IFE identified by condition and the overall growth impact measured on *E. coli*.

To gain insight into the interaction mechanisms among microbes, we next analyzed the functions of the genes of the interaction-associated mutants (i.e., genes associated with a significant IFE). Here, the vast majority of the genes associated with interaction-associated mutants are part of an interaction network (Fig. 1c). These STRING networks connect genes that code for proteins that have been shown or are predicted to contribute to a shared function, with or without having to form a complex²¹. A significant fraction of the interaction-associated mutants associated with a negative IFE are part of amino acid biosynthesis and transport (17%—Fig. 1c and Supplementary Figs. 2, 4), and more specifically with histidine, tryptophan and arginine biosynthesis. This points to competition for these nutrients between *E. coli* and the other species. Another large set of interaction-associated mutants is related to nucleotide metabolism and transport (14%—Fig. 1c and Supplementary Figs. 2, 5), highlighting competitive interactions for nucleotides and/or their precursors. The majority of the associated genes relate to purine nucleotides and more specifically to the initial steps of their de novo biosynthesis associated with the biosynthesis of 5-aminoimidazole monophosphate (IMP) ribonucleotide. Of the genes associated with interaction-mutants with a positive IFE, 15% are related to amino acid biosynthesis and transport (Fig. 1c and Supplementary Figs. 3, 4), suggesting cross feeding of amino acids between *E. coli* and the other species.

More specifically, this includes phosphoserine, serine, homoserine, threonine, proline and arginine. The presence of amino acid biosynthetic genes among both negative and positive IFEs indicate that trophic interactions (competition versus cross-feeding) depend on the type of amino-acid and/or the species interacting with *E. coli*. For both negative and positive IFEs, numerous genes of the associated interaction-mutants were annotated as transcriptional regulators (Fig. 1c and Supplementary Figs. 2, 3) emphasizing the importance of transcriptional reprogramming in response to interactions. These transcriptional regulators include metabolism regulators as well as regulators of growth, cell cycle and response to stress. Finally, these interaction-associated mutants and the inferred interaction mechanisms are consistent with previous findings in this microbiome¹⁴ as well as in a study of bacterial-fungal interactions involving *E. coli* and cheese rind isolated fungal species¹⁵. While this approach allows us to infer the interaction mechanisms that are happening between the transposon library and the other species, further experimental validation would be needed to confirm that these interactions more generally happen between a WT strain and the other species.

Introduction of a third interacting species deeply reshapes microbial interactions. The differences in the number and sign of significant IFEs observed among the different interactive conditions, with different numbers of interaction species, suggest that the number and type of interacting partners influence interaction mechanisms. To characterize how the interactions are reorganized with community complexity, we then investigated if and how the genetic basis of interactions changes when the number of interacting partners increases by comparing the genes associated with interaction-associated mutants with significant IFE in 2-species cultures, in 3-species cultures and then in 4-species cultures.

First, we have identified 104 IFEs associated with 98 genes in 2-species cultures as well as 168 IFEs associated with 136 unique genes in 3-species conditions (Supplementary Fig. 6 and Supplementary Data 2). Comparing these gene sets, we can identify how the interaction-associated mutants change when a third-species is added to a 2-species culture. We identified 45 genes associated with 2-species interaction-associated mutants maintained in at least one 3-species condition (maintained interaction-mutants), 55 genes associated with 2-species interaction-associated mutants no longer associated with interaction in any 3-species condition (dropped interaction-mutants) and 100 genes associated with 3-species interaction-associated mutants that aren't related to any 2-species interaction-associated mutants (emergent interaction-mutants) (Fig. 2a, Supplementary Fig. 6 and Supplementary Data 3). Both dropped and emerging interaction-associated mutants represent 3-species HOIs; the third species either removes an existing interaction or brings about a new one.

We further carried out functional analysis of the genes related to maintained, dropped and emerging interaction-mutants to elucidate whether maintained and HOIs interaction-mutants would be associated with specific functions and thus interaction mechanisms (Fig. 2b). For each set of genes, we calculated the fraction of genes of that set associated with a given COG ontology category. Metabolism and transport is the most observed COG group (Fig. 2b—teal dots). For genes related to maintained interaction-mutants, this indicates that some trophic interactions can be maintained from 2-species to 3-species conditions. For instance, serine biosynthetic genes *serA*, *serB* and *serC* as well as threonine biosynthetic genes *thrA*, *thrB* and *thrC* are associated with positive IFEs in the 2-species condition with *G. candidum* as well as in the 3-species conditions involving *G. candidum* (Supplementary Fig. 4). This suggests that, (i) *G. candidum* facilitates serine and threonine cross feeding and (ii) this cross-feeding is still observed when another species is introduced. However, metabolism-related genes identified among the dropped and emerging interaction-mutants indicate that many trophic interactions are also rearranged through HOIs. Genes associated with lactate catabolism (*lldP* and *lldD*) and lactate metabolism regulation (*lldR*) have a negative IFE in the 2-species culture with *H. alvei*, suggesting competition for lactate between *E. coli* and *H. alvei*. Yet, mutants of these genes are no longer associated with a significant IFE when at least another partner is introduced (Supplementary Fig. 7). Histidine biosynthesis genes *hisA*, *hisB*, *hisD*, *hisH* and *hisI* are associated with interaction-mutants with negative IFE in the 2-species culture with *H. alvei* and sometimes in the 3 species culture with *H. alvei* + *P. camemberti*. However, the negative IFE is alleviated whenever *G. candidum* is present, suggesting that potential competition for histidine between *E. coli* and *H. alvei* is alleviated by this fungal species (Supplementary Fig. 4). Also, genes related to the COG section “Information storage and processing” are mostly found among genes of HOIs-mutants suggesting a fine-tuning of specific cellular activity depending on the interacting condition. For instance, we identified many transcriptional regulators of central metabolism among the dropped interaction-mutants genes (*rbsR* and *lldR*) and the emerging interaction-mutants genes (*purR*, *puuR*, *gcvR* and *mngR*), highlighting again the reorganization of trophic interactions associated with HOIs. Also, many transcriptional regulators broadly associated with growth control, cell cycle and response to stress were found among the emerging interaction-mutants genes with 3-species (*hyfR*, *chpS*, *sdiA*, *slyA* and *rssB*), underlining a noticeable modification of *E. coli*'s growth environment with 3-species compare to with 2-species.

Finally, we further aimed to understand whether HOIs are associated with the introduction of any specific species (Fig. 2c and Supplementary Fig. 8). We observe that interaction-associated mutants with *H. alvei* are more likely to be dropped, as 65% of them are alleviated by the introduction of a fungal species (Fig. 2c). This can be seen, for instance, with the reorganization of *E. coli* and *H. alvei* trophic interactions following the introduction of *G. candidum* (alleviation of lactate and histidine competition for instance). Also, we observe that 76% of the interactions in the 3-species cultures with *H. alvei* + *P. camemberti* and 65% in the 3-species culture with *H. alvei* + *G. candidum* are emerging interaction-mutants (compared to 38% of emerging interaction-associated mutants in the 3-species condition with *G. candidum* + *P. camemberti*) (Fig. 2c). For the interaction-associated mutants found in the 3-species with *H. alvei* + *P. camemberti*, they include for instance the genes associated with purine de novo biosynthesis (*purR*, *purF*, *purN*, *purE*, *purC*) and the genes associated with pyrimidine de novo biosynthesis (*pyrD*, *pyrF*, *pyrC*, *carA* and *ulaD*), suggesting important trophic HOIs. For the 3-species condition with *H. alvei* + *G. candidum*, emerging interaction-mutants include for example the transcriptional regulator

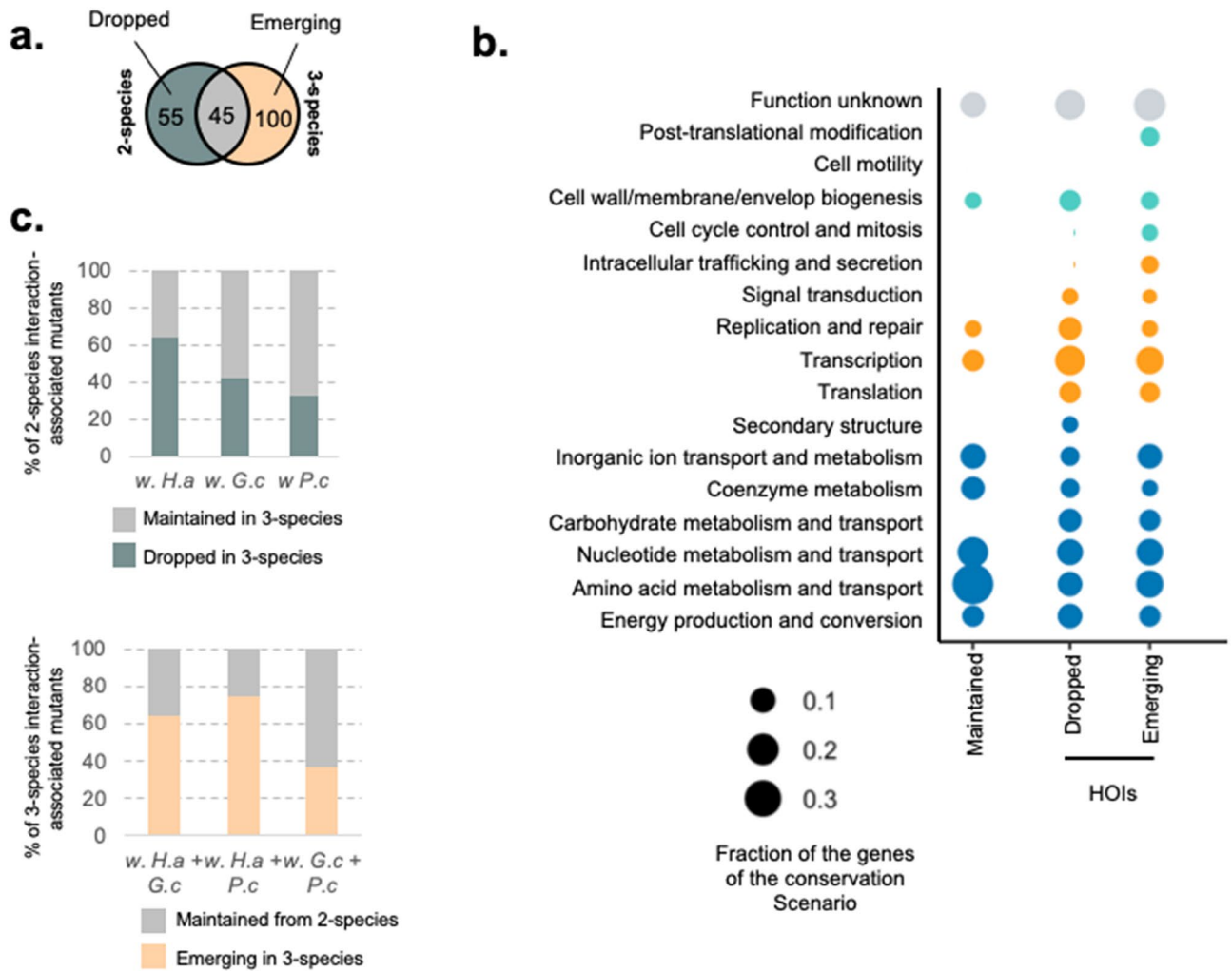


Figure 2. Comparison of the genetic basis of interaction for 2-species and 3-species conditions. **(a)** Venn Diagram of 2-species and 3-species sets of genes related to interaction-associated mutants. This Venn Diagram identifies 2-species interaction-mutants that are dropped when a third species is introduced (Left side; Dropped interaction-mutants = any 2-species gene that is not found in any 3-species condition), 2-species interaction-mutants that are maintained in at least one associated 3-species condition (Intersection; Maintained interaction-mutants) and interaction-mutants that are specific to 3-species condition (Right side; Emerging interaction-mutants). **(b)** Functional analysis of the genes associated with dropped, maintained and emerging interaction-mutants from 2-species to 3-species. Each dot represents the fraction of genes of the studied gene set associated with a given COG category (Number of genes found in the category / Total number of genes in the gene set). The color of the dots indicates the general COG group of the COG category: Teal: Metabolism; Blue: Information storage and processing; Orange: Cellular Processes and Signaling; Grey: Unknown or no COG category. **(c)** Species-level analysis of 3-species HOIs: for each 2-species condition, we measure the fraction of interaction-mutants that are dropped in associated 3-species cultures (Dropped in 3-species) or maintained in at least one of the 3-species cultures (Maintained in 3-species); for each 3-species condition, we measure the fraction of interaction-mutants that have been conserved from at least one associated 2-species condition (Maintained from 2-species) or that are emerging with 3-species (Emerging in 3-species).

genes *chpS*, *sdiA* and *slyA*, indicating the presence of a stress inducing environment. Together, these observations suggest that the introduction of a fungal partner may introduce multiple 3-species HOIs by both canceling existing interactions and introducing new ones.

HOIs are prevalent in a 4-species community. To further decipher whether microbial interactions continue to change with increasing community complexity, we investigated the changes in the genetic basis of interactions going from 3-species to 4-species experiments. We identified 58 interaction-associated mutants in the 4-species condition (*E. coli* with *H. alvei* + *G. candidum* + *P. camemberti*), compared with 145 interaction-associated mutants in any 3-species condition. Comparing the two sets of interaction-associated mutants and corresponding genes we identify: 26 3-species interaction-mutants that are maintained in the 4-species condition (including 16 directly from 2-species interactions), 115 3-species interaction-mutants that are no longer associ-

ated with interactions in the 4-species condition (dropped interaction-mutants) and 32 interaction-mutants that are observed solely in the 4-species condition (emerging interaction-mutants) (Fig. 3a, Supplementary Fig. 6 and Supplementary Data 3). Both dropped and emerging interaction-mutants represent 4-species HOIs. Here, HOIs are remarkably abundant when introducing a single new species and moving up from 3-species interactions to 4-species interactions. Functional analysis of the genes of maintained-mutants and HOI-mutants reveals the presence of many metabolism related genes in every gene set (Fig. 3), suggesting that some trophic interactions can be maintained from 3-species to 4-species interactions while some other trophic interactions are rearranged with HOIs. For instance, most of the genes of the initial steps of de novo purine biosynthesis have been found to be associated with a negative IFE in the 3 species condition with *H. alvei* + *P. camemberti* (*purC*, *purE*, *purF*, *purL* and *purN*) as well as in the pairwise condition with *H. alvei* for *purH* and *purK* (Supplementary Fig. 5), suggesting competition for purine initial precursor IMP in these conditions. Yet, the introduction of the yeast *G. candidum* as a fourth species cancels the negative IFE value, suggesting that the competition is no longer happening in its presence. Altogether, the observation of noticeable trophic HOIs moving up from 2 to 3 species and then from 3 to 4-species interaction highlights a consistent reorganization of trophic interactions along with community complexity. Also, genes related to Cell wall/membrane/envelope biogenesis are found abundantly among the 4-species emerging-mutants (Fig. 3b) and they represent the largest functional fraction of this gene set. These genes are associated with a negative IFE and are related to Enterobacterial Common Antigen (ECA) biosynthetic processes (*wecG*, *wecB* and *wecA*) (Supplementary Fig. 9). While the roles of ECA can be multiple but are not well defined²², they have been shown to be important for response to different toxic stress, suggesting the development of a specific stress in the presence of the four species.

As for the 2 to 3 species comparison, we investigated whether the introduction of a specific fourth species would be most likely associated with HOIs. The 3-species culture that appears to be the least affected by the introduction of a fourth member is with *G. candidum* + *P. camemberti* where 34% of the observed interactions are still conserved in the 4-species condition after the introduction of *H. alvei* (versus 22% for with *H. alvei* + *G. candidum* when *P. camemberti* is added and 21% for with *H. alvei* + *P. camemberti* when *G. candidum* is added) (Fig. 3c and Supplementary Fig. 10). Together, these observations suggest that, again, the introduction of a fungal partner may introduce multiple 4-species HOIs.

Finally, by increasing the number of interacting species in our system and investigating interaction-mutants maintenance and modification with every increment of community complexity, we are able to build our understanding of the architecture of interactions in a microbial community. Altogether, we have observed a total of 218 individual interaction-associated mutants in any experiment. Only 16 of them (7%) were conserved across all levels of community complexity (Fig. 3d). Starting from 2-species interaction-mutants, 48% of them were maintained with 3-species and only 15% (16 out of 104) were still maintained with 4-species. Thus, we demonstrate here a progressive loss and replacement of 2-species interactions as community complexity increases and the prevalent apparition of HOIs. Tracking back the origins of the genetic basis of interactions in the 4-species experiment that represents the full community of our model, we identify that 28% of the full community interactions can be traced back to 2-species interactions, 18% are from 3-species interaction and 54% are specific to the 4-species interaction (Fig. 3d,e). Most of the maintained interaction-mutants from 2-species as well as from 3-species are associated with metabolism (Fig. 3d and Supplementary Fig. 11) while Signal transduction and cell membrane biosynthesis genes are most abundant among the 4-species interaction-mutants as previously mentioned. To conclude, this shows that the genetic basis of interactions and thus the sets of microbial interaction are deeply reprogrammed at every level of community complexity and illustrates the prevalence of higher order interactions (HOIs) even in simple communities.

The majority of maintained 2-species interaction-mutants in the 4-species culture follows an additive conservation behavior.

While HOIs are abundant in the 4-species condition, our data yet suggest that up to 28% of the interactions are maintained from 2-species interactions. However, we don't know whether and how 2-species interactions are quantitatively affected by the introduction of other species and whether they would follow specific quantitative models of conservation. For instance, we can wonder how the strength of a given 2-species interaction is modified by the introduction of one or two other species, or how two 2-species interactions associated with the same gene will combine when all the species are present. In other words, can we treat species interactions as additive when we add multiple species? Such information would generate a deeper mechanistic understanding of the architecture of microbial interactions while allowing us to potentially predict some whole community interactions from 2-species interactions. Here, two main hypothetical scenarios can be anticipated. First, the conservation of 2-species interactions follows a linear or additive behavior, where the introduction of other species either doesn't affect the strength of the conserved 2-species interaction or two similar 2-species interactions combine additively. The second scenario identifies non-linear or non-additive conservation of 2-species interactions, where the strength of the conserved 2-species interaction is modified by the introduction of other species or two similar 2-species interactions are not additive. The second scenario would encompass for instance synergistic effects or inhibitory effects following the introduction of more species. We next use an epistasis and quantitative genomics approach to understand whether interactions that are conserved follow a linear, or additive, pattern. For the 16 interaction-associated mutants that are associated with interaction in 2-species cultures, in associated 3-species cultures and in the 4-species condition, we use epistasis analysis to test the linear behavior of their IFE when the number of interacting species increases, as IFEs are quantitative traits related to the interaction strength. In multi-dimensional systems, an epistasis analysis quantifies the additive (or linear) behavior of conserved quantitative traits. In quantitative genetics, for instance, epistasis measures the quantitative difference in the effects of mutations introduced individually versus together^{18,23,24}. Using a similar rationale, we can use IFEs as a quantitative proxy for interaction strength and test

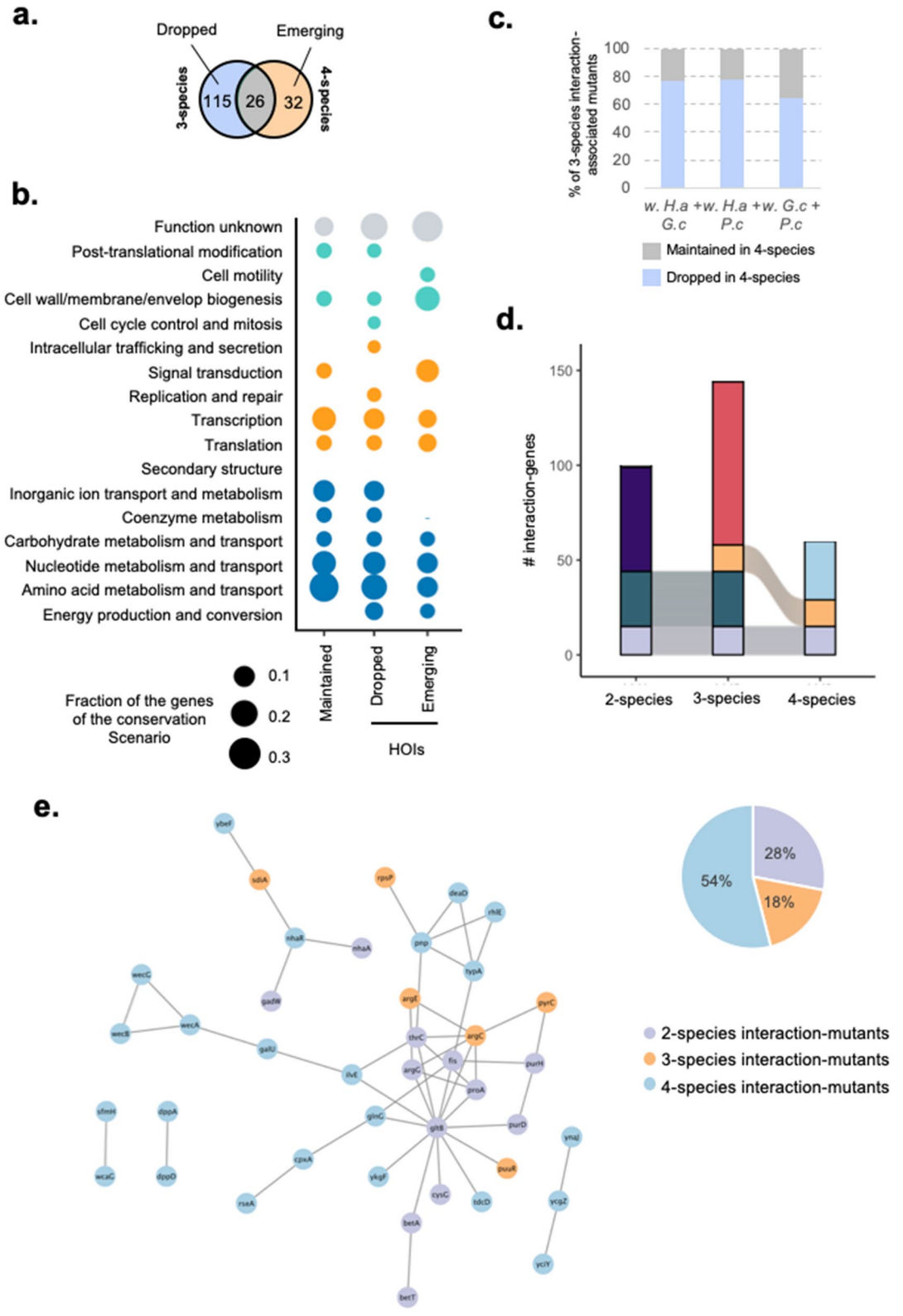


Figure 3. Organization of the interactions in the 4-species community. **(a)** Venn Diagram of 3-species and 4-species sets of genes related to interaction-associated mutants. This Venn Diagram identifies 3-species interaction-mutants that are dropped when a fourth species is introduced (Left side; Dropped interaction-mutants = any 3-species interaction-associated mutant that is not found in the 4-species condition), 3-species interaction-mutants that are maintained in the 4-species condition (Intersection; Maintained interaction-mutants) and interaction-mutants that are specific to 4-species condition (Right side; Emerging interaction-mutants). **(b)** Functional analysis of the genes associated with dropped, maintained and emerging interaction-mutants from 3-species to 4-species. Each dot represents the fraction of genes of the studied gene set associated with a given COG category (Number of genes found in the category/Total number of genes in the gene set). The color of the dots indicates the general COG group of the COG category: Teal: Metabolism; Blue: Information storage and processing; Orange: Cellular Processes and Signaling; Grey: Unknown or no COG category. **(c)** Species-level analysis of 4-species HOIs: for each 3-species cultures we measure the fraction of interaction-genes that is conserved in the 4-species culture (Maintained in 4-species) and the fraction of interaction-genes that has been dropped (Dropped in 4-species). **(d)** Alluvial plots of the interaction genes across community complexity levels. **(e)** STRING network of the 4-species interaction genes (Nodes). Edges connecting the genes represent both functional and physical protein association and the thickness of the edges indicates the strength of data support (minimum required interaction score: 0.4—medium confidence). Nodes are colored based on the level of community complexity the genes are conserved from.

whether the IFEs of the maintained interaction genes in 3-species and in 4-species conditions result from the linear combination of associated 2-species IFEs (Fig. 4a). Nonlinear combination, or non-additivity of 2-species IFEs in higher community level also highlights higher-order interactions.

