microLife, 2023, **4**, 1–15
DOI: 10.1093/femsml/uqad038
Advance access publication date: 12 September 2023
Research Article

Emergence of novel non-aggregative variants under negative frequency-dependent selection in *Klebsiella variicola*

Amandine Nucci, Juliette Janaszkiewicz, Eduardo P.C. Rocha, Olaya Rendueles*

Institut Pasteur, Université Paris Cité, CNRS, UMR3525, Microbial Evolutionary Genomics, F-75015, Paris, France
*Corresponding author. Institut Pasteur, Université Paris Cité, CNRS, UMR3525, Microbial Evolutionary Genomics, F-75015, Paris, France. E-mail: olaya.rendueles-garcia@pasteur.fr
Editor: [Carmen Buchrieser]

Abstract

Klebsiella variicola is an emergent human pathogen causing diverse infections, some of which in the urinary tract. However, little is known about the evolution and maintenance of genetic diversity in this species, the molecular mechanisms and their population dynamics. Here, we characterized the emergence of a novel rdar-like (rough and dry) morphotype which is contingent both on the genetic background and the environment. We show that mutations in either the nitrogen assimilation control gene *(nac)* or the type III fimbriae regulator, *mrkH,* suffice to generate rdar-like colonies. These morphotypes are primarily selected for the reduced intercellular aggregation as a result of MrkH loss-of-function which reduces type 3 fimbriae expression. Additionally, these clones also display increased growth rate and reduced biofilm formation. Direct competitions between rdar and wild type clones show that mutations in *mrkH* provide large fitness advantages. In artificial urine, the morphotype is under strong negative frequency-dependent selection and can socially exploit wild type strains. An exhaustive search for *mrkH* mutants in public databases revealed that *ca* 8% of natural isolates analysed had a truncated *mrkH* gene many of which were due to insertions of IS elements, including a reported clinical isolate with rdar morphology. These strains were rarely hypermucoid and often isolated from human, mostly from urine and blood. The decreased aggregation of these mutants could have important clinical implications as we hypothesize that such clones could better disperse within the host allowing colonisation of other body sites and potentially leading to systemic infections.

Keywords: evolution, negative frequency-dependent selection, morphological diversification

Introduction

How diversity emerges and is maintained in microbial populations are central questions in evolution and ecology. During evolution in sympatry, diversity can be driven by genetic drift, epistasis, or intercellular interactions (mostly competition) (Kassen 2014). For instance, competition for resources can select for key innovations allowing the use of previously unavailable nutrients (Blount et al. 2008). Similarly, historical contingency, that is, existing mutations in one clone may alter the range of subsequent mutations available and define different adaptive outcomes (Blount et al. 2008, Blount et al. 2012, Debray et al. 2022, Batarseh et al. 2023). One of the most potent forces of diversification is divergent selection. This is exemplified by structured environments where microniches can be generated owing to gradients of nutrient or oxygen (Rainey and Travisano 1998). Thus, adaptive morphotypes may emerge in different microniches giving rise to specialists (Kassen 2002, Baquero et al. 2021). Additionally, the existence of microniches may support growth of subpopulations with lower effective sizes which may be more subject to genetic drift. Similarly, balancing selection, by which a polymorphism in a given locus is sustained, can promote genetic diversification (Hedrick 2007). This is specially so under temporally or spatially variable environments or in fluctuating environments (Abdul-Rahman et al. 2021).

Once it has emerged, microbial diversity can be maintained by the aforementioned spatial structure (Comins and Hassell 1996, Kassen and Rainey 2004) that reduces or limits migration across subpopulations. This allows for the coexistence of different morphotypes even when this could not be so in well-mixed environments (Habets et al. 2006). Diversity can also be maintained by cross-feeding (D'Souza et al. 2018), interference competition (Czárán et al. 2002, Kerr et al. 2002), dormancy (Jones and Lennon 2010), positive (Rendueles et al. 2015) and negative frequency-dependent selection (Rainey et al. 2000, Velicer et al. 2000, Lemonnier et al. 2008). Specifically, positive-frequency dependent selection contributes to global diversity allowing genotypes that are less fit when present at intermediate frequencies to persist in patchily distributed populations provided that they are locally more abundant (Rendueles et al. 2015). Negative frequency-dependent selection occurs when increasing abundance results in decreasing fitness. This allows relatively rare variants to be maintained because they have a selective advantage over more common variants, avoiding local extinction.

For over forty years, evolution experiments have fueled the phenotypic and subsequent genotypic studies of bacterial diversification. Trait diversification has been documented in many of the phenotypes studied including cell size (Travisano et al. 1995, Baselga-Cervera et al. 2023), resource utilization (Tyerman et al.

Received 11 July 2023; **revised** 5 September 2023; **accepted** 9 September 2023

2005, Spencer et al. 2008) or swarming rate (Rendueles and Velicer 2017), among others (Travisano 1997, Rendueles and Velicer 2020, La Fortezza et al. 2022). One of the most visible features of these evolution experiments is the emergence of different colony morphotypes. Among the best well-known are the so-called wrinkly (WS) and fuzzy spreaders (FS) in *Pseudomonas fluorescens* when grown in static microcosms of nutrient-rich medium (Rainey and Travisano 1998). Similar wrinkly or rough phenotypes also emerge during growth of *Burkholderia cenocepacia* as a biofilm (Poltak and Cooper 2011) or of *Bacillus subtilis* cells during colonization of *Arabidopsis thaliana* roots (Blake et al. 2021). Many others have also been described and are reviewed in (Xu et al. 2022). Such rapid diversification also occurs in more natural settings like complex soil microcosms (Gómez and Buckling 2013) or as exemplified by the variety of *E. coli* colonies varying in size and motility during gut colonization (De Paepe et al. 2011).

The *Klebsiella pneumoniae* species complex is a metabolically versatile and diverse group belonging to the *Enterobacteriaceae* family and includes the ubiquitous *Klebsiella variicola* (Rosenblueth et al. 2004, Barrios-Camacho et al. 2019). The latter is commonly isolated as part of plant microbiomes as it can promote plant growth by nitrogen fixation (Pinto-Tomás et al. 2009). It is also an emerging human pathogen (Rodríguez-Medina et al. 2019) causing urinary tract infections (Potter et al. 2018), as well as in other diseases in wild and farm animals (Giannattasio-Ferraz et al. 2022). However, *K. variicola* has been largely neglected due to a historic taxonomic misclassification (Martínez-Romero et al. 2018, Rodrigues et al. 2018). Recent developments including fast multiplex PCR (Garza-Ramos et al. 2015), mass spectrometry profiles (Rodrigues et al. 2018), and phylogenomic analyses (Lam et al. 2021) allowed the appropriate *K. variicola* identification and prompted studies addressing this species' biology, population structure and virulence determinants (Potter et al. 2018). Analyses of the *K. variicola* pangenome revealed the existence of nine different pili, including two well-studied chaperon usher systems type 1 fimbriae encoded in the *fim* operon and the type 3 fimbriae (T3F) encoded by the *mrk* operon (Potter et al. 2018).

To study how morphotypic diversity emerges and how it is maintained in the *Klebsiella pneumoniae* species complex, and more specifically in *K. variicola*, we used a previous evolution study (Nucci et al. 2022), in which we evolved in parallel two hypervirulent *K. pneumoniae* strains (Kpn NTUH and Kpn BJ1) and one environmental *K. variicola* strain (Kva 342) as well as their non-capsulated isogenic mutants. The non-capsulated mutants were generated by an in-frame deletion of *wcaJ*, the first gene of the biosynthetic pathway and the gene most commonly mutated in lab-evolved non-capsulated clones (Buffet et al. 2021) and genomic datasets (Haudiquet et al. 2021). We propagated these six different genotypes for 675 generations in two nutrient-rich environments (artificial lung sputum and LB), and three nutrient-poor environments (artificial urine, soil and minimal media supplemented with glucose). Prior to each daily transfer, populations were homogenized by vigorous pipetting, and a bottleneck of 1% was applied. The described evolution experiment allowed us to show that originally capsulated populations rapidly diversified generating several morphotypes, all of which relied on the capacity of each clone to produce the polysaccharidic capsule. On the one hand, we observed small and translucent non-capsulated colonies, most often generated by IS insertions or point mutations in *wcaJ*, the initial glycosyltransferase. On the other hand, large, bulky and hypermucoviscous capsulated colonies also evolved frequently due to point mutations in *wzc*, another gene involved in capsule synthesis (Nucci et al. 2022). However, morphological diversification of the originally non-capsulated populations was rarer and remained to be addressed.

Here, we followed throughout the evolution experiment the emergence of a novel rough and dry morphotype (i.e. rdar-like) in non-capsulated Kva 342 populations, similar to the rdar-like morphotypes of other Enterobacteria, including *E. coli* and *Salmonella* (Römling 2005, Cimdins et al. 2017). We identified its genetic basis, quantified its fitness effects, and determined the underlying selection forces. Our work revealed that novel morphotypes under negative frequency-dependent selection emerged in non-capsulated populations. Finally, we searched in the literature and in a large genomic dataset for the prevalence of the identified mutations leading to this morphotype in natural populations. We found similar mutations in both capsulated and non-capsulated *Klebsiella pneumoniae* strains isolated from clinical settings, mostly in classical non-hypermucoid strains, suggesting that these morphotypes could also play an important ecological and evolutionary role in natural contexts.

Results

Emergence of rdar-like morphotypes in non-capsulated populations of *K. Variicola*

The regular plating of independent populations during the evolution experiment (see Materials and Methods) revealed the emergence of colonies displaying peculiar morphologies in six independent populations of non-capsulated *Klebsiella variicola* 342, *i.e.* twenty % of all non-capsulated Kva 342 populations (Fig. 1A). Such morphotype was not observed in populations descending from the capsulated ancestor nor from the two *Klebsiella pneumoniae* strains, Kpn NTUH or Kpn BJ1. These morphotypes evolved thrice in M02 and thrice in AUM, two environments with low-carrying capacity, but not in ASM or LB, two environments with high-carrying capacity. We isolated one clone from each population for further analyses. Despite the similarities of the morphotypes at the single-colony level (originating from single cells), at the population-level (originating from an overnight culture), a morphological difference was evident between 4D2, 4D6 and 6B3 (morphotype 1) vs 4D4, 6B1 and 6B4 (morphotype 2) clones (Fig. 1A).

To determine the frequency of the morphotypes and their population dynamics, we plated the six evolving populations at all intermediate time-points. After 50 cycles, *ca* 330 generations, the morphotypes had already emerged in all populations. Clones displaying morphotype 2 reached significantly higher maximum frequencies compared to those of morphotype 1(t-test, $P = 0.009$). These ranged between 12.7 and 66.6% of the population (Fig. 1B). However, after reaching maximum frequency, most morphotypes experienced a strong decrease in frequency, and towards the end of the experiment reached a frequency of ∼16% of the total population.

The observed morphotypes are similar to the abovementioned rdar phenotype observed in other Enterobacteria (Römling 2005), or the wrinkly spreader of *Pseudomonas fluorescens* (Rainey and Travisano 1998). These well-known morphotypes rely on the expression of specific surface adhesins and exopolysaccharides, most notably curli and cellulose (White and Surette 2006, Cimdins et al. 2017, Di Sante et al. 2018). A bioinformatic search for the *csgA-G* operon responsible for curli formation in the genome of Kva 342 revealed no match (see Methods, Table S1). We then looked for the *bcsABCDEFGQZ* operon responsible for cellulose biosynthesis. The Kva 342 had a complete cellulose synthesis (Fig. S1A). We tested the ability of the ancestral strain and the

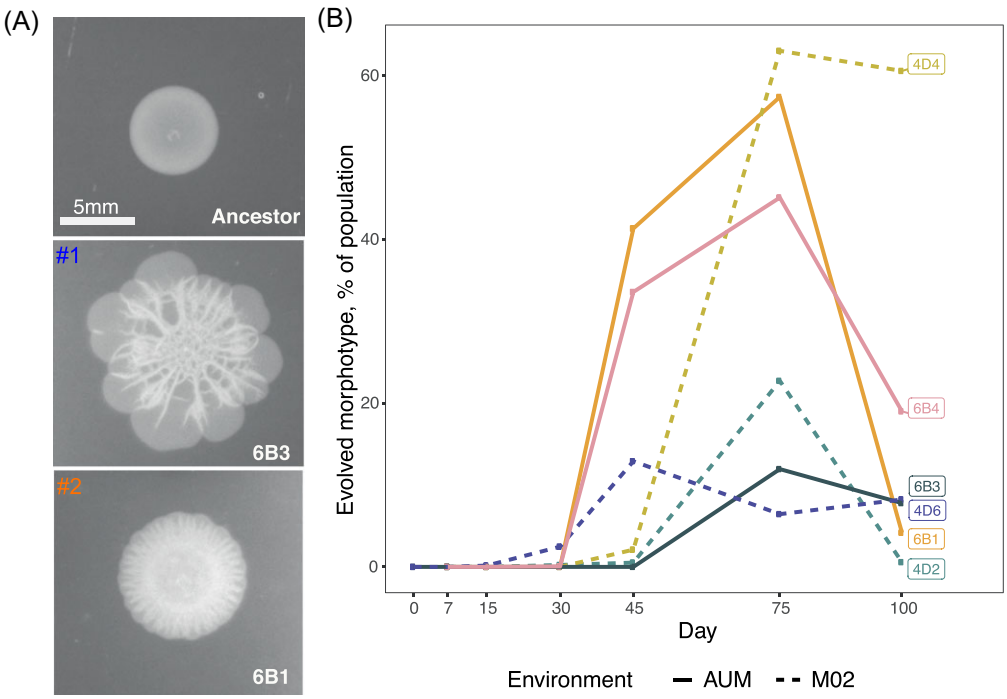

Figure 1. Phenotype and evolution of rdar- like morphotype in *K. variicola* populations. **(A)** Representative images of the ancestor and two rdar-like population-level colonies from clone 6B3 (morphotype 1) and 6B1 (morphotype 2). About 5 μL of an overnight culture were spotted on an LB plate and allowed to grow for 24 hours. **(B)** Population dynamics of evolved morphotypes in the six independent populations. Populations with morphotype 1 (4D2, 4D6 and 6B3) are displayed in cold colors (blue), whereas populations belonging to morphotype 2 (4D4, 6B1, and 6B4) are depicted with warmer colors (orange).

evolved clones to produce cellulose by the specific binding to calcofluor dye. No differences in colony morphotype were observed between the ancestor and the evolved clones (Fig. S1B). We thus conclude that the emergence of these two novel rdar-like morphotypes in *K. variicola* does not rely on the expression of curli or exopolysaccharides, the two known mechanisms involved in the formation of rdar-like morphotypes in Enterobacteria (Römling 2005, White and Surette 2006).

Mutations in two type 3 fimbriae regulators are responsible for the rdar-morphotypes

To determine the genetic basis of the rdar-like morphotype, we performed whole-genome sequencing of four randomly chosen rdar-like clones from distinct populations, two of each morphotype. We then compared the Illumina reads to the ancestral sequence of Kva 342 using *breseq* (0.30.1) (Deatherage and Barrick 2014) (Table S2) (See Methods). We observed that different mutational events (either SNPs or insertion of IS) occurred in two genes *mrkH* (morphotype 1) and *nac* (morphotype 2) (Table 1). These mutations were only present in rdar clones but absent from other wild type-like clones in their respective populations (Nucci et al. 2022). Both *mrkH* and *nac* code for nucleic acid binding proteins, recognizing either RNA (confidence score = 0.85) and DNA (confidence score = 0.98), respectively, as predicted by DeepFRI (Gligorijević et al. 2021). Indeed, MrkH binds to the α-CTD of RNA polymerase and acts as a transcriptional activator of the *mrk* operon involved in the synthesis of T3F (Wilksch et al. 2011). NAC is a LysR-type transcriptional activator known to be expressed in nitrogen-limited conditions and activating sigma70-dependent genes (Bender 1991). PCR and sanger sequencing of the two non-sequenced clones (4D2 and 6B1) also revealed mutations in *mrkH* and *nac* respectively (Table 1). Interestingly, a search for the 15-bp consensus sequence for the LysR-binding box in *Klebsiella* (ATA-N9-TAT) (Frisch and Bender 2010) revealed two hits in the mrk operon; one 200 base pairs upstream the start codon of *mrkA* and a second one upstream the promotor of *mrkJ* (Fig. S2). Interestingly, this LysR-binding box, overlaps the 'mrkH box', a palindromic sequence to which MrkH binds to control transcription of *mrkHI* and *mrkABCDF* clusters. Deletion of the 'mrkH box' also leads to reduction of *mrkJ* expression (Ares et al. 2017). This suggests that mutations in either *mrkH* or *nac* ultimately impact type 3 fimbriae.

To test whether the rdar-like morphotype was dependent on the *mrkH* and *nac*, we restored the ancestral allele in four different evolved clones from different environments by markerless allele exchange (See Methods). Ancestral *mrkH* was amplified and recombined into the evolved clones. Restoration of the ancestral *mrkH* allele resulted in the loss of morphotype 1 (Fig. S3A) and restoration of the *nac* allele in clones 4D4 and 6B1 resulted in the loss of morphotype 2 (Fig. S3B). We then showed that SNPs in *mrkH* or *nac* were sufficient to produce the evolved morphotypes (Fig. S3C) when inserted in an ancestral genotype. Finally, to show that IS insertions in the 3' of the gene, due to IS903 from IS5 family in *mrkH* were equivalent to a loss of function, we generated an in-frame deletion of the entire *mrkH* gene. This led to the emergence of the rdar-morphotype (Fig. S3C). Overall, our results show that mutations in known regulators of T3F are responsible for the rdar-like morphotype.

Mutations in *mrkH* and *nac* decrease aggregation and limit biofilm formation but increase growth rate

Rdar-like morphotypes are well known for having altered intercellular interactions which result in differences in aggregation and formation of biofilm (Da Re and Ghigo 2006). To quantify aggre-

Table 1. Convergent mutations in clones displaying rdar-like morphotypes

Clone	Morphotype	Environment	Position	Mutation	Change	Annotation
4D2	1	M02	844 700	IS insertion	608/705 nt	*mrkH*
4D4	2	M02	5 497 313	C→T	G275S	*nac*
4D6	1	M02	844 713	IS insertion	621/705 nt	*mrkH*
6B1	2	AUM	5 498 042	C→T	R228C	*nac*
6B3	1	AUM	845 042	G→A	P98S	*mrkH*
6B4	2	AUM	5 498 183	C→A	intergenic (–48/+212)	*nac/argH*

gation, we measured the absorbance of the top layer of sitting cultures through time as a proxy for sedimentation, and thus for increased cell-to-cell interactions. In aggregation tests, higher relative absorbance represents lower sedimentation. Biofilm formation was measured using the traditional crystal violet staining method and speaks to cell-to-surface interactions and biomass. We showed that all evolved clones, independently of the gene and the mutational event (SNP or IS insertion) had a diminished capacity to aggregate compared to the ancestor (Fig. S4A). As expected, this was mostly associated with a decreased capacity to form biofilm except for the evolved clones 4D2 and 4D6, both of which had an IS inserted in *mrkH* (Fig. S4B).

We hypothesize that mutations in *mrkH* and *nac* could be either responsible for the decreased aggregation, for changes in biofilm formation or for both. We tested this using the abovementioned mutants. The reversion of SNPs in *nac* to the ancestral allele restored wild type levels of both aggregation (Fig. 2A) and biofilm formation (Fig. 2C). A change of the ancestral sequence by one SNP resulting in the amino acid change G275C in the ancestral background recapitulated the evolved phenotype (Fig. 2C). We thus conclude that SNPs in *nac* alone are enough to explain the changes observed in the evolved morphotype 2.

Reversal of the amino acid change P98S or reversal of the IS insertion in the gene *mrkH* to the ancestral allele restored wild type levels of aggregation. Similarly, insertion of the evolved allele or deletion of *mrkH* in the ancestor recapitulated the evolved phenotype and decreased aggregation (Fig. 2B). Absence of significant differences between the deletion of the gene *mrkH* and P98S in the ancestral background indicated that this amino acid change results in a loss of function (Two-sided t-test, $P > 0.05$). However, we do observe interesting differences in biofilm formation, depending on the environment. As expected, in AUM, *mrkH* mutations alone could explain the differences between evolved and ancestral phenotypes. This was not so in M02. For instance, the deletion of *mrkH*, or the reconstitution of a full *mrkH* in the evolved clone did not significantly alter biofilm formation compared to their wild type or the IS-interrupted genes, respectively. This suggests that *mrkH* is not being selected for changes in biofilm formation in our evolution experiment, but rather for changes in aggregation.

We then tested whether *nac* and *mrkH* mutations in evolved clones resulted in increased growth rate, as we would expect due to adaptation to the environment. We calculated the area under the growth curve (AUC), a measure that considers the lag time, maximum growth rate and population yield. All evolved clones (Fig. S4C), as well as those with the insertion of *mrkH* or *nac* evolved allele (or deletion of the gene) in an ancestral background, had a growth advantage irrespective of the environment (Fig. 2EF). Overall, the differences in growth observed in *mrkH* and *nac* mutants suggest that these mutations are adaptive.

Finally, to test if phenotypic effect in *mrkH* mutants results from changes in the expression of T3F, we performed quantitative RT-PCR. Loss-of-function of *mrkH* reduced significantly the expression of *mrkA* (the major subunit pilin of T3F) in both AUM and M02 (Two-way ANOVA, dF= 5, F = 4.8, $P = 0.007$) (Fig. S5).

Taken together, mutations in *mrkH* and *nac* are responsible for the decreased aggregation and faster growth of evolved rdar-like morphotypes. Further, mutations in *mrkH* also diminish biofilm formation, but in an environment-dependent way, suggesting that its role in biofilm formation was not the primary force selecting for this morphotype.

Mutations in *mrkH* are adaptive and under negative frequency-dependent selection in AUM

The parallel emergence of rdar-like morphotype in populations that evolved in M02 and AUM suggested that it is adaptive. To test this hypothesis, we focused on the clones with mutations in *mrkH*. We excluded *nac* in this analysis because it belongs to one of the largest family of regulator proteins known and could alter numerous other phenotypes via activation of sigma70-dependent genes (Bender 1991). As expected, evolved clones had a significant fitness advantage compared to the ancestor in both environments (Fig. S6A). We then performed direct competitions in a 1:1 ratio of the ancestral strain against its isogenic P98S or the loss-of-function *mrkH* deletion in the two environments. Additionally, we also competed an evolved clone against its isogenic mutant in which the ancestral *mrkH* was restored. In AUM, mutations in *mrkH* are enough to explain the increase in fitness of the evolved clones (Fig. 3AB). In M02, *mrkH* increases fitness in an evolved background, but not in an ancestral background. This suggests that it requires other mutations to improve fitness (Fig. 3AB).

Some rdar-like morphotypes reached high frequencies during the evolution experiment, but towards the end, their frequencies decreased. This could be indicative of negative frequency-dependent selection. To test this, we repeated the competitions experiments but with the evolved alleles in an initial ratio of 1:9. In M02, fitness gains were equivalent across the different inoculation ratios (Two-sampled t-test, $P = 0.71$, N=3). However, in AUM, we found that when inoculated in the minority, evolved alleles were not only fitter than the ancestor (Fig. 3CD, Fig. S6B), but the fitness gain was significantly higher, compared to competitions in which both clones were inoculated at similar frequencies (Two-sampled t-test, $P = 0.04$, N=3). Further, fitness of evolved allele correlated negatively with proportion at the beginning of the co-culture (Fig. 3E) in AUM but not in M02. Finally, even when inoculated in the minority, rdar-like morphotypes were more frequent at the end of the competition.

Taken together, *mrkH* mutations result in reduced *mrkA* expression and reduced aggregation. In M02, they are adaptive in the genetic context in which they emerged, whereas, in AUM, these mutations are adaptive even in the ancestral background and are under negative-frequency-dependent selection.

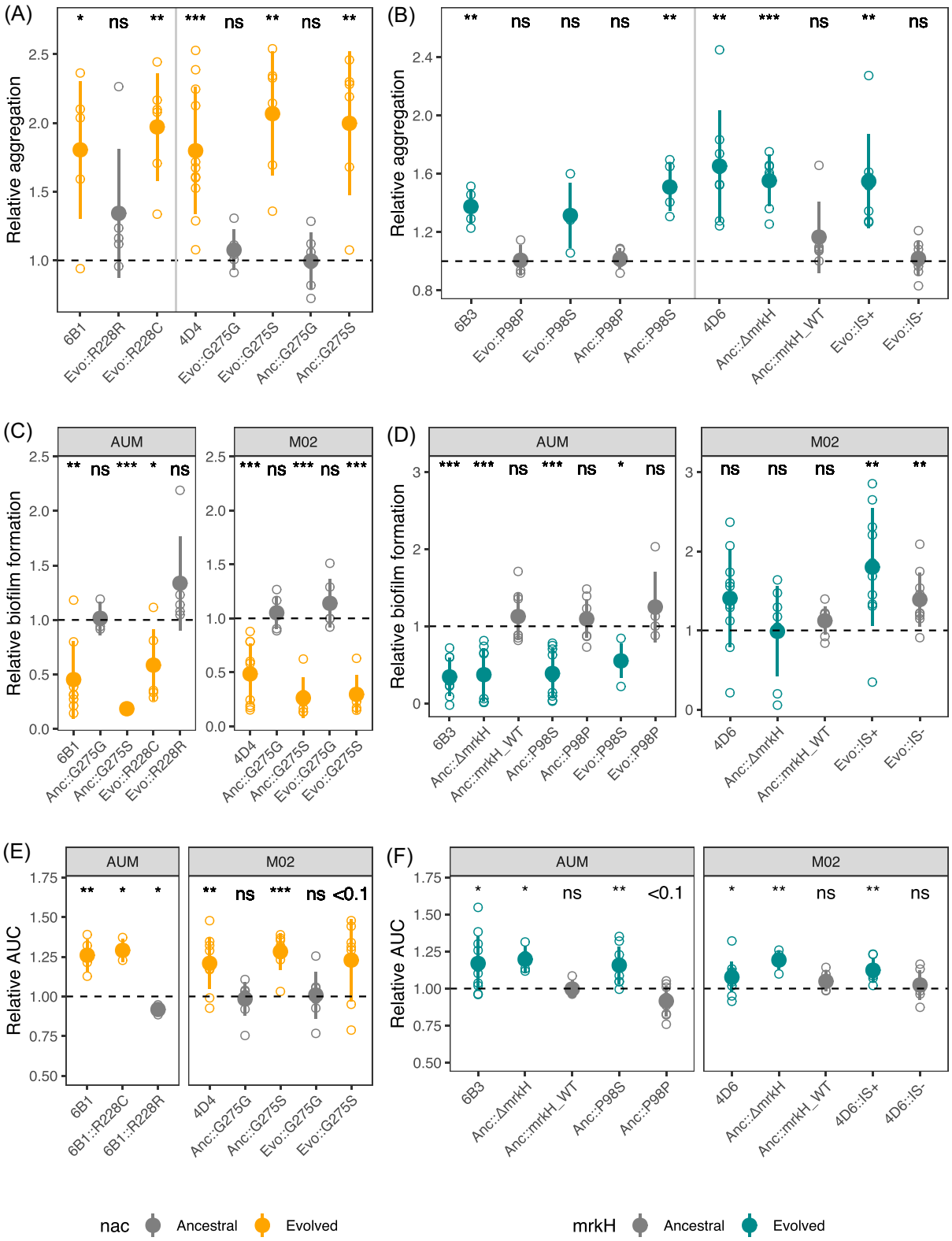

Figure 2. Mutations in *mrkH* and *nac* diminish aggregation, biofilm formation and increase growth. **(A and B)** Relative aggregation of *nac* **(A)** and *mrkH* **(B)** mutants is quantified by the absorbance (OD_{600nm}) of the top layer of culture in static conditions for 24 hours and the results are the ratio of the values for evolved over the initial clones. Reported data correspond to differences after 4.5 hours. High absorbance results from low aggregation levels. Calculation of the area under the aggregation curve results in qualitatively similar results. **(C and D)** Biofilm formation of *nac* **(C)** and *mrkH* **(D)** mutants was measured in the evolution growth media in which each mutation emerged. **(E and F)** Area under the growth curve of evolved and ancestral *nac* **(E)** and *mrkH* **(F)** alleles. Data is represented relative to the ancestral strain (dashed line). The AUC was calculated using the formula *trapz* from the pracma package for R. Grey points represent clones with ancestral alleles whereas colors indicate clones with evolved *mrkH* (green) or *nac* (orange) alleles. Small open points indicate independent biological replicates. Large, closed points represent the average of biological replicates and error bars indicate the standard deviation. Statistical analysis was performed to compare all alleles to its non-capsulated ancestor. One-sample two-sided t-test, difference from 1. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, ns $P \geq 0.05$.

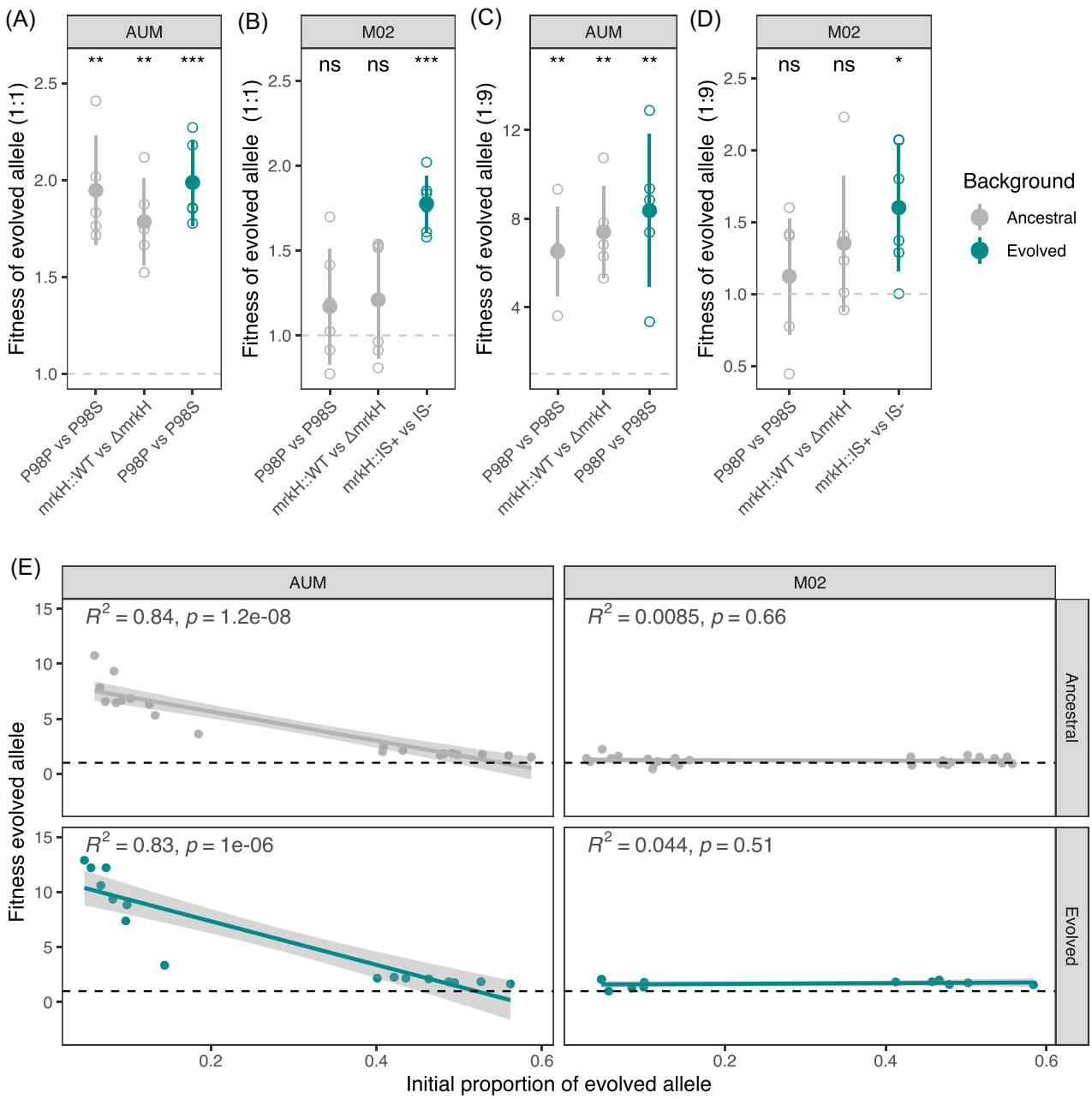

Figure 3. Fitness effects of mutations in *mrkH*. Competitions were performed at an initial ratio of 1:1 **(A and B)** and of 1:9 **(C and D)**, with the ancestral allele in the majority. Grey dots correspond to competitions in which the ancestor was competed against isogenic mutants with carrying either a SNP in *mrkH* or a *mrkH* deletion. Green dots correspond to competitions among evolved clones which were reverted or not to ancestral allele. Each allele was tested in its respective evolutionary environment. Large, closed points represent the average of biological replicates and error bars indicate the standard deviation. Open points indicate individual biological replicates. One-sample two-sided t-test, difference from 1. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, ns $P > 0.05$. **(E)** Correlation between fitness of evolved clone versus the initial proportion of the population. Fitness of each evolved allele was calculated using the ratio $\frac{T24}{T0}$. Each dot represents an independent biological replicate of an evolved: ancestral competition, across different genetic backgrounds. P-values were calculated and plotted using the stat_cor function in the *ggpubr* package.

The presence of the capsule limits effect of *mrkH* mutations on aggregation

The emergence of the rdar-like morphotype was exclusively observed in the non-capsulated background. This is striking as Kva 342 is non-mucoid and mildly capsulated in all media, but especially in AUM (Buffet et al. 2021). We first hypothesized this could be due to potentially deleterious effects for the cell in a capsulated background. However, *mrkH* mutant clones (both gene deletions and substitution with evolved allele) were obtained easily by allele exchange (see Methods) thereby suggesting that these clones are not deleterious and could have emerged in the experiments with capsulated strains. Of note, insertion of *mrkH* and *nac* mutations produced a very noticeable morphotype, which could have gone unnoticed (Fig. S7A). The resulting colonies were slightly rough but not dry, and thus, strictly speaking they are not rdar-like, as the capsule masks the dryness, but roughness at the borders can be observed. Next, we hypothesized that the *mrk* operon was not expressed in our evolutionary conditions in capsu-

lated background and was therefore 'invisible' to selection. However, qRT-PCR showed marginal or no differences in expression of the *mrkA* between capsulated and non-capsulated genetic background (Fig. S7B). We then hypothesized that mutations in *mrkH* had milder fitness effects, if any, on capsulated cells and were thus not strongly selected for. Direct competitions between the capsulated ancestor and evolved *mrkH* alleles showed that the latter were fitter. This was significantly so in AUM, both at 1:1 and at 1:9 ratio (Fig. 4A). Comparison of fitness effects across capsulated and non-capsulated clones showed that these effects were only marginally larger in non-capsulated than in capsulated clones. This suggests that rdar-like mutations could be under weaker selection in capsulated clones (Fig. S7C).

To test which phenotype is driving the selection for rdar-like clones in non-capsulated, we tested growth, aggregation and biofilm formation in a capsulated background (Fig. 5BCD) and compared it to the respective non-capsulated construction (Fig. S6 D–F). Growth curves of capsulated mutants revealed that the relative effect of ancestral and evolved *mrkH* and *nac* alleles (Fig. 4B) is similar in capsulated and non-capsulated strains (Fig. S7D). Of note, in M02, overall increased growth rate of capsulated clones is due to the large fitness advantage provided by the capsule in Kva 342 in M02 (Fig. S7D) (Buffet et al. 2021). These results imply that the absence of rdar-like morphotypes in capsulated strains is not related to changes in growth rate, suggesting that selection would be acting on another phenotype other than growth. Mutations in *mrkH* and *nac* decreased biofilm formation in capsulated strains (Fig. 4D) in a similar magnitude than in the non-capsulated strains (Fig. S7F). This again suggests that selection would not be acting on biofilm formation.

Capsulated strains aggregate poorly as the capsule can mask adhesins (Schembri et al. 2005). Despite such diminished aggregation in capsulated backgrounds, evolved *mrkH* or nac alleles did not further limit cell-to-cell interactions. This reveals significant differences in aggregation between ancestral and evolved *mrkH* and *nac* alleles in function of the presence of a capsule (Figs 2B, 4C; Fig. S7E).

Taken together, among the three traits we assessed (growth, aggregation, and biofilm) that could directly contribute to fitness of the evolved rdar-like morphotype, only aggregation was different in capsulated clones when compared to non-capsulated. This strongly suggests that rdar-like clones alleles are being selected for their impact in cell-to-cell aggregation.

Truncations and IS insertions in *mrkH* are found in natural isolates

Recently, a similar rdar-like morphotype was described in a clinical isolate of *K. pneumoniae* from a blood sample (PBIO3459) (Sydow et al. 2022). As the rdar-morphotype evolved more readily in the non-capsulated background in our experiments, we first abalyze the genomic sequence of PBIO3459 to test if it could be non-capsulated. Using ISFinder (Siguier et al. 2006), we identified IS903 belonging to the IS5 family inserted in the *wza* gene (position 1127/1140), the outer membrane exporter of the capsule. Loss-of-function of Wza is associated to a loss of capsulation (Buffet et al. 2021). We then searched the genome for MrkH and NAC protein sequences and compared them to those of our reference strains. The protein sequences of NAC are highly conserved between *K. pneumoniae* and *K. variicola*, revealing only one amino acid changes between the two species in position 108. Accordingly, in PBIO3459, we only observed one additional amino acid change corresponding to N229K compared to the other *K. pneumoniae* proteins. Such change is one amino acid away from that of clone 6B1 (R228C) and is also located in the DNA-binding domain (Fig. S8). This could potentially explain the rdar-like morphotype. We then analysed the MrkH sequence. The MrkH sequence of PBIO3459 seemed truncated (Fig. S8). We hypothesized that this could be the result of an IS insertion. Indeed, ISFinder revealed the presence of IS903 overlapping the end of *mrkH* sequence. Interestingly, IS903 is also responsible for mutations in *mrkH* in *K. variicola* 342. Independently of the mutation in *nac*, insertion of IS in *mrkH* could also explain the rdar phenotype in such clinical sample.

To further verify whether *nac* or *mrkH* mutations lead to a rdar-like morphotype in clinical isolates, we took advantage of a recently available *Klebsiella pneumoniae* panel of diverse clinical isolates (Martin et al. 2023). This collection consists of 100 sequenced *K. pneumoniae* clinical isolates worldwide and representative of the species diversity, hereafter referred to as the Klebsiella Diversity Panel. We searched for NAC and MrkH protein sequences in the genomes. We detected three strains, MRSN- 560539 and MRSN-375436 isolated from urine, and MRSN- 730567 isolated from blood with a truncated *mrkH* (Table 2). In all instances, the truncated *mrkH* was found at the border of the contig. This could be by chance, but could also be suggestive of an IS insertion, as these often complicate genome assemblies. To test whether there was an anomaly in *mrkH* sequence, we performed a PCR and confirmed that an IS element was inserted in all three genomes in *mrkH* (Fig. 5A, Table 2). We then analysed the colony morphology. We did not observe dry colonies, as all remained capsulated. However, we do observe roughness around the edges of the colony, similar to what was observed in the capsulated Kva 342 mutants we constructed (Fig. 5B). NAC protein of this strain was identical to that of our wild type *K. pneumoniae* strains. This confirms that loss-of-function mutations in *mrkH* can also lead to a wrinkly phenotype in capsulated *K. pneumoniae* isolated in the clinic.

To test the broader prevalence of truncated *mrkH* gene or IS insertion events, we downloaded and annotated the genomes (complete & draft) of all available *K. variicola* and 10 000 random *K. pneumoniae* genomes in the Pathosystems Resource Integration Center (PATRIC) genome database (Wattam et al. 2014). We searched for MrkH homologs and observed that ∼4% and ∼17% of MrkH proteins of *K. variicola* and *K. pneumoniae,* respectively, were either truncated or found at the border of contigs (Table S3) and could potentially exhibit rdar-like morphotypes. We then tested whether these proteins with abnormal lengths were interrupted by an IS. Indeed, we observed at least three *K. variicola* genomes had an IS inserted in *mrkH* (Table S4). In *K. pneumoniae,* we observed 133 (76 at the contig border and 57 inside the contig) of such events, most of which were mediated by IS5-like elements, among which IS903 (Fig. 5C; Table S4). Isolates with interrupted *mrkH* were mostly isolated from a human host, 35 of which (26%), were isolated from urine and another 35 from blood cultures (Fig. 5D). This correlates with the observations in the Klebsiella Diversity Panel.

The two *K. pneumoniae* strains used in our evolution experiment present a hypermucoid phenotype due to the presence of *rmpA*, a well-known mucoid regulator (Nassif et al. 1989). Given that in these two strains we did not observe mutations in *mrkH*, we hypothesized that specific mucoid capsule-fimbriae interactions could be at play and that non-mucoid *K. pneumoniae* strains would behave and evolve more like *K. variicola*. To test this, we searched for *rmpA* in our genomic datasets. In the Klebsiella Diversity Panel, five strains encoded *rmpA* (Martin et al. 2023), none of which presented truncations in *mrkH*. Similarly, in the *K. variicola* dataset, four strains encoded *rmpA* (0.6%) all of which had a

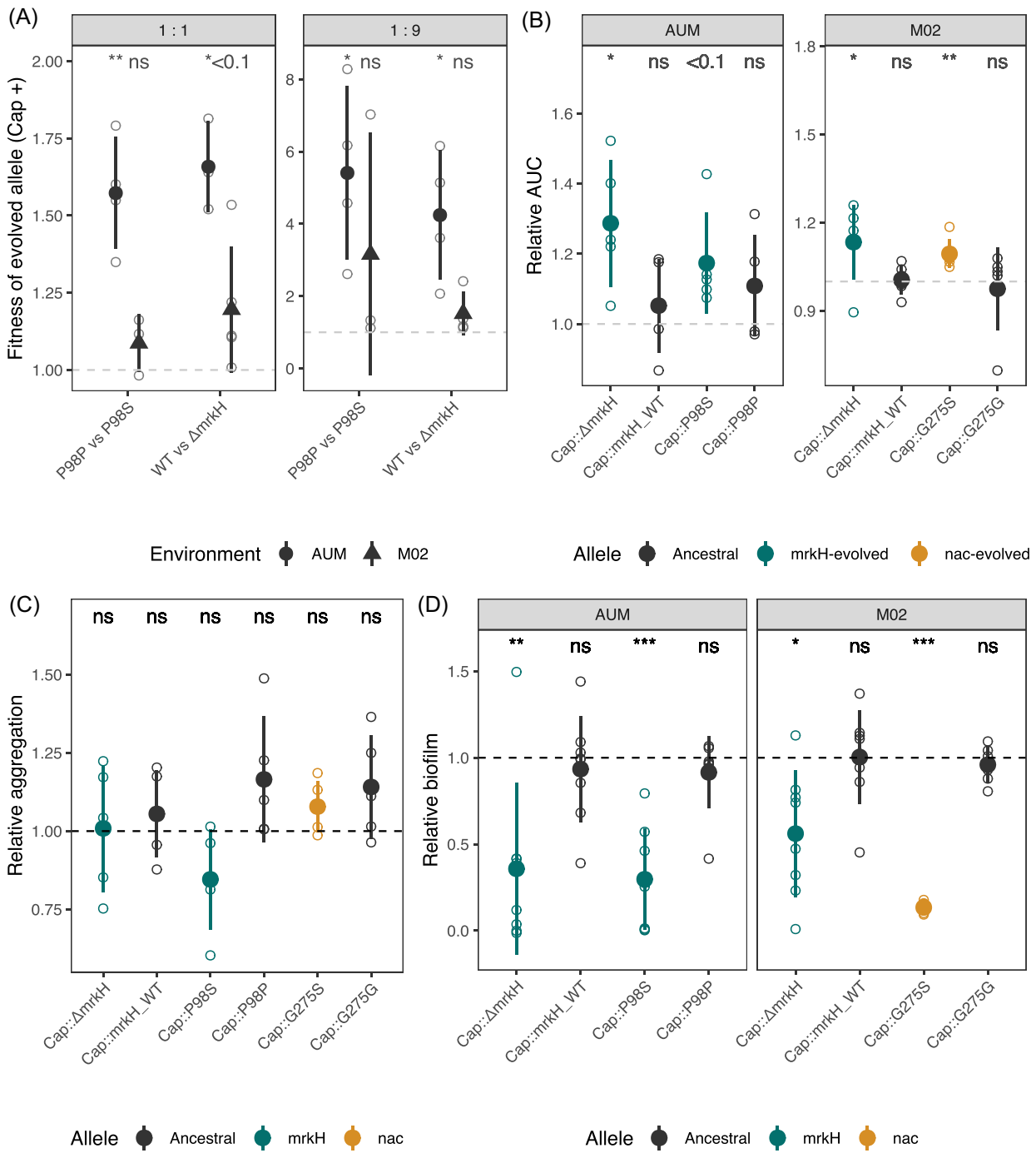

Figure 4. Effects of *mrkH* and *nac* mutations in capsulated ancestors. **(A)** Fitness effects of evolved alleles in capsulated strains when mixed in a 1:1 (0.5) and in a 1:9 (0.1) ratio. **(B)** Area under the growth curves relative to the capsulated ancestor. **(C)** Aggregation of *mrkH* and *nac* mutants relative to the capsulated ancestor. Reported data correspond to differences after 4.5 hours in absorbance (OD_{600nm}). **(D)** Biofilm formation of *mrkH* and *nac* alleles was measured in the evolution growth media in which each mutation emerged and compared to the non-capsulated mutants (dashed line). Black points represent capsulated clones with ancestral alleles whereas dark green points indicate capsulated clones with evolved *mrkH* (green) or *nac* (dark orange) alleles. Small open points indicate independent biological replicates. Large and closed points represent the average of biological replicates and error bars indicate the standard deviation. Statistical analysis was performed to compare all alleles to its non-capsulated ancestor. One-sample two-sided t-test, difference from 1. * $P < 0.05$, **$P < 0.01$, *** $P < 0.001$, ns $P > 0.05$.

complete *mrkH* gene. However, one strain presented 5 SNPs compared to the reference sequence of Kva 342, potentially suggesting a rdar-like phenotype. To increase the statistical power, we performed the same analyses in the *K. pneumoniae* PATRIC dataset, in which 4.1% of the genomes encoded *rmpA* (N=391). Of those, only 22 presented *mrkH* truncations (5.6%). This contrasts with the amount of truncations present in the entire database (17%) and suggests a strong association between *mrkH* truncations and absence of *rmpA*, *i.e.* non-mucoid strains (Fisher's Exact test, $P < 0.001$).

Table 2. Strains from the Klebsiella Diversity Panel in which truncation of *mrkH* gene was detected.

MRSN ID	Catalog NR	Sample type	Region	Year	MLST	K locus	O locus	IS family	Position in *mrkH*
560 539	55 571	Urine	N. America	2018	3050	KL36	O4	IS1202	591/729
375 436	55 550	Urine	N. America	2016	13	KL3	O1/O2v2	IS21	428/729
730 567	55 595	Blood	N. America	2019	1621	KL46	03b	IS110	546/729

Strain information (sample type, K locus type, O locus type, region, year and ST) as reported previously (Martin et al. 2023). IS family was determined by blast to the ISFinder database (Siguier et al. 2006). The nucleotide position in which the IS is inserted is also reported.

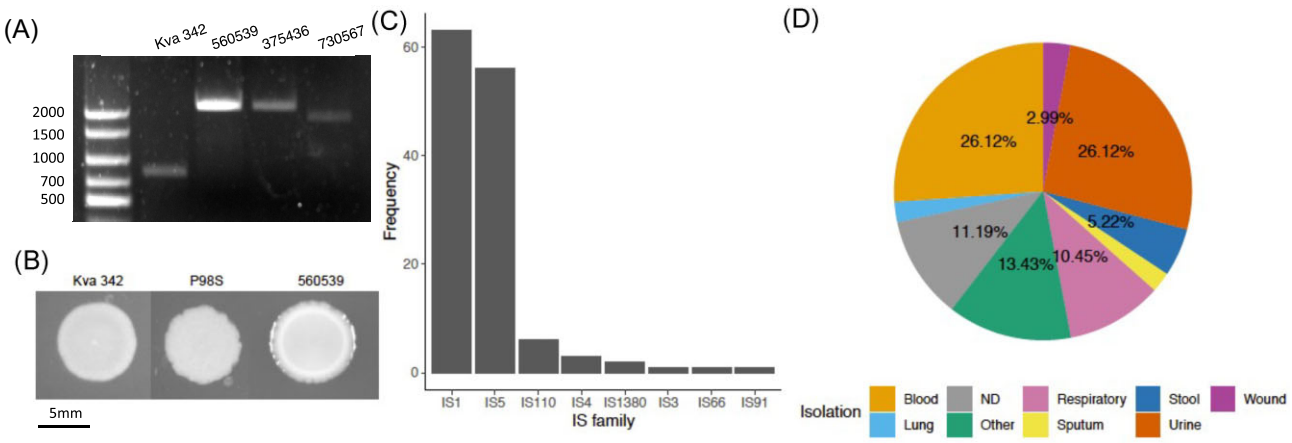

Figure 5. Interruption of *mrkH* gene by IS insertion in natural isolates. **(A)** PCR confirmation that the *mrkH* truncation and the contig breaks are due to the presence of an IS in the *mrkH* gene in strains from the Klebsiella Diversity Panel. **(B)** Colony morphology of capsulated Kva 342, its isogenic *mrkH* mutant resulting in P98S and strain MRSN-560539. **(C)** Family of ISs found to interrupt *mrkH* in genomes from the PATRIC database. **(D)** The source of isolation of genomes in which an IS was identified co-localizing with *mrkH*. Metadata was retrieved from the PATRIC database and manually curated. Lung and sputum isolates represented less than 2% each.

Taken together our analyses suggest that the unique morphotype here described can be isolated in natural populations, including clinical isolates, and are mostly generated by IS sequences. Yet, these morphotypes are rare, as one would expect from traits that are under negative-frequency-dependent selection.

Discussion

We have studied the outcome of an evolution experiment with three capsulated Klebsiella strains (and their non-capsulated mutants). This experiment has shed light on the general patterns of adaptation of the species independently of strain specificities, including the pervasive evolution of hypermucoidy (Nucci et al. 2022). Similarly, it revealed how biofilm formation could evolve as a latent phenotype and how it coevolves with other fitness-related traits like surface-attached polysaccharides and population yield. This led to the identification of repeated mutations in the tip adhesin of T3F and revealed its conspicuous role in biofilm formation (Nucci et al. 2023).

Here, we focus on *K. variicola* and its morphotypic diversification. We expand the genomic analyses to a larger set of *K. variicola* and *K. pneumoniae* genomes to suggest that this morphotype is also observed in clones isolated from the urinary and blood environments. Such morphotype is contingent on the genetic background and on the environment. Its emergence relies on the absence of the bacterial capsule and on growth in nutrient-poor conditions. The latter suggests that in nutrient-rich conditions, these morphotypes were not significantly fitter than the wild type and were thus not selected for. In addition, we identified that mutations in two regulators, *mrkH* and *nac*, were directly responsible for such morphotype. These mutations significantly reduce aggregation via decreased expression of surface fimbriae. Collectively, our data showed that the rdar-like morphotype in *K. variicola* is primarily selected due to its role in cell-to-cell aggregation and not in growth or biofilm formation. This relies on several pieces of evidence. First, the increased growth rate of rdar-morphotypes which could easily explain increases in fitness and strong selective coefficient, is also observed in capsulated clones, where the morphotype is not selected for. Second, changes in the ability to form biofilm were environment-dependent, and thus could not explain the emergence of rdar morphotype in M02. The only trait which was consistent across environments and different across capsulated and non-capsulated populations was aggregation, suggesting this was the major selective force. Indeed, the fitness benefits of different *mrkH* mutations in capsulated clones are slightly less than in non-capsulated genetic backgrounds, especially in urine media (Fig. S7C).

Then, why are non-aggregative phenotypes selected during the evolution experiment? Ancestral-like clones express T3F leading to clump formation and sedimentation. This can also lead to surface adhesion and biofilm formation at the bottom of the microtiterplate well. Rdar-like mutants have evolved to decrease expression of T3F, reduce costs associated to fimbriae production and grow faster (Fig. 2). We foresee two main advantages for these 'escape mutants'. First, our selection regime included vigorous pipetting before each transfer. Thus, larger, more cohesive cell clumps would be excluded in benefit of smaller, or isolated cells. Second, non-aggregative variants remain in suspension and access other resources, such as increased oxygen concentrations, which may be lacking at the bottom of the microtiterplate-well

where other aggregative clones may have fallen. Within a host, we speculate rdar morphotypes at low frequencies could favour dispersal of *Klebsiella* cells from a population during infectious episodes. It was recently shown that in the urinary tract, non-capsulated variants are often recovered (Ernst et al. 2020). Our observations that *mrkH* mutants are more often found in the blood and in urine suggests that these mutations constitute the first step for the evolution of rdar-like phenotypes prior to systemic dispersion by the blood, like the clinical isolate PBIO3459 (Sydow et al. 2022).

Our genomic analyses reveal that *mrkH* truncations and the resulting rough phenotypes are mostly associated to strains that do not encode the mucoidy regulator RmpA. This would suggest specific physical interactions between T3F and a mucoid capsule which would limit the fitness advantage of rdar phenotypes. It could also be that in these strains, T3F is regulated differently. In the specific case of hypervirulent strains, characterized by the presence of the virulence plasmid encoding *rmpA*, but also other siderophores like salmochelin (*iro* operon) and aerobactin (*iuc* operon), it was recently shown that T3F and capsule are both tightly regulated by *iroP* also present in the virulence plasmid (Chu et al. 2023). IroP acts as a switch, by which T3F is repressed during capsule expression. Such regulation was not observed in classical *K. pneumoniae*. If mucoidy precludes selection of *mrkH* mutations, this would also explain why these are more commonly observed in strains isolated from urine. A recent study showed that urine suppresses mucoidy by repression the *rmpA* regulator, *rmpD* (Khadka et al. 2023), which ultimately alters capsule chain length (Ovchinnikova et al. 2023) without affecting total amount of capsule production *per se*. Overall, these results are in line with previous studies in which changes in the capsule, either in physical properties or in overall presence, have a major impact in Klebsiella evolution (Nucci et al. 2022), and further suggest that classical *K. pneumoniae* could follow evolutionary trajectories more similar to *K. variicola* than to hypermucoid *K. pneumoniae*.

In artificial urine, during competition experiments between rdar-like and wild type clones, we observed that even when inoculated in a minority, rdar-like clones not only increased their relative frequency, indicating higher fitness (Fig. 4C, D and S4), but they reached a majority (frequency > 0.5). This, together with the negative-frequency dependence advantage of the phenotype, could be suggestive of cheating. Cheating can be defined as a fitness relationship where a strain which performs poorly at a social trait in pure culture exploits another strain with high performance at the focal trait (Travisano and Velicer 2004, Smith and Schuster 2019). The effect of mixing would result in a relative within-group fitness advantage of the low-performance strain. This can be explained because the cheater cells do not pay their fair share of the cooperative act. Well-known cooperative acts are production of siderophores (West and Buckling 2003), or fruiting body formation during multicellular development (Velicer and Vos 2009). Here, if we consider that the main selective pressure in our experiment is aggregation and production of extracellular adhesins, *mrkH* mutants are unable to properly aggregate in isolation and could thus potentially behave as *bona fide* cheaters.

Numerous evolution experiments performed in static environments or with some degree of spatial structure have revealed a remarkable morphotypic convergence in the emergence of rough, 'wrinkly' or rdar-like colonies. Formation of such colonies has been attributed to changes in exopolysaccharides (Spiers et al. 2002, Lin et al. 2022), or changes in expression of amyloid fibers like curli. Most of these morphotypes rely on regulatory pathways that respond to c-di-GMP levels (Bantinaki et al. 2007). Indeed, this second messenger has been identified as an important determinant of diversification in *P. aeruginosa* biofilms (Flynn et al. 2016). Similarly, in *Vibrio cholerae*, c-di-GMP regulates a switch between motility driven by the flagella and biofilm formation, via regulation of the Vibrio polysaccharide, *vps* (Wu et al. 2020). This is similar to the abovementioned switch in hypervirulent *K. pneumoniae* between fimbriae and capsule (Chu et al. 2023). Lastly, rdar-like or wrinkly morphotypes are very often under negative frequency-dependent selection (Spiers et al. 2002, Poltak and Cooper 2011, Udall et al. 2015). Indeed, they are stably present at low frequencies in the population, and their benefit is highest when they are rare. Such morphotypes in *K. variicola* may also respond to c-di-GMP (via *mrkH*, (Wilksch et al. 2011)). Here, we show that they emerge by mutations in pathways independent of c-di-GMP (*nac*), and that their negative-frequency dependence is environment-specific. *K. variicola* morphotypes emerge independent of EPS or amyloid fiber production, but rather in changes in type 3 fimbriae. Most importantly, in most species these morphotypes are strongly aggregative allowing efficient colonisation of air-liquid interface (Poltak and Cooper 2011, Udall et al. 2015, Blake et al. 2021) but in *K. variicola,* biofilm formation and aggregation is strongly impaired. Hence, even if these morphotypes are a common adaptive strategy across Bacteria, as shown by the incredible parallel evolution across microbial systems, they do not seem to be functionally convergent, as they seem to impact differently the ability to form biofilm.

This work further contributes to highlight the important role of insertion sequences in shaping genomes (Siguier et al. 2014). For long, transposition of IS was associated to fitness loss due to gene inactivation or expression changes. Here, we provide more evidence that IS can increase fitness and play an important role in adaptation, as was already shown across different species ranging from Enterobacteria (Consuegra et al. 2021, Frazão et al. 2022) to Cyanobacteria (Miller et al. 2021). A recent study in *Bacillus subtilis* and *Bacillus thuringensis* revived end-point populations from four different evolution experiments and showed that, independently of growth in either biotic or abiotic conditions, insertion sequences are critical in adaptation (Hu et al. 2023). No differences were found across *Bacillus* species. Such conservation across the *Bacillus* genus contrasts with our previous studies in *Klebsiella*, in which we showed that evolution by insertion sequences was strain-dependent. More specifically, it relied on the presence of specific IS families, including IS903B, belonging to the IS5 family also found in Kva 342 (Nucci et al. 2022). IS903B from IS5 family is known to insert in *mgrB* and cause resistance to colistin, a last-resort antimicrobial peptide (Fordham et al. 2022). Here, we observe that IS903B action can also extend to other morphological diversification mediated surface structures like fimbriae. Further, we reveal that IS1 family, which also codes for a DDE transposase, like IS5, is also a primary driver of *Klebsiella* adaptation and its morphological diversification. Genomic analyses highlighted similar trends are also observed in *Klebsiella* natural populations. We showed that at least 0.4% (3 out of 676) *K. variicola* and 1.4% (133 out of 9532) *K. pneumoniae* MrkH proteins were interrupted by an IS. This is most likely an underestimation as genes encoding MrkH were occasionally located at the border of the contig and a direct association to an IS insertion could not be determined. Yet, contig breaks are oftentimes an indication of an IS insertion, as we confirmed in the three different strains from the Klebsiella Diversity Panel.

Taken together, our work provides important insight into the biology and evolution of *K. variicola,* a plant endosymbiont but also an emerging pathogen in cattle and humans, which has been

significantly less studied than *K. pneumoniae sensu stricto*. Along our previous findings that one mutation in Kva 342 results in the *de novo* emergence of hypermucoidy (a proxy of hypervirulence) even in the absence of an immune system (Nucci et al. 2022), this work further highlights the complexity of cell-to-cell and cell-to-surface interactions and how these evolve. Given the renewed interest for microbial products as fertilizers in agricultural settings to enable bacterial nitrogen fixation (Wen et al. 2021), and particularly of *K. variicola*, our study also contributes to critically advance our understanding of how this species evolves and responds to different environments, a requirement prior to any commercial exploitation.

Materials and methods

Bacterial strains and growth conditions

(i) *Strain. K. variicola* 342 (Kva 342, serotype K30) is a non-mucoid, mildly capsulated environmental strain, isolated from maize in the USA (Fouts et al. 2008). The *K. pneumoniae* Diversity Panel was acquired at BEI resources (https://www.beiresources.org/) and is available for research purposes under catalogue #NR-55604. Strains in this panel were previously characterized in (Martin et al. 2023). (ii) *Growth media*. AUM (artificial urine medium) was prepared as described previously (Brooks and Keevil 1997). AUM is mainly composed of 1% urea and 0.1% peptone with trace amounts of lactic acid, uric acid, creatinine and peptone. Artificial Sputum Medium (ASM) is composed of 0.5% mucin, 0.4% DNA, 0.5% egg yolk and 0.2% amino acids. Lysogeny Broth (LB) is composed of 1% tryptone, 1% NaCl and 0.5% yeast extract. M02 corresponds to minimal M63B1 supplemented with 0.2% of glucose as a sole carbon source. (iii) *Evolution experiment*. The evolution experiment was previously described (Nucci et al. 2022). Briefly, six ancestral genotypes (Kva 342, Kpn NTUH and Kpn BJ1, and their respective non-capsulated mutants) were evolved in parallel for 100 days, accounting for *ca* 675 generations. To initiate the experiment, a single colony of each ancestral genotype was inoculated in 5 mL of LB and allowed to grow under shaking conditions at 37°C overnight. Twenty microliters of the diluted (1:100) overnight culture were used to inoculate each of the six independent replicates in the five environments. Each population was grown in a final volume of 2 mL in independent wells of 24 well microtiter plates. Every 24 hours, 20 μL of each culture was propagated into 1980 μL of fresh media and grown for 37°C under static conditions. Although each growth media had different carrying capacities, *i.e.* the maximum population size an environment can sustain, all cultures reached bacterial saturation in late stationary phase, ensuring that the different populations underwent a similar number of generations across media. Independently evolving populations were plated 28 times, every second day for ten days and every four days until the end of the experiment. Cross contamination checks were routinely performed (Nucci et al. 2022). (iv) *Growth conditions*. Unless stated otherwise, experiments performed in the evolutionary environment were initiated by an overday culture in LB prior to an overnight culture in evolutionary media under shaking conditions. Pre-grown cultures were then diluted 1:100 into 1980 μL 24-well plates and allowed to grow for 24 hours at 37° under static conditions. (v) *Primers*. Primers used in this study are listed in Table S5.

Frequency of rdar-like morphotypes

To test the frequencies of rdar-like morphotypes in the evolved populations, we aliquoted the glycerol stocks belonging to days 7, 15, 30, 45, 75, and 100. Stocks were serially diluted in LB and appropriate dilutions were plated on LB agar plates for CFU counting (at least three plates for each population). The number of rdar colonies as well as the total number of CFUs was quantified and the proportion of rdar-like clones computed.

Whole genome sequencing and variant analyses

A single rdar clone from each population was isolated for whole genome sequencing. DNA was extracted from pelleted cells grown overnight in LB supplemented with 0.7 mM EDTA with the guanidium thiocyanate method. Extra RNAse A treatment (37°C, 30 min) was performed before DNA precipitation. Each clone was sequenced by Illumina with 150pb paired-end reads. Each evolved clone was compared to ancestral sequence using *breseq* (0.30.1) (Deatherage and Barrick 2014) with default parameters. The SNPs identified by *breseq* were further confirmed using *snippy* (https://github.com/tseemann/snippy) with default parameters. When generating appropriate mutants, specific PCR and Sanger sequencing confirmed that these mutations were present in the evolved clones.

Mutant construction

Isogenic mutants were constructed by allelic exchange. (i) *mrkH deletion*. 500 bp upstream and downstream of the gene of interest were amplified by PCR. Cloning vector pKNG101 plasmid was also amplified using Phusion Master Mix (Thermo Scientific). Afterwards, pKNG101 was digested by DpnI (NEB BioEngland) restriction enzyme for 30 minutes at 37°C. Inserts and vector were then assembled using the GeneArt™ Gibson Assembly HiFi kit (Invitrogen) for 30 minutes at 50°C. The reaction was dyalised and electroporated into competent *E. coli* DH5α strain and selected on Streptomycin LB plates (100 μg/mL for *E. coli*). Correct assemblies were checked by PCR. pKNG101 containing insert of interest was extracted using the QIAprep Spin Miniprep Kit then electroporated again into *E. coli* MFD λ-pir strain, used as a donor strain for conjugation in strains of interest. Single cross-over mutants (transconjugants) were selected on Streptomycin plates (200 μg/mL for *Klebsiella*) and double cross-over mutants were selected on LB without salt, supplemented with 5% sucrose, at room temperature. From each double-recombination, a mutant and a wild-type were isolated. Mutants were verified for their sensitivity to Streptomycin and by Sanger sequencing. (ii) *Insertion and reversion* of evolved alleles was done in ancestor and evolved clone, respectively. The gene of interest (ancestor or evolved allele) was amplified using Phusion Master Mix (Thermo Scientific) and cloned into pKNG101 vector as described above. All mutants generated and used in this study are listed in Table S6.

Biofilm formation

The capacity of a population or isolated clones to form a biofilm was performed as previously described (Buffet et al. 2021). Briefly, each population or clones was pre-conditioned by allowing growth in LB overday, prior to inoculating overnight cultures in the environments in which the populations or clones evolved. Then, 20 μL of each overnight culture was inoculated into 1980 μL in 24-well microtiter plates and allowed to grow for 24 hours without shaking at 37°C. Unbound cells were removed by washing once in distilled water. To stain biofilms, 2100 μL of 1% crystal violet was added to each well for 20 minutes. The crystal violet was decanted and washed thrice with distilled water. The plates were allowed to dry under a laminar flow hood. Then, the biofilm was solubilized for 10 min in 2300 μL of mix with 80% ethanol and 20% acetone.

About 200 μL of each mix was transferred in a well of a 96-well plate. The absorbance of the sample was read at OD590.

Aggregation test

An isolated colony was allowed to grow in 5 mL overnight in M02 medium at 37° under shaking conditions. Prior to the experiment, the absorbance (OD_{600nm}) was measured and adjusted to OD_{600} = 2, and the cultures were transferred to static test tubes. Two hundred μL samples were sampled, and the absorbance (OD_{600nm}) was measured at defined time points (0; 1,5; 3; 4,5 and 24 hours) using an automatic plate reader Spark Control Magellan (TECAN). Samples were removed from the uppermost layer of tube cultures, roughly at the 4 mL mark. Decreasing absorbance represents the settling of agglutinated cell clumps. Calculation of the area under the aggregation curve with the function *trapz* from the R package pracma results in qualitatively similar interpretations. Values above 1 represent decreased aggregation compared to the ancestor.

Fitness of *mrkH* mutants

To estimate the fitness advantage of mutations in *mrkH*, we performed direct competitions between the deletion, insertion and reversion mutants and their respective associated wild types. Additionally, competitions between 6B3 and 4D6 evolved clones and the ancestor were performed. To initiate the competition experiments, individual clones were grown overnight in LB, and mixed in a 1:1 or 1:9 proportions. An aliquot was taken to estimate the initial ratio of each genotype by serial dilution and CFU counting as control of T_0. The co-culture was then diluted 1:100 in 2 mL in the evolutionary environment in which the mutations emerged (AUM & M02). After 24 h of competition (T_{24}) in 24-well microplate plates under static conditions, each culture was re-homogenized by vigorous pipetting and then serially diluted and plated. Rdar-like and wild-type colonies are clearly differentiated visually and counted separately. The competitive index of each genotype was calculated using their ratio $\frac{T24}{T0}$. Competitions were only taken into account if initial frequencies for rdar-morphotype were in between 0.4 and 0.6 for 1:1 competitions and in between 0.02 and 0.18 for 1:9.

Search for curli and cellulose operons

The proteic sequences of the experimentally validated curli biogenesis apparatus and the cellulose operon, both from *E. coli* were downloaded (accession numbers: Table S1). BlastP (v2.7.1+) with default parameters was used to search for each protein in the Kva 342 proteome. (i) *Curli*. No hits were obtained (E-value < 10^{-5} & identity > 60%). (ii) *Cellulose*. Each protein searched matched (E-value < 10^{-5} & identity > 60%) a protein in Kva 342 genomes. Further, these proteins were found in consecutive positions in the genome (*bcsGFEQABZC*).

Search for MrkH proteins

The sequences for the *K. pneumoniae* Diversity Panel were downloaded on June 12, 2023, from NCBI under BioProject PRJNA717739. All genomes corresponding to *K. variicola* in the Pathosystems Resource Integration Center (PATRIC) genome database (Wattam et al. 2014), filtered by good quality and text mined for *K. variicola* species (751 out of the 767), were downloaded on March 11 2022. Same procedure was applied for the first 10 000 genomes of *K. pneumoniae* (of which 239 were discarded). The genomes were checked for quality control and annotated with the pipeline PaNaCoTa (Perrin and Rocha 2021) and the –prodigal option. Protein-Protein Blast (BLAST 2.7.1+) against either the Kva 342 MrkH was performed with the following option -max_target_seqs 100 000 and an E-value smaller than 10^{-5}. Sequences with an identity percentage of less than 90% were discarded. Reducing the threshold to 80% did not alter the number of sequences discarded.

Search for RmpA proteins

The genomic databases (PATRIC -downloaded on March 11 2022-, and the Klebsiella Diversity Panel) were mined for the presence of RmpA using RmpA protein sequence of Kpn NTUH-K2044 as reference (Accession number: CDO11653.1). Protein-Protein Blast (BLAST 2.7.1+) searches were performed with the following option -max_target_seqs 10 000 and an e-value smaller than 10^{-5}. Hits were found in 459 genomes. Hits were further filtered to exclude proteins with an identity of less than 75% or shorter than 70% of the query protein length (156 out of 575). This excluded 67 genomes, resulting in 391 genomes with *rmpA* (4.1%).

Quantitative RT-PCR

Twenty-four-hour cultures in 24-well plates in the respective evolution environment were grown as abovementioned. RNA was extracted with the miRNA Extraction Kit (Macherey Nagel). Total RNA was measured and 400 ng were used as matrix for cDNA amplification using the iScript cDNA Synthesis Kit. Samples were then treated for 30 minutes with DNase. Quantitive RT-PCR was performed using 1.5 μL of amplified cDNA in a total volume of 15 μL. As suggested by Gomes et al (Gomes et al. 2018), we used *rho* and *recA* as housekeeping genes. Data presented is based on calculations taking into account the geometric average of the two housekeeping genes. Taking one housekeeping gene or the other individually does not qualitatively alter any of the results.

Author contributions

OR conceived and designed the details of the study. AN, JJ and OR performed the experiments. AN and OR performed statistical analysis. OR performed the bioinformatics work, analyzed the data and wrote the manuscript. OR and EPCR secured funding, provided the resources and materials necessary for this study and revised the manuscript. All authors approved the final version of the manuscript.

Acknowledgements

The authors would like to thank Katharina Schaufler and Stefan Heiden for providing the genomic sequence of strain PBIO3459 and Samay Pande for helpful discussions during the writing of the manuscript. We are also grateful for the constructive comments of Ákos Kovács and two anonymous reviewers. The sequencing of clones 4D4 and 6B4 was made at the Biomics Platform, C2RT, Institut Pasteur, Paris, France, supported by France Génomique (ANR-10-INBS-09) and IBISA.

Supplementary data

Supplementary data is available at *FEMSML Journal* online.

Conflict of interest: Authors declare that we do not have any competing financial interests in relation to the work described.

Funding

This work was funded by an ANR JCJC (Agence national de recherche) grant [ANR 18 CE12 0001 01 ENCAPSULATION] awarded to O.R. The laboratory is funded by a Laboratoire d'Excellence 'Integrative Biology of Emerging Infectious Diseases' (grant ANR-10-LABX-62-IBEID) and the FRM [EQU201903007835]. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Data availability

Raw reads are available at the European Nucleotide Archive (ENA), project number PRJEB54810. The sample names for rdar-like clones 4D4, 4D6 and 6B3 and 6B4 are ERS15562389, ERS12546754, ERS12546780 and ERS15562390 respectively.

All raw data generated in this study has been deposited in the public repository Figshare https://doi.org/10.6084/m9.figshare.23268791.

References

Abdul-Rahman F, Tranchina D, Gresham D. Fluctuating environments maintain genetic diversity through neutral fitness effects and balancing selection. *Mol Biol Evol* 2021;**38**:4362–75. https://doi.org/10.1093/molbev/msab173.

Ares MA, Fernández-Vázquez JL, Pacheco S *et al.* Additional regulatory activities of MrkH for the transcriptional expression of the *Klebsiella pneumoniae* mrk genes: antagonist of H-NS and repressor. *PLoS One* 2017;**12**:e0173285. https://doi.org/10.1371/journal.pone.0173285.

Bantinaki E, Kassen R, Knight CG *et al.* Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity. *Genetics* 2007;**176**:441–53. https://doi.org/10.1534/genetics.106.069906.

Baquero F, Coque TM, Galán JC *et al.* The origin of niches and species in the bacterial world. *Front Microbiol* 2021;**12**:657986. https://doi.org/10.3389/fmicb.2021.657986.

Barrios-Camacho H, Aguilar-Vera A, Beltran-Rojel M *et al.* Molecular epidemiology of *Klebsiella variicola* obtained from different sources. *Sci Rep* 2019;**9**:10610. https://doi.org/10.1038/s41598-019-46998-9.

Baselga-Cervera B, Jacobsen KA, Ford Denison R *et al.* Experimental evolution in the *Cyanobacterium trichormus; variabilis*: increases in size and morphological diversity. *Evol; Intern J Organic Evol* 2023;qpad037. https://doi.org/10.1093/evolut/qpad037.

Batarseh TN, Batarseh SN, Rodríguez-Verdugo A *et al.* Phenotypic and genotypic adaptation of *Escherichia coli* to thermal stress is contingent on genetic background. *Mol Biol Evol* 2023;**40**:msad108. https://doi.org/10.1093/molbev/msad108.

Bender RA. The role of the NAC protein in the nitrogen regulation of *Klebsiella aerogenes*. *Mol Microbiol* 1991;**5**:2575–80. https://doi.org/10.1111/j.1365-2958.1991.tb01965.x.

Blake C, Nordgaard M, Maróti G *et al.* Diversification of Bacillus subtilis during experimental evolution on Arabidopsis thaliana and the complementarity in root colonization of evolved subpopulations. *Environ Microbiol* 2021;**23**:6122–36. https://doi.org/10.1111/1462-2920.15680.

Blount ZD, Barrick JE, Davidson CJ *et al.* Genomic analysis of a key innovation in an experimental Escherichia coli population. *Nature* 2012;**489**:513–8. https://doi.org/10.1038/nature11514.

Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proc Nat Acad Sci USA* 2008;**105**:7899–906. https://doi.org/10.1073/pnas.0803151105.

Brooks T, Keevil CW. A simple artificial urine for the growth of urinary pathogens. *Lett Appl Microbiol* 1997;**24**:203–6. https://doi.org/10.1046/j.1472-765x.1997.00378.x.

Buffet A, Rocha EPC, Rendueles O. Nutrient conditions are primary drivers of bacterial capsule maintenance in Klebsiella. *Proceed Biolog Sci* 2021;**288**:20202876. https://doi.org/10.1098/rspb.2020.2876.

Chu WHW, Tan YH, Tan SY *et al.* Acquisition of regulator on virulence plasmid of hypervirulent Klebsiella allows bacterial lifestyle switch in response to iron. *Mbio* 2023;**0**:e01297–23. https://doi.org/10.1128/mbio.01297-23.

Cimdins A, Simm R, Li F *et al.* Alterations of c-di-GMP turnover proteins modulate semi-constitutive rdar biofilm formation in commensal and uropathogenic *Escherichia coli*. *MicrobiologyOpen* 2017;**6**:e00508. https://doi.org/10.1002/mbo3.508.

Comins HN, Hassell MP. Persistence of multispecies host-parasitoid interactions in spatially distributed models with local dispersal. *J Theor Biol* 1996;**183**:19–28. https://doi.org/10.1006/jtbi.1996.0197.

Consuegra J, Gaffé J, Lenski RE *et al.* Insertion-sequence-mediated mutations both promote and constrain evolvability during a long-term experiment with bacteria. *Nat Commun* 2021;**12**:980. https://doi.org/10.1038/s41467-021-21210-7.

Czárán TL, Hoekstra RF, Pagie L. Chemical warfare between microbes promotes biodiversity. *Proc Nat Acad Sci USA* 2002;**99**:786–90. https://doi.org/10.1073/pnas.012399899.

D'souza G, Shitut S, Preussger D *et al.* Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Nat Prod Rep* 2018;**35**:455–88. https://doi.org/10.1039/c8np00009c.

Da Re S, Ghigo J-M. A CsgD-independent pathway for cellulose production and biofilm formation in *Escherichia coli*. *J Bacteriol* 2006;**188**:3073–87. https://doi.org/10.1128/JB.188.8.3073-3087.2006.

De Paepe M, Gaboriau-Routhiau V, Rainteau D *et al.* Trade-off between bile resistance and nutritional competence drives *Escherichia coli* diversification in the mouse gut. *PLos Genet* 2011;**7**:e1002107. https://doi.org/10.1371/journal.pgen.1002107.

Deatherage DE, Barrick JE. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. in Sun L., Shou W. (eds.), *Engineering and Analyzing Multicellular Systems: Methods and Protocols*. New York, NY: Springer (Methods in Molecular Biology): 2014;165–88. https://doi.org/10.1007/978-1-4939-0554-6_12.

Debray R, De Luna N, Koskella B. Historical contingency drives compensatory evolution and rare reversal of phage resistance. *Mol Biol Evol* 2022;**39**:msac182. https://doi.org/10.1093/molbev/msac182.

Di Sante L, Pugnaloni A, Biavasco F *et al.* Multicellular behavior of environmental *Escherichia coli* isolates grown under nutrient-poor and low-temperature conditions. *Microbiol Res* 2018;**210**:43–50. https://doi.org/10.1016/j.micres.2018.03.004.

Ernst CM, Braxton JR, Rodriguez-Osorio CA *et al.* Adaptive evolution of virulence and persistence in carbapenem-resistant *Klebsiella pneumoniae*. *Nat Med* 2020;**26**:705–11. https://doi.org/10.1038/s41591-020-0825-4.

Flynn KM, Dowell G, Johnson TM *et al.* Evolution of ecological diversity in biofilms of Pseudomonas aeruginosa by altered cyclic diguanylate signaling. *J Bacteriol* 2016;**198**:2608–18. https://doi.org/10.1128/JB.00048-16.

Fordham SME, Mantzouratou A, Sheridan E. Prevalence of insertion sequence elements in plasmids relating to mgrB gene disruption causing colistin resistance in *Klebsiella pneumoniae*. *MicrobiologyOpen* 2022;**11**:e1262. https://doi.org/10.1002/mbo3.1262.

Fouts DE, Tyler HL, Deboy RT *et al.* Complete genome sequence of the N2-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. *PLos Genet* 2008;**4**:e1000141. https://doi.org/10.1371/journal.pgen.1000141.

Frazão N, Konrad A, Amicone M *et al.* Two modes of evolution shape bacterial strain diversity in the mammalian gut for thousands of generations. *Nat Commun* 2022;**13**:5604. https://doi.org/10.1038/s41467-022-33412-8.

Frisch RL, Bender RA. Properties of the NAC (Nitrogen Assimilation Control Protein)-binding site within the ureD promoter of *Klebsiella pneumoniae*. *J Bacteriol* 2010;**192**:4821–6. https://doi.org/10.1128/JB.00883-09.

Garza-Ramos U, Silva-Sánchez J, Martínez-Romero E *et al.* Development of a multiplex-PCR probe system for the proper identification of *Klebsiella variicola*. *BMC Microbiol* 2015;**15**:64. https://doi.org/10.1186/s12866-015-0396-6.

Giannattasio-Ferraz S, Ene A, Johnson G *et al.* Multidrug-resistant Klebsiella variicola isolated in the urine of healthy bovine heifers, a potential risk as an emerging Human pathogen. *Appl Environ Microbiol* 2022;**88**:e00044–22. https://doi.org/10.1128/aem.00044-22.

Gligorijević V, Renfrew PD, Kosciolek T *et al.* Structure-based protein function prediction using graph convolutional networks. *Nat Commun* 2021;**12**:3168. https://doi.org/10.1038/s41467-021-23303-9.

Gomes AÉI, Stuchi LP, Siqueira NMG *et al.* Selection and validation of reference genes for gene expression studies in *Klebsiella pneumoniae* using Reverse Transcription Quantitative real-time PCR. *Sci Rep* 2018;**8**:9001. https://doi.org/10.1038/s41598-018-27420-2.

Gómez P, Buckling A. Real-time microbial adaptive diversification in soil. *Ecol Lett* 2013;**16**:650–5. https://doi.org/10.1111/ele.12093.

Habets MGJL, Rozen DE, Hoekstra RF *et al.* The effect of population structure on the adaptive radiation of microbial populations evolving in spatially structured environments. *Ecol Lett* 2006;**9**:1041–8. https://doi.org/10.1111/j.1461-0248.2006.00955.x.

Haudiquet M, Buffet A, Rendueles O *et al.* Interplay between the cell envelope and mobile genetic elements shapes gene flow in populations of the nosocomial pathogen *Klebsiella pneumoniae*. *PLoS Biol* 2021;**19**:e3001276. https://doi.org/10.1371/journal.pbio.3001276.

Hedrick PW. Balancing selection. *Curr Biol* 2007;**17**:R230–1. https://doi.org/10.1016/j.cub.2007.01.012.

Hu G, Wang Y, Liu X *et al.* Species and condition shape the mutational spectrum in experimentally evolved biofilms. 2023; 2022.12.07.519423. https://doi.org/10.1101/2022.12.07.519423. bioRxiv

Jones SE, Lennon JT. Dormancy contributes to the maintenance of microbial diversity. *Proc Nat Acad Sci USA* 2010;**107**:5881–6. https://doi.org/10.1073/pnas.0912765107.

Kassen R, Rainey PB. The ecology and genetics of microbial diversity. *Annu Rev Microbiol* 2004;**58**:207–31. https://doi.org/10.1146/annurev.micro.58.030603.123654.

Kassen R. *Experimental Evolution and the Nature of Biodiversity*. Greenwood Village, Colorado,USA: Roberts & Company Publishers. 2014.

Kassen R. The experimental evolution of specialists, generalists, and the maintenance of diversity. *J Evol Biol* 2002;**15**:173–90. https://doi.org/10.1046/j.1420-9101.2002.00377.x.

Kerr B, Riley MA, Feldman MW *et al.* Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* 2002;**418**:171–4. https://doi.org/10.1038/nature00823.

Khadka S, Ring BE, Walker RS *et al.* Urine-mediated suppression of Klebsiella pneumoniae mucoidy is counteracted by spontaneous Wzc variants altering capsule chain length. *mSphere* 2023;e0028823. https://doi.org/10.1128/msphere.00288-23.

La Fortezza M, Rendueles O, Keller H *et al.* Hidden paths to endless forms most wonderful: ecology latently shapes evolution of multicellular development in predatory bacteria. *Commun Biol* 2022;**5**:977. https://doi.org/10.1038/s42003-022-03912-w.

Lam MMC, Wick RR, Watts SC *et al.* A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex. *Nat Commun* 2021;**12**:4188. https://doi.org/10.1038/s41467-021-24448-3.

Lemonnier M, Levin BR, Romeo T *et al.* The evolution of contact-dependent inhibition in non-growing populations of Escherichia coli. *Proc Biolog Sci* 2008;**275**:3–10. https://doi.org/10.1098/rspb.2007.1234.

Lin Y, Xu X, Maróti G *et al.* Adaptation and phenotypic diversification of *Bacillus thuringiensis* biofilm are accompanied by fuzzy spreader morphotypes. *NPJ Biofilms and Microbiomes* 2022;**8**:1–14. https://doi.org/10.1038/s41522-022-00292-1.

Martin MJ, Stribling W, Ong AC *et al.* A panel of diverse *Klebsiella pneumoniae* clinical isolates for research and development. *Microbial Genomics* 2023;**9**:000967. https://doi.org/10.1099/mgen.0.000967.

Martínez-Romero E, Rodríguez-Medina N, Beltrán-Rojel M *et al.* Genome misclassification of *Klebsiella variicola* and *Klebsiella quasipneumoniae* isolated from plants, animals and humans. *Salud Publica Mex* 2018;**60**:56–62. https://doi.org/10.21149/8149.

Miller SR, Abresch HE, Ulrich NJ *et al.* Bacterial adaptation by a transposition burst of an invading IS element. *Genome Biol Evol* 2021;**13**:evab245. https://doi.org/10.1093/gbe/evab245.

Nassif X, Fournier JM, Arondel J *et al.* Mucoid phenotype of Klebsiella pneumoniae is a plasmid-encoded virulence factor. *Infect Immun* 1989;**57**:546–52. https://doi.org/10.1128/iai.57.2.546-552.1989.

Nucci A, Rocha EPC, Rendueles O. Adaptation to novel spatially-structured environments is driven by the capsule and alters virulence-associated traits. *Nat Commun* 2022;**13**:4751. https://doi.org/10.1038/s41467-022-32504-9.

Nucci A, Rocha EPC, Rendueles O. Latent evolution of biofilm formation depends on life-history and genetic background. *NPJ Biofilms Microbiomes* 2023;**9**:53. https://doi.org/10.1038/s41522-023-00422-3. PMID: 37537176.

Ovchinnikova OG, Treat LP, Teelucksingh T *et al.* Hypermucoviscosity regulator RmpD interacts with Wzc and controls capsular polysaccharide chain length. *Mbio* 2023;**14**:e0080023. https://doi.org/10.1128/mbio.00800-23.

Perrin A, Rocha EPC. PanACoTA: a modular tool for massive microbial comparative genomics. *NAR Genomics Bioinformatics* 2021;**3**:lqaa106. https://doi.org/10.1093/nargab/lqaa106.

Pinto-Tomás AA, Anderson MA, Suen G *et al.* Symbiotic nitrogen fixation in the fungus gardens of leaf-cutter ants. *Science* 2009;**326**:1120–3. https://doi.org/10.1126/science.1173036.

Poltak SR, Cooper VS. Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *ISME J* 2011;**5**:369–78. https://doi.org/10.1038/ismej.2010.136.

Potter RF, Lainhart W, Twentyman J *et al.* Population structure, antibiotic resistance, and uropathogenicity of *Klebsiella variicola*. *Mbio* 2018;**9**:e02481–18. https://doi.org/10.1128/mBio.02481-18.

Rainey PB, Buckling A, Kassen R *et al.* The emergence and maintenance of diversity: insights from experimental bacterial populations. *Trends Ecol Evol* 2000;**15**:243–7. https://doi.org/10.1016/S0169-5347(00)01871-1.

Rainey PB, Travisano M. Adaptive radiation in a heterogeneous environment. *Nature* 1998;**394**:69–72. https://doi.org/10.1038/27900.

Rendueles O, Amherd M, Velicer GJ. Positively frequency-dependent interference competition maintains diversity and pervades a natural population of cooperative microbes. *Curr Biol* 2015;**25**:1673–81. https://doi.org/10.1016/j.cub.2015.04.057.

Rendueles O, Velicer GJ. Evolution by flight and fight: diverse mechanisms of adaptation by actively motile microbes. *ISME J* 2017;**11**:555–68. https://doi.org/10.1038/ismej.2016.115.

Rendueles O, Velicer GJ. Hidden paths to endless forms most wonderful: complexity of bacterial motility shapes diversification of latent phenotypes. *BMC Evol Biol* 2020;**20**:145. https://doi.org/10.1186/s12862-020-01707-3.

Rodrigues C, Passet V, Rakotondrasoa A *et al.* Identification of *Klebsiella pneumoniae*, *Klebsiella quasipneumoniae*, *Klebsiella variicola* and related phylogroups by MALDI-TOF mass spectrometry. *Front Microbiol* 2018;**9**:3000. https://doi.org/10.3389/fmicb.2018.03000.

Rodríguez-Medina N, Barrios-Camacho H, Duran-Bedolla J *et al. Klebsiella variicola*: an emerging pathogen in humans. *Emerg Microbes Infect* 2019;**8**:973–88. https://doi.org/10.1080/22221751.2019.1634981.

Römling U. Characterization of the rdar morphotype, a multicellular behaviour in Enterobacteriaceae. *Cell Mol Life Sci CMLS* 2005;**62**:1234–46. https://doi.org/10.1007/s00018-005-4557-x.

Rosenblueth M, Martínez L, Silva J *et al. Klebsiella variicola*, a novel species with clinical and plant-associated isolates. *Syst Appl Microbiol* 2004;**27**:27–35. https://doi.org/10.1078/0723-2020-00261.

Schembri MA, Blom J, Krogfelt KA *et al.* Capsule and fimbria interaction in *Klebsiella pneumoniae*. *Infect Immun* 2005;**73**:4626–33. https://doi.org/10.1128/IAI.73.8.4626-4633.2005.

Siguier P, Gourbeyre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev* 2014;**38**:865–91. https://doi.org/10.1111/1574-6976.12067.

Siguier P. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 2006;**34**:D32–36. https://doi.org/10.1093/nar/gkj014.

Smith P, Schuster M. Public goods and cheating in microbes. *Curr Biol* 2019;**29**:R442–7. https://doi.org/10.1016/j.cub.2019.03.001.

Spencer CC, Tyerman J, Bertrand M *et al.* Adaptation increases the likelihood of diversification in an experimental bacterial lineage. *Proc Nat Acad Sci USA* 2008;**105**:1585–9. https://doi.org/10.1073/pnas.0708504105.

Spiers AJ, Kahn SG, Bohannon J *et al.* Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics* 2002;**161**:33–46. https://doi.org/10.1093/genetics/161.1.33.

Sydow K, Eger E, Schwabe M *et al.* Geno- and phenotypic characteristics of a *Klebsiella pneumoniae* ST20 isolate with unusual colony morphology. *Microorganisms* 2022;**10**:2063. https://doi.org/10.3390/microorganisms10102063.

Travisano M, Vasi F, Lenski RE. Long-term experimental evolution in *Escherichia coli*. III. Variation among replicate populations in correlated responses to novel environments. *Evol; Intern J Organic Evol* 1995;**49**:189–200. https://doi.org/10.1111/j.1558-5646.1995.tb05970.x.

Travisano M, Velicer GJ. Strategies of microbial cheater control. *Trends Microbiol* 2004;**12**:72–78. https://doi.org/10.1016/j.tim.2003.12.009.

Travisano M. Long-term experimental evolution in *Escherichia coli*. VI. Environmental constraints on adaptation and divergence. *Genetics* 1997;**146**:471–9. https://doi.org/10.1093/genetics/146.2.471.

Tyerman J, Havard N, Saxer G *et al.* Unparallel diversification in bacterial microcosms. *Proc Biolog Sci* 2005;**272**:1393–8. https://doi.org/10.1098/rspb.2005.3068.

Udall YC, Deeni Y, Hapca SM *et al.* The evolution of biofilm-forming wrinkly Spreaders in static microcosms and drip-fed columns selects for subtle differences in wrinkleality and fitness. *FEMS Microbiol Ecol* 2015;**91**:fiv057. https://doi.org/10.1093/femsec/fiv057.

Velicer GJ, Kroos L, Lenski RE. Developmental cheating in the social bacterium *Myxococcus xanthus*. *Nature* 2000;**404**:598–601. https://doi.org/10.1038/35007066.

Velicer GJ, Vos M. Sociobiology of the myxobacteria. *Annu Rev Microbiol* 2009;**63**:599–623. https://doi.org/10.1146/annurev.micro.091208.073158.

Wattam AR, Abraham D, Dalay O *et al.* PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res* 2014;**42**:D581–91. https://doi.org/10.1093/nar/gkt1099.

Wen A, Havens KL, Bloch SE *et al.* Enabling biological nitrogen fixation for cereal crops in fertilized fields. *ACS Synthetic Biology* 2021;**10**:3264–77. https://doi.org/10.1021/acssynbio.1c00049.

West SA, Buckling A. Cooperation, virulence and siderophore production in bacterial parasites. *Proc Biolog Sci* 2003;**270**:37–44. https://doi.org/10.1098/rspb.2002.2209.

White AP, Surette MG. Comparative genetics of the rdar morphotype in Salmonella. *J Bacteriol* 2006;**188**:8395–406. https://doi.org/10.1128/JB.00798-06.

Wilksch JJ, Yang J, Clements A *et al.* MrkH, a novel c-di-GMP-dependent transcriptional activator, controls *Klebsiella pneumoniae* biofilm formation by regulating type 3 fimbriae expression. *PLoS Pathog* 2011;**7**:e1002204. https://doi.org/10.1371/journal.ppat.1002204.

Wu DC, Zamorano-Sánchez D, Pagliai FA *et al.* Reciprocal c-di-GMP signaling: incomplete flagellum biogenesis triggers c-di-GMP signaling pathways that promote biofilm formation. *PLos Genet* 2020;**16**:e1008703. https://doi.org/10.1371/journal.pgen.1008703.

Xu A, Wozniak DJ, Zhou J *et al.* Toward a unified nomenclature for strains with hyper-biofilm phenotypes. *Trends Microbiol* 2022;**30**:1019–21. https://doi.org/10.1016/j.tim.2022.07.007.

Received 11 July 2023; **revised** 5 September 2023; **accepted** 9 September 2023