

Single-cell RNA sequencing reveals plasmid constrains bacterial population heterogeneity and identifies a non-conjugating subpopulation

Received: 13 March 2024

Accepted: 18 June 2024

Published online: 12 July 2024

 Check for updates

Valentine Cyriaque^{1,2}✉, Rodrigo Ibarra-Chávez¹, Anna Kuchina^{3,4,5}, Georg Seelig^{5,6}, Joseph Nesme¹ & Jonas Stenlørkke Madsen¹✉

Transcriptional heterogeneity in isogenic bacterial populations can play various roles in bacterial evolution, but its detection remains technically challenging. Here, we use microbial split-pool ligation transcriptomics to study the relationship between bacterial subpopulation formation and plasmid-host interactions at the single-cell level. We find that single-cell transcript abundances are influenced by bacterial growth state and plasmid carriage. Moreover, plasmid carriage constrains the formation of bacterial subpopulations. Plasmid genes, including those with core functions such as replication and maintenance, exhibit transcriptional heterogeneity associated with cell activity. Notably, we identify a cell subpopulation that does not transcribe conjugal plasmid transfer genes, which may help reduce plasmid burden on a subset of cells. Our study advances the understanding of plasmid-mediated subpopulation dynamics and provides insights into the plasmid-bacteria interplay.

To survive and reproduce, bacteria perform numerous tasks, including importing and processing nutrients and metabolites, cell division, defense/resistance, exogenous DNA acquisition, and many others. Yet, the number of tasks is limited by intracellular competition for relevant resources¹. Transcriptional heterogeneity at the single-cell level, resulting in subpopulations², expands the number of simultaneous processes that can be performed at the population level. This can, for example, facilitate division of labor where complementary metabolic processes are carried out by different subpopulations³. Heterogeneity also enables bet-hedging strategies where mal-adapted phenotypes may become an asset in a changing environment⁴. The formation of subpopulations in isogenic populations² can thus increase the fitness of bacteria while preserving the genotype. Subpopulations can emerge through a number of processes and phenomena such as cell cycle progression, responsive switching in spatial gradients, epigenetic

determination⁵, compartmentalization⁶ and stochasticity during gene expression¹. This results in transcriptionally and phenotypically distinct subpopulations⁷ fluctuating over space and time, and has been shown to be important during utilization of different metabolic substrates¹ and survival of sudden physiological stress (e.g., via detoxification³, competence⁸ or persistence⁹).

Plasmids are extrachromosomal semi-autonomous mobile genetic elements (MGEs) that interact with their host's genome to varying degrees. Plasmids use host factors¹⁰ such as nucleotides, tRNAs, amino acids, ribosomes¹¹ and thus take part in the intracellular competition for resources². More crude interactions also transpire, facilitated via plasmid-encoded regulatory elements^{9,12}, toxin-antitoxin systems¹³ or other defense and anti-defense systems¹⁴. The association of plasmids with a host can change its transcriptomic profile in a plasmid-host specific manner at the population level^{15–17}. Plasmids can,

¹Section of Microbiology, University of Copenhagen, Copenhagen, Denmark. ²Proteomics and Microbiology Laboratory, Research Institute for Biosciences, UMONS, Mons, Belgium. ³Institute for Systems Biology, Seattle, WA, USA. ⁴Molecular Engineering and Sciences Institute, University of Washington, Seattle, WA, USA. ⁵Department of Electrical and Computer Engineering, University of Washington, Seattle, WA, USA. ⁶Paul G. Allen School for Computer Science & Engineering, University of Washington, Seattle, WA, USA. ✉e-mail: valentinecyriaque@gmail.com; jmadsen@bio.ku.dk

for example, induce the use of alternative carbon sources or alter the metabolism of carbohydrates¹⁶, amino acids and nucleotides¹¹. Additionally, host respiratory activities¹⁶ and regulatory pathways^{16,17} can be directly influenced by plasmids. Over evolutionary time, compensatory mutations may reduce such genetic conflicts^{11,18} between plasmids and their hosts. However, while the above illustrates that plasmids interact with host chromosomes, very little is currently known about plasmid-host interactions at the single-cell level and how these interactions may affect the subpopulation dynamics of their host.