We adapted the pipeline Epistasis¹⁷, originally designed for quantitative genetics investigation. We implemented the linear model with the gene fitness values of the interaction-associated mutants for growth alone, for each of the 2-species conditions, for each of the 3-species cultures and for the 4-species condition. For each gene, the software finds the simplest mathematical model that reproduces the observed IFEs across all levels of community complexity. In the simplest case, the model will have a term describing the effects for adding each species individually to the *E. coli* alone culture; that term corresponds to the 2-species IFE. Then, if the IFE for two *E. coli*'s partners combined (3-species IFE) differs from the sum of their individual effects (corresponding 2-species IFE), the software adds a term capturing this epistasis (Fig. 4a). Here, we call that term 3-species epistatic coefficient or ϵ_{ij} . Finally, if the IFE for the combined community (*E. coli* plus all three species; 4-species condition) differs from the prediction based on the 2-species and 3-species terms, the software will add a high-order interaction term to the model (Fig. 4b). Here, we name that term 4-species epistatic coefficient or ϵ_{ijk} .

We performed this analysis on the 16 interaction-associated mutants that are associated with interactions at every level of community complexity. To identify real additive behavior of IFE from non-additivity, we screen for 3-species epistatic coefficients and 4-species epistatic coefficients that are significantly different from 0 (adjusted p-value ≤ 0.01 , Benjamini–Hochberg correction for multiple testing). We found that 13 interaction-associated mutants behaved additively from 2-species to 4-species culture, with no epistatic contributions in the 3-species conditions nor in the 4-species condition (Fig. 4c, (i)). One interaction-associated mutant (gene (*gadW*)) exhibited nonlinear conservation of IFE only in the 4-species condition, but additive IFE conservation from 2-species to 3-species (Fig. 4c, (ii)). Another interaction-associated mutant (gene (*lsrG*)) showed epistasis in one 3-species condition but no epistasis in the 4-species condition (Fig. 4c, (iii)) Finally, one interaction-associated mutant (gene (*gltB*)) displayed both non-additivity in 3-species and 4-species conditions (Fig. 4c, (iv)). If we look more closely at the genes related to interaction-associated mutant with an additive behavior, we find genes (*betA*, *betT*, *purD* and *purH*) that are associated with the conservation of negative IFEs (Supplementary Fig. 12). While *betA* and *betT* are associated with choline transport (*betT*) and glycine betaine biosynthesis from choline (*betA*)²⁵, *purD* and *purH* are associated with de novo purine biosynthesis²⁶. This suggests that requirements for glycine betaine biosynthesis from choline and for purine biosynthesis caused by microbial interactions, possibly due to competition for the nutrients used as precursors, are additively conserved from individual 2-species interactions requirements. Also, 5 genes associated with amino acid biosynthesis (*serA*, *thrC*, *cysG*, *argG* and *proA*) are associated with the additive conservation of positive IFE (Supplementary Fig. 12), suggesting that cross feeding can be additive when the community complexity increases. Altogether, this highlights the existence of 2-species interactions, including trophic ones, conserved in an additive fashion in the highest-level of complexity.

This leaves 3 interaction-associated mutants (18%) of the maintained 2-species interaction-mutants, that are associated with non-additive behavior, and thus HOIs, at at least one higher level of community complexity (Fig. 4c—(ii), (iii) and (iv)). The interaction-associated mutant for the gene *gadW* is associated with non-additivity at the 4-species level, suggesting that while IFEs are additive in 3-species cultures, the introduction of a fourth species introduces HOI. Moreover, the observed 4-species IFE is greater than the IFE predicted by a linear model (Fig. 4d), highlighting a potential synergistic effect when the 4 species are together. The interaction-associated mutant for the gene *lsrG* is associated with non-additivity only at the 3-species culture w *G.c* + *P.c*. More specifically, this indicates that HOI arise when these 2 fungal species are interacting together with *E. coli*, but that no more HOI emerge when *H. alvei* is introduced (i.e., the 4-species IFE can be predicted by the linear combination of the lower levels IFEs). As the observed IFE for the 3-species condition w *G.c* + *P.c* is greater than the predicted IFE (Fig. 4c), this suggests a synergistic effect between the 2 fungal species. Finally, the interaction-associated mutants for the gene *gltB* is associated with non-additivity at both the 3-species and 4-species levels. For this interaction-associated mutant, the conservation of IFE is never associated with an additive model. Here, the observed 4-species IFE is not as negative as it would be as the result of the linear combination of the associated lower IFE (Fig. 4d), suggesting the existence of a possible IFE threshold, or plateau effect. Altogether, this

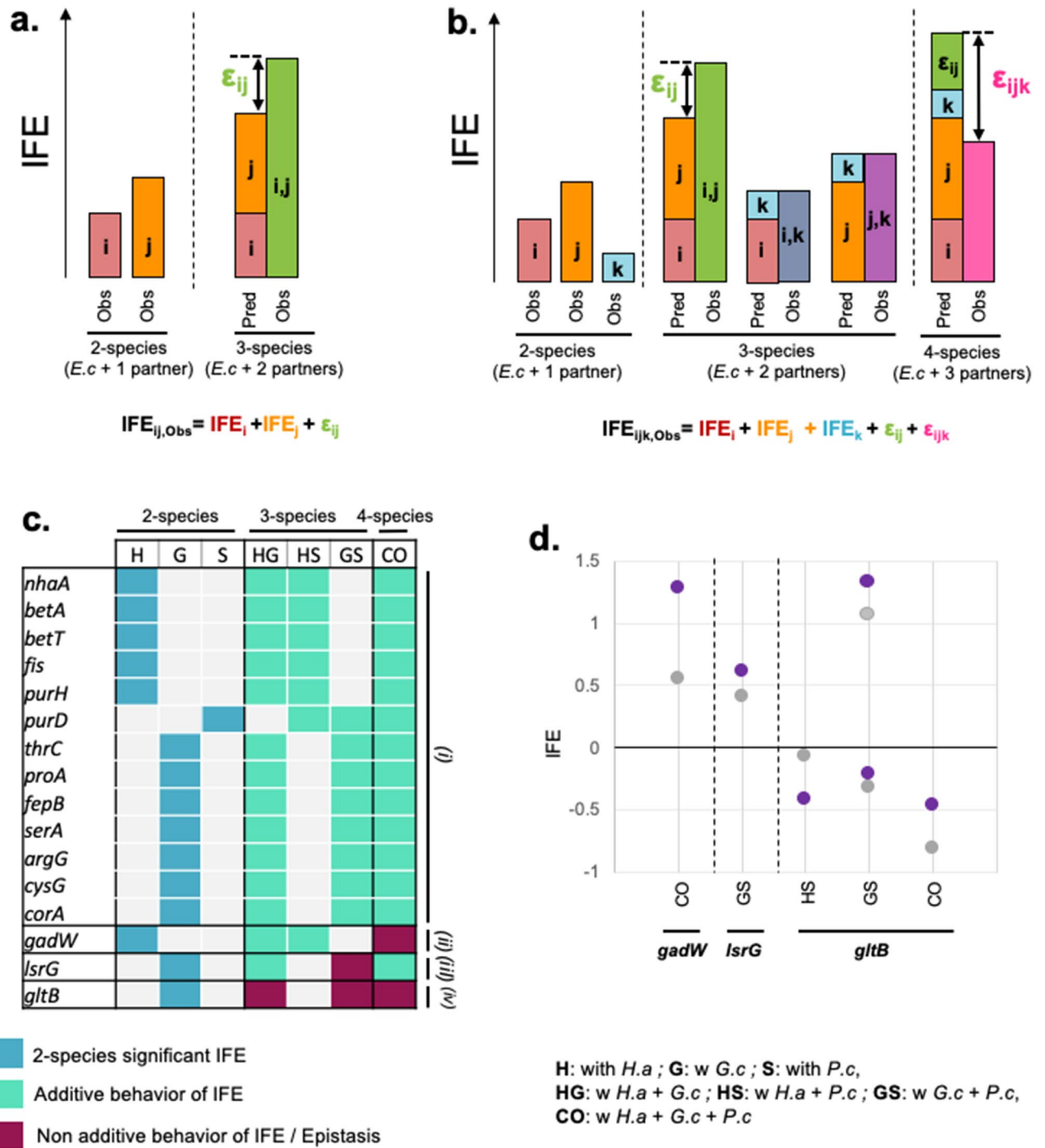


Figure 4. Quantitative analysis of IFE conservation for the interaction-associated mutants conserved from 2-species to 4-species conditions. **(a)** Schematized quantitative epistasis/non-linearity measured in 3-species conditions (with partner *i* and *j*). Epistasis (ϵ_{ij}) is the difference between the individual IFE of partner *i* and partner *j* (red and orange bars) versus placing them together (green). Mathematically, we need three terms (IFE_i , IFE_j , and ϵ_{ij}) to reproduce the observed IFE for the 3-species condition. **(b)** This analysis can be extended to higher levels of community complexity: 4-species (*E. coli* with 3-partners *i*, *j*, and *k*). The model first accounts for epistasis between *i/j*, *i/k*, and *j/k*. In this example, *i* and *j* exhibit epistasis; *i/k* and *j/k* are additive (dark blue and purple). The predicted IFE for the 4-species community is the sum of the individual 2-species effects (red, orange, light blue) and the 3-species epistatic terms (green). The 4-species epistatic coefficient is the difference between this low-order prediction and the observed IFE for the *i,j,k* community (pink). **(c)** Conservation profiles of the 16 2-species interaction-associated mutants conserved up to 4-species. 2-species conditions: a colored square indicates the 2-species condition(s) in which the interaction-associated mutant was identified; a grey square indicates non-significant 2-species IFEs. 3-species conditions: a teal square indicates that the associated IFE is associated with additive behavior from associated 2-species IFE (no ϵ_{ij} epistatic coefficient), a red square indicates that the associated IFE displays non-additivity from 2-species IFE and thus epistasis, a grey square corresponds to a 3-species condition that is not associated with significant 2-species IFE (no epistasis analysis performed); 4-species condition: a teal square indicates that the associated IFE is associated with additive behavior (no ϵ_{ijk} epistatic coefficient), a red square indicates that the associated IFE is associated with non-additivity from lower-order IFE. **(d)** Comparison of the observed and predicted IFE for the genes and condition associated with 3-species and 4-species non-additive IFE.

indicates that maintained 2-species-interactions can follow nonlinear behaviors that could involve synergistic effects, inhibitory effects or constraints.

Discussion

Interactions between microbes are responsible for the specific and multiple phenotypes observed in microbial communities compared to monocultures. Knowledge of microbial interactions could be the key to controlling microbial communities, but deciphering these interactions is challenging in complex microbiomes. Moreover, pairwise culture interactions are often insufficient to predict what happens in more complex microbiomes, suggesting an important reprogramming of interactions as community complexity increases in terms of the number of species present^{3,5,7,13,14}. Understanding the restructuring of these interactions in complex communities is thus essential to comprehending the biology of microbial communities. Using an in vitro multi-kingdom community, we performed a molecular investigation of the reorganization of interaction profiles as community complexity changed. Relying on the tractability of our model system and an RB-TnSeq-based interaction assay, we tracked interactions in 2-species, 3-species and 4-species cultures. In this work, the combination of a qualitative and quantitative comparison of interaction profiles at the molecular level underlines the complex dynamics of interaction reorganization with community complexity and the existence of multiple forms of higher-order interactions. While this work has been conducted on solid medium, we would expect similar observations in liquid batch cultures given the nature of the inferred interactions. For example, amino acid cross-feeding would likely occur both in liquid and solid media.

This work offers an example of the different forms that HOIs can take in biological systems. We report multiple mechanistic HOIs represented here by any pairwise interaction-associated mutants that are not observed in 3 and more species conditions as well as any interaction-associated mutant observed in higher-levels than pairwise cultures¹². We also report another form of HOIs, as defined in quantitative genetics, which are associated with the non-additive behavior of conserved 2-species interactions in 3 or 4 species communities²³. Here, each HOI level refers to different biological phenomena occurring in the same biological system. Yet, they are essential and complementary to decipher the extremely convoluted architecture and dynamics of microbial interactions and microbiome biology. It is important to point out here, that these HOIs are measured for individual mutants that are part of a larger pool of a total of 152,018 mutants and they only represent a small fraction of the full population. Thus, it is possible that such HOIs don't translate at the overall community growth level, and that community growth could still be predicted from pairwise observations. However, these HOIs could underlie other community and emerging traits we haven't measured in this present work.

As the number of interaction-associated mutants strongly decreased in the 4-species culture compared to 2 or 3 species setups, our work points out a strong reduction of the interaction landscape with 4 species that is associated with the loss of many lower-level interactions and the emergence of context-specific interactions. While 43% of the 2-species interaction-associated mutants are still found with 3 species, only 15% are still found with 4 species, representing less than a third (28%) of the total of 4-species interactions. This highlights the increasing dilution and replacement of original 2-species interactions. To summarize, the more complex a community gets, the more mechanistic HOIs emerge. While our work was limited to 4 species, it remains necessary to verify whether this statement will be true for more complex communities: whether more 2-species interactions will be lost and more HOIs will keep emerging at each level or whether the interaction landscape will stabilize. Indeed, Friedman et al. have highlighted that growth observation for 2- and 3-species combination could predict the assembly of 7- to 8-species²; this would suggest that interactions, or at least key interactions driving community assembly, could change less dramatically in higher complexity microbiomes.

In a sense, the *E. coli* library is allowing us insight into the environment created by the community and allows us to further hypothesize about microbial interactions profiles (for instance the release of amino acids by certain species that can potentially be used for cross-feeding). Yet this does not necessarily reflect what would happen with the WT strain, as specific follow-up experiments would have to be conducted for each hypothesized mechanism. Nevertheless, our molecular approach enabled us to identify that most of the reorganization of the interaction profile is associated with the reprogramming of metabolic or trophic interactions including both competition for nutrients and cross-feeding. As more species are introduced it appears that the dynamics of nutrient consumption is rearranged. For instance, we observed that some 2-species competition for amino acids and for lactate between *E. coli* and *H. alvei* can be alleviated by the introduction of *G. candidum*. While the competition for amino acids is likely alleviated by amino acid cross-feeding from *G. candidum*, as this species is known to release amino acids into the environment through digestion of proteins and peptides^{27–29}, the mechanism relieving competition for lactate is unclear. Possibly, *G. candidum* provides other nutrients sources like amino acids that alleviate the need for *E. coli* to rely on lactate. Trophic interactions are described as core determinants in community assembly^{30,31} and simple rules of metabolic interdependencies and metabolic specialization can be sufficient to predict the assembly of rather complex communities^{32–34}. Yet in complex systems, such as the gut microbiome, with complex nutrient composition, the important reorganization of metabolic interactions as community complexity changes could likely explain the difficulty in predicting community assembly and composition in some studies⁷. As previously suggested³⁵, ecological and metabolic factors associated with the present species such as niche overlap, degree of metabolic specialization and species similarities are likely to be drivers of these metabolic HOIs. Indeed, microorganisms display incredible metabolic abilities, from nutrient usage to rapid metabolic switching, and in the presence of multiple nutrient sources and/or other microorganisms they are likely to readjust the sequence of nutrient uptake. Niche occupation and nutrient access are also two crucial aspects in the success or failure of microbial invasion^{4,36}. The reorganization of metabolic interactions in different community composition could likely explain the poor predictability of invasion resistance from simple species-combinations and the emergence resistance-specific phenotype.

To some extent, our work also highlights the importance of fungal species as major actors in reshaping interaction networks. Here, more 3-species HOIs and 4-species HOIs were observed when one fungal species was added either as a third species with two bacteria or as a fourth species with another fungus and two bacteria. In microbial communities, fungi are known to impact community structure and access to nutrients through the formation of hyphal highways^{37,38}, to impact community assembly through environmental modification^{39–41} as well as to impact community protection through the production of multiple secondary metabolites with antibiotic or antimicrobial properties^{15,40,42}. In this specific context, fungi-associated HOIs seem to be related to their metabolic characteristics, whether through cross-feeding or competition for nutrients. Yet, the emergence of specific requirements for ECA biosynthetic genes in *E. coli* when all species are present, suggesting a potential toxic stress, may reveal other fungal context-specific properties, while it could also inform about a more precise role of the ECA in *E. coli*. We believe this work contributes, along with other recent studies, to advocate for the need to include fungal species more frequently into microbiome work studies.

Finally, to understand principles behind the conservation of pairwise interactions, we used epistasis analysis to quantify the additive or non-additive behavior of conserved 2-species interactions in 3 and 4 species conditions. Epistasis analysis offers an adaptable approach to test the linear behavior of quantitative traits in systems with increasing complexity, whether it is the effects of accumulation of mutations in a genome^{18,43}, the effect of drug combinations⁴⁴ or the dynamics of predators-prey ecosystems⁴⁵. Recent studies in the field of microbiome research have also relied on epistasis analysis to elucidate fundamentals of microbial communities such as how the nutrient composition of the environment determines the assembly and the diversity of a microbial community^{35,46}, how the functional landscape of microbial community is built³ or how the commensal microbiome determines the development, lifespan and reproduction of its host⁵. In this work, using the quantitative metric IFE for the strength of interactions, we used an epistatic model to characterize the additivity or non-linearity of IFE of conserved 2-species interactions when *E. coli* is growing with 2 and 3 other species (3-species and 4-species conditions respectively). We observed that most of the conserved 2-species interaction genes followed an additive model of conservation, including trophic-interaction related genes. While our study was limited to a small number of interacting species, it remains to be investigated if this linear behavior is propagated if the community complexity increases again. While we have observed both synergistic effects and limitations effects for nonlinear conservation of interaction strength, we believe that such behavior are likely to arise in more complex contexts and that additivity will eventually stop. For instance, additivity of pairwise interactions is likely to saturate and to stop at higher-levels of complexity due to environmental constraints (nutrient supply for instance) and/or the metabolic capacities of the different species as highlighted in³⁵. While this theory can apply to conserved trophic interactions, like the ones highlighted in our work, other scenarios could include molecule-driven or environment-mediated interactions such as exchange of metabolites, quorum sensing, anti-biosis and modification of the physiochemical properties of the environment.

While trait additivity is fairly easy to quantify in high-dimensional systems, our study, along with other work using epistasis to study microbial communities^{3,5,35,46}, suggests that additivity of quantitative microbial features is not the only scenario in nature, and strongly emphasizes the need to develop mathematical approaches to accurately understand the dynamics and functionalities of microbiomes.

Overall, our work identifies the reorganization of microbial interactions along with community complexity at the molecular level. While more advanced modeling is still required, this knowledge of the mechanistic reorganization of interactions and patterns of HOIs is essential for bottom-up approaches to be able predict, from minimal information, the biology of complex microbiomes and will offer novel opportunities to design synthetic microbial communities or control natural ones.

Methods

Strains and media. *Strains.* The bloomy rind cheese community was reconstructed with the same strains from¹⁴: *H. alvei* JB232 isolated from cheese⁴⁷ and two industrial cheese strains: *G. candidum* (Geotrichum candidum GEO13 LYO 2D, Danisco—CHOOZITM, Copenhagen, Denmark) and *P. camemberti* (PC SAM 3 LYO 10D, Danisco—CHOOZITM).

Medium. All assays have been carried out on 10% cheese curd agar, pH7 (CCA) (10% freeze-dried Bayley Hazen Blue cheese curd (Jasper Hill Farm, VT), 3% NaCl, 0.5% xanthan gum and 1.7% agar). The pH of the CCA was buffered from 5.5 to 7 using 10 M NaOH.

Competition assay—RB-TnSeq assays. The *E. coli* barcoded transposon library Keio_ML9¹⁹ was used for all RB-TnSeq assays on CCA during a 3-day growth in eight different culture conditions: alone, 2-species conditions (with *H. alvei*; with *G. candidum*; and with *P. camemberti*), 3-species conditions (with *H. alvei* + *G. candidum*; with *H. alvei* + *P. camemberti*; and with *G. candidum* + *P. camemberti*), and in the single 4-species condition (with *H. alvei* + *G. candidum* + *P. camemberti*) (Fig. 1).

From the species and library inoculation to cell harvest after 3 days of growth, we followed the same procedure as described in¹⁴. We amplified the *E. coli* barcoded transposon library Keio_ML9 into 25 mL of liquid LB-kanamycin (50 mg/mL) up to an OD of 0.6–0.8. A 5 mL sample of this preculture was spun down and kept at –80°C as the T0 sample required for fitness calculation. The remaining cells were then washed in PBS1x-Tween0.05% and used to inoculate the competition assays with 7×10^6 cells of the library on each 100 mm petri dish plate. When necessary, the other community members were then inoculated at the following densities: for *H. alvei*: 7×10^6 cells; for *G. candidum*: 7×10^6 cells; for *P. camemberti*: 7×10^5 cells.

After 3 days, cells were harvested by flooding the plates with 1.5 mL of PBS1X-Tween0.05%, gentle scraping and then transferring the resuspended cells into 1.5 mL tubes. Cells were centrifuged for 3 min at room temperature (RT) at 10 KRPM and stored at -80°C until gDNA extraction.

All assays have been performed in triplicate.

Epistasis analysis of *E. coli*'s growth. The Python package Epistasis¹⁷ was used to run the epistasis analysis on *E. coli*'s final CFUs in 3 and 4-species conditions. The model was implemented with the average final CFUs across three biological replicates along with the corresponding standard deviation for each of the culture conditions (Alone, w *H.a*, w *G.c*, w *P.c*, w *H.a* + *G.c*, w *H.a* + *P.c*, w *G.c* + *P.c*, w *H.a* + *G.c* + *P.c*). The growth conditions correspond to the 'genotypes' expected in the model and they are binary coded based on the presence or not of *E. coli*'s partners *G. candidum*, *H. alvei* and *P. camemberti*. The final CFU counts corresponds to the 'phenotype' implemented in the model. The model was run using the 'local' parameter. 7,000 pseudoreplicates were generated to determine whether 3-species and 4-specie epistatic coefficient were significant and P-values were then adjusted for multiple comparison testing using Benjamini Hochberg Correction²⁰.

gDNA extraction, library preparation and sequencing. gDNA from the samples of the competition assays was extracted by phenol-chloroform extraction (pH 8) as described in¹⁴. For each sample, we resuspended the cells into 500 mL of buffer B (200 mM NaCl, 20 mM EDTA) and then transferred them into a 2 mL screw-capped tube previously filled with 125 mL of 425–600 mm acid-washed beads and 125 mL of 150–212 mm acid-washed beads. Then, 210 mL of SDS 20% and 500 mL of Phenol:Chloroform (pH 8) were added to each sample before mechanical lysis by vortexing the tubes for 2 min at maximum speed. Tubes were then centrifuged for 3 min at 8 KRPM at 4 °C and 450 mL of aqueous phase were recovered for each sample. 45 mL of sodium acetate 3 M and 450 mL of ice-cold isopropanol were added and tubes were incubated for 10 min at -80 °C. Tubes were then centrifuged for 5 min at 4 °C at 13 KRPM, supernatant was removed and the pellet was washed in 750 mL of 70% ice-cold ethanol before being resuspended in 50 mL of DNase/RNase free water.

For library preparation, the 98C BarSeq PCR described in¹⁹ was used to amplify the barcoded region of the transposons and PCR was performed in a final volume of 50 mL: 25 mL of Q5 polymerase master mix (New England Biolab), 10 mL of GC enhancer buffer (New England Biolab), 2.5 mL of the common reverse primer (BarSeq_P1¹⁹) at 10 mM, 2.5 mL of a forward primer from the 96 forward primers (BarSeq_P2_ITXXX¹⁹) at 10 mM and 50 ng to 2 mg of gDNA. The following PCR program was used: (i) 98 °C–4 min, (ii) 30 cycles of: 98 °C–30 s; 55 °C–30 s; 72 °C–30 s, (iii) 72 °C–5 min. After the PCR, 10 mL of each of the PCR products were pooled together to create the BarSeq library. 200 mL of the pooled library were purified using the MinElute purification kit (Qiagen) and final elution of the BarSeq library was performed in 30 mL in DNase and RNase free water.

The BarSeq library was quantified using Qubit dsDNA HS assay kit (Invitrogen) and then sequenced on HiSeq4000 (50 bp, single-end reads), by the IGM Genomics Center at the University of California San Diego.

Data processing and interaction fitness effects (IFE) analysis. For each library, BarSeq reads were first processed using the Perl script BarSeqTest.pl from¹⁹ to obtain the count file (all.poolcount) containing the number of reads per barcode for each sample. This pipeline requires a table where each barcode is mapped to a location in the genome. The Arkin lab (Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA) kindly provided the TnSeq table for the *E. coli* library. The original script used for this analysis originates from¹⁹ is publicly available on <https://bitbucket.org/berkeleylab/feba>.

The generated all.poolcount information was then implemented into custom R scripts from¹⁵ to determine the average fitness scores for each gene across three RB-TnSeq assay replicates for each of the eight cultures (<https://github.com/DuttonLab/RB-TnSeq-Microbial-interactions>). Insertion mutants that did not have a sufficient T0 count in each condition or that were not centrally inserted (10–90% of gene) were removed from the analysis and counts were then normalized using a set of five reference genes (*glgP*, *acnA*, *modE*, *leuA*—average of 52 strains each). Detailed explanation of the fitness calculation strategy can be found in the Readme document as well as in¹⁵. In each condition, to assess the possible variability between replicates, we measured the correlation between the three replicates (Supplementary Fig. 13). In this set of experiments, Pearson coefficient between replicates varied from 0.79 to 0.86 suggesting low technical noise.

Gene fitness values were then compared between interactive conditions and *E. coli* growth alone condition using two-sided *t* tests (when the equality of variance was verified by Fisher test) and correction for multiple comparison (Benjamini–Hochberg method²⁰). During the fitness calculation process, the sequencing data have been log₂-transformed that allows to avoid any issue with the compositional nature of the data that could interfere with a two-sided *t* test⁴⁸. Comparisons associated with an adjusted *P* value lower than 10% were considered a significant interaction fitness effect. The choice of the adjusted *P* value cutoff allows us to broadly screen for interaction-associated mutants. While a more stringent adjusted *P* value cutoff would reduce the number of insertion-mutants associated with interactions, the overall observations and conclusions of this study are consistent regardless of the *P* value selected (Supplementary Data 4).

Epistasis model analysis. We used the Python package Epistasis¹⁷ to run the epistasis analysis on *E. coli*'s genes. We implemented the model with the average fitness values across the three biological replicates along with the corresponding variance, for each gene and for each of the culture conditions (Alone, w *H.a*, w *G.c*, w *P.c*, w *H.a* + *G.c*, w *H.a* + *P.c*, w *G.c* + *P.c*, w *H.a* + *G.c* + *P.c*). The model is run for each gene individually. The 'genotypes' expected in the model correspond to the culture conditions and are binary coded based on the presence or not of *E. coli*'s partners *G. candidum*, *H. alvei* and *P. camemberti*. For instance, the Alone condition corresponds to

phenotype 000, the w *G.c* condition is coded 100, the w *H.a + P.c* condition is coded 011, the w *H.a + G.c + P.c* is coded 111 and so on. The ‘phenotype’ implemented in the model corresponds to the average fitness value across three biological replicates. After generating the genotype–phenotype map for each gene, the model was run a first time using the ‘local’ parameter to set-up the center of the map, and returns the set of epistatic coefficients for each gene calculated as follows:

$$\text{IFE}_{\text{obs,H.a,G.c}} = \text{IFE}_{\text{H.a}} + \text{IFE}_{\text{G.c}} + e_{\text{H.a,G.c}}$$

$$\text{IFE}_{\text{obs,H.a,P.c}} = \text{IFE}_{\text{H.a}} + \text{IFE}_{\text{P.c}} + e_{\text{H.a,P.c}}$$

$$\text{IFE}_{\text{obs,G.c,P.c}} = \text{IFE}_{\text{G.c}} + \text{IFE}_{\text{P.c}} + e_{\text{G.c,P.c}}$$

$$\text{IFE}_{\text{obs,H.a,G.c,P.c}} = \text{IFE}_{\text{H.a}} + \text{IFE}_{\text{G.c}} + \text{IFE}_{\text{P.c}} + e_{\text{H.a,G.c}} + e_{\text{H.a,P.c}} + e_{\text{G.c,P.c}} + e_{\text{H.a,P.c,G.c}}$$

To determine whether each coefficient was different from zero, we generated 7000 pseudoreplicates from our experimental variance from each gene and then calculated the probability each coefficient was above or below zero from this distribution. P-values were then adjusted for multiple comparison testing using Benjamini Hochberg Correction²⁰. As it has been shown previously that epistatic analyses can fail if the effects of mutations combine non-linearly instead of linearly (e.g. mutational effects multiply rather than add), we used the epistasis package to look for such nonlinearity. However, no nonlinearity was observed (Supplementary Fig. 14).

Data availability

Sequence data that support the RB-TnSeq analysis of this work have been deposited in the NCBI SRA database with SRA accession codes and BioProject code PRJNA843125 and SRA accession codes SRR19434865–SRR19434889. In addition to these sources, the fitness data calculated from these sequencing data and RB-TnSeq analysis are available in the Supplementary Data 1.

Received: 21 May 2022; Accepted: 28 November 2022

Published online: 31 December 2022

References

- Ortiz, A., Vega, N. M., Ratzke, C. & Gore, J. Interspecies bacterial competition regulates community assembly in the *C. elegans* intestine. *ISME J.* <https://doi.org/10.1038/s41396-021-00910-4> (2021).
- Friedman, J., Higgins, L. M. & Gore, J. Community structure follows simple assembly rules in microbial microcosms. *Nat. Ecol. Evol.* **1**(5), 109 (2017).
- Sanchez-Gorostiaga, A., Bajić, D., Osborne, M. L., Poyatos, J. F. & Sanchez, A. High-order interactions distort the functional landscape of microbial consortia. *PLoS Biol.* **17**(12), e3000550 (2019).
- Stressmann, F. A. *et al.* Mining zebrafish microbiota reveals key community-level resistance against fish pathogen infection. *ISME J.* **15**(3), 702–719 (2021).
- Gould, A. L. *et al.* Microbiome interactions shape host fitness. *Proc. Natl. Acad. Sci. U. S. A.* **115**(51), E11951–E11960 (2018).
- Vandermeer, J. H. The competitive structure of communities: An experimental approach with protozoa. *Ecology* **50**(3), 362–371 (1969).
- Sundarraman, D. *et al.* Higher-order interactions dampen pairwise competition in the zebrafish gut microbiome. *MBio* <https://doi.org/10.1128/mBio.01667-20> (2020).
- Grilli, J., Barabás, G., Michalska-Smith, M. J. & Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**(7666), 210–213 (2017).
- Mayfield, M. M. & Stouffer, D. B. Higher-order interactions capture unexplained complexity in diverse communities. *Nat. Ecol. Evol.* **1**(3), 62 (2017).
- Levine, J. M., Bascompte, J., Adler, P. B. & Allesina, S. Beyond pairwise mechanisms of species coexistence in complex communities. *Nature* **546**(7656), 56–64 (2017).
- Bairey, E., Kelsic, E. D. & Kishony, R. High-order species interactions shape ecosystem diversity. *Nat. Commun.* **7**, 12285 (2016).
- Billick, I. & Case, T. J. Higher order interactions in ecological communities: What are they and how can they be detected?. *Ecology* **75**(6), 1529–1543 (1994).
- Mickalide, H. & Kuehn, S. Higher-order interaction between species inhibits bacterial invasion of a phototroph-predator microbial community. *Cell Syst.* <https://doi.org/10.1016/j.cels.2019.11.004> (2019).
- Morin, M., Pierce, E. C. & Dutton, R. J. Changes in the genetic requirements for microbial interactions with increasing community complexity. *Elife* <https://doi.org/10.7554/eLife.37072> (2018).
- Pierce, E. C. *et al.* Bacterial-fungal interactions revealed by genome-wide analysis of bacterial mutant fitness. *Nat. Microbiol.* <https://doi.org/10.1038/s41564-020-00800-z> (2020).
- Heckendorn, R. B. & Whitley, D. Predicting epistasis from mathematical models. *Evol. Comput.* **7**(1), 69–101 (1999).
- Sailer, Z. R. & Harms, M. J. Detecting high-order epistasis in nonlinear genotype-phenotype maps. *Genetics* **205**(3), 1079–1088 (2017).
- Weinreich, D. M., Lan, Y., Wylie, C. S. & Heckendorn, R. B. Should evolutionary geneticists worry about higher-order epistasis?. *Curr. Opin. Genet. Dev.* **23**(6), 700–707 (2013).
- Wetmore, K. M. *et al.* Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *MBio* **6**(3), e00306–e315 (2015).
- Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **57**(1), 289–300 (1995).
- Szklarczyk, D. *et al.* STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**(D1), D607–D613 (2019).
- Rai, A. K. & Mitchell, A. M. Enterobacterial common antigen: Synthesis and function of an enigmatic molecule. *MBio* <https://doi.org/10.1128/mBio.01914-20> (2020).
- Sanchez, A. Defining higher-order interactions in synthetic ecology: Lessons from physics and quantitative genetics. *Cell Syst.* **9**(6), 519–520 (2019).

24. Taylor, M. B. & Ehrenreich, I. M. Higher-order genetic interactions and their contribution to complex traits. *Trends Genet.* **31**(1), 34–40 (2015).
25. Rkenes, T. P., Lamark, T. & Strøm, A. R. DNA-binding properties of the BetI repressor protein of *Escherichia coli*: The inducer choline stimulates BetI-DNA complex formation. *J. Bacteriol.* **178**(6), 1663–1670 (1996).
26. Jensen, K. F., Dandanell, G., Hove-Jensen, B. & Willemoës, M. Nucleotides, nucleosides, and nucleobases. *EcoSal Plus* <https://doi.org/10.1128/ecosalplus.3.6.2> (2008).
27. Boutrou, R., Kerriou, L. & Gassi, J.-Y. Contribution of *Geotrichum candidum* to the proteolysis of soft cheese. *Int. Dairy J.* **16**(7), 775–783 (2006).
28. Aziza, M., Couriol, C., Amrane, A. & Boutrou, R. Evidences for synergistic effects of *Geotrichum candidum* on *Penicillium camembertii* growing on cheese juice. *Enzyme Microb. Technol.* **37**(2), 218–224 (2005).
29. Ozturkoglu-Budak, S., Wiebenga, A., Bron, P. A. & de Vries, R. P. Protease and lipase activities of fungal and bacterial strains derived from an artisanal raw ewe's milk cheese. *Int. J. Food Microbiol.* **237**, 17–27 (2016).
30. Zengler, K. & Zaramela, L. S. The social network of microorganisms—How auxotrophies shape complex communities. *Nat. Rev. Microbiol.* **16**(6), 383–390 (2018).
31. Zuñiga, C. *et al.* Synthetic microbial communities of heterotrophs and phototrophs facilitate sustainable growth. *Nat. Commun.* **11**(1), 3803 (2020).
32. Guo, X. & Boedicker, J. Q. The contribution of high-order metabolic interactions to the global activity of a four-species microbial community. *PLoS Comput. Biol.* **12**(9), e1005079 (2016).
33. Enke, T. N. *et al.* Modular assembly of polysaccharide-degrading marine microbial communities. *Curr. Biol.* **29**(9), 1528–1535.e6 (2019).
34. Campbell, E. A. *et al.* Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* **104**(6), 901–912 (2001).
35. Pacheco, A. R., Osborne, M. L. & Segrè, D. Non-additive microbial community responses to environmental complexity. *Nat. Commun.* **12**(1), 1–11 (2021).
36. Kurkjian, H. M., Akbari, M. J. & Momeni, B. The impact of interactions on invasion and colonization resistance in microbial communities. *PLoS Comput. Biol.* **17**(1), e1008643 (2021).
37. Zhang, Y., Kastman, E. K., Guasto, J. S. & Wolfe, B. E. Fungal networks shape dynamics of bacterial dispersal and community assembly in cheese rind microbiomes. *Nat. Commun.* **9**(1), 336 (2018).
38. Warmink, J. A., Nazir, R., Corten, B. & van Elsas, J. D. Hitchhikers on the fungal highway: The helper effect for bacterial migration via fungal hyphae. *Soil Biol. Biochem.* **43**(4), 760–765 (2011).
39. Cosetta, C. M., Kfoury, N., Robbat, A. & Wolfe, B. E. Fungal volatiles mediate cheese rind microbiome assembly. *Environ. Microbiol.* **22**(11), 4745–4760 (2020).
40. Frey-Klett, P. *et al.* Bacterial-fungal interactions: Hyphens between agricultural, clinical, environmental, and food microbiologists. *Microbiol. Mol. Biol. Rev.* **75**(4), 583–609 (2011).
41. Lessard, M.-H., Viel, C., Boyle, B., St-Gelais, D. & Labrie, S. Metatranscriptome analysis of fungal strains *Penicillium camemberti* and *Geotrichum candidum* reveal cheese matrix breakdown and potential development of sensory properties of ripened Camembert-type cheese. *BMC Genomics* **15**, 235 (2014).
42. Boutrou, R. & Guéguen, M. Interests in *Geotrichum candidum* for cheese technology. *Int. J. Food Microbiol.* **102**(1), 1–20 (2005).
43. Fisher, R.A. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinb.* **52**(2), 99–433 (1919).
44. Bliss, C. I. The toxicity of poisons applied jointly. *Ann. Appl. Biol.* **26**(3), 585–615 (1939).
45. Sih, A., Englund, G. & Wooster, D. Emergent impacts of multiple predators on prey. *Trends Ecol. Evol.* **13**(9), 350–355 (1998).
46. Estrela, S., Sanchez-Gorostiaga, A., Vila, J. C. & Sanchez, A. Nutrient dominance governs the assembly of microbial communities in mixed nutrient environments. *Elife* <https://doi.org/10.7554/eLife.65948> (2021).
47. Wolfe, B. E., Button, J. E., Santarelli, M. & Dutton, R. J. Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. *Cell* **158**(2), 422–433 (2014).
48. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* **8**, 2224 (2017).

Acknowledgements

The authors would like to thank the following people and groups: the Arkin Lab and the Deutschbauer Lab at UC Berkeley for the *E. coli* Keio_ML9 RB-TnSeq library; K. Jepsen at the IGM Genomics Center at the University of California, San Diego for assistance with sequencing; and all members of the Dutton Lab for constructive comments on the manuscript. This work was supported by the National Institute of Health New Innovator grant #DP2 AT010401 (R.J.D.), the NIH Institutional Training grant NIH-T32GM007413 (A.J.M.), National Science Foundation CAREER Award DEB-1844963 (M.J.H.).

Author contributions

R.J.D., M.J.H. and M.A.M. conceptualized the study. M.A.M. performed the experiments. The Epistasis pipeline was written and adapted by M.J.H. and A.J.M. Data analyses were performed by M.A.M. and A.J.M. The article was written by M.A.M. and revised with input from all authors. The figures were made by M.A.M. with input from all authors. The study was supervised by R.J.D. and M.J.H.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-25303-1>.

Correspondence and requests for materials should be addressed to M.A.M. or R.J.D.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022