Recently, several techniques successfully enabled single-cell transcriptomics in bacteria either by transcriptome imaging (Parseq¹⁹), physically separating the cells with fluorescence-activated cell sorting (MATQ-seq^{20,21}) or using random molecular-barcode ligation to concomitantly tag mRNAs stemming from the same cell (smRandom-seq²², ProBac-seq²³, PETRI-seq²⁴, BacDrop²⁵, M3-seq²⁶, microSPLIT⁷). These latter approaches have made high-throughput single-cell RNA sequencing possible for a large number of prokaryotic cells in parallel.

Here, we successfully applied the microSPLIT (microbial split-pool ligation transcriptomics)⁷ approach to assess how a conjugative broad-host range plasmid influenced the subpopulation dynamics of *Pseudomonas putida* KT2440. This bacterium is widely used as a model in studies of soil bacteria and is important for biotechnology²⁷. *P. putida* KT2440 has also been considered for bioremediation (e.g.,²⁸) and repeatedly used in studies of plasmid transfer in natural (e.g.,^{29,30}) and synthetic communities^{31,32}. We sought to investigate transcriptional interactions between a conjugative broad-host-range plasmid and the host genome at the single-cell level hypothesizing that; (i) *P. putida* would form distinct transcriptomic subpopulations despite growing in a homogeneous environment; (ii) single-cell transcriptomics would be successful at segregating plasmid free and carrier cells; (iii) the broad-host-range plasmid would interact distinctly with the host, in a subpopulation dependent manner, and; (iv) the transcription level of plasmid-encoded genes would follow the growth status of its host to reduce its burden. In this work, bacteria grew with constant shaking in large bottles, making it unlikely that microenvironmental differences arose. Comparing bulk and single-cell transcriptomes, we observe clear differences at the individual-cell level leading to the formation of subpopulations. We find that the *P. putida* population is heterogeneous and

that the presence of a broad-host-range plasmid, which inflicts a population-wide burden, changes the subpopulation dynamics. The experimental design made it unlikely that transcriptional subpopulations identified were generated by arising mutations, as the cells only grow for a few generations and the mutation rate of *P. putida* is relatively low. Furthermore, our data demonstrate a differential transcript abundance of plasmid-encoded genes amongst bacterial subpopulations and that the transcription of operons, essential for conjugation, are absent in a subset of cells. These observations support that transcriptional heterogeneity has a determining role in the evolutionary and ecological success of plasmids, as seems to be the case for many bacteria.

Results

Plasmid affects growth but scantily chromosomal transcription

First, we sought to compare the growth of *P. putida* cultures in absence and presence of the broad-host range plasmid pJKJ5 to evaluate any cost or benefit in the overall population. Notably, we detected a decrease in growth rate when carrying the plasmid at the start of the exponential phase (AUC 23.5 ± 0.3 vs. 24.9 ± 0.4 , $p = 1.8 \times 10^{-9}$; $r = 0.76 \pm 0.03$ vs. 1.14 ± 0.08 , $p = 5.12 \times 10^{-10}$) (Fig. 1A, Table S1). However, towards the end of the growth phase, no difference was observed. When growing plasmid-free and plasmid-carrier cells in coculture, in a competition experiment, the plasmid showed no significant fitness cost or advantage ($w = 1.016 \pm 0.009$) after 16 hours of cocultivation (Table S2). First, unsupervised cluster analysis of the population-level (bulk) gene transcripts (excluding tRNA and rRNA) separated samples according to growth state (OD0.5 vs. OD1.5), and the presence/absence of the plasmid (Fig. S1). This separation, however, was mainly attributed to plasmid-encoded gene transcripts since, after excluding these, the separation between plasmid-free and carrier populations was no longer supported statistically (Fig. 1B), despite differential transcript abundance of a few chromosome-encoded genes.

microSPLIT generates quality scRNA-seq data

Few studies have currently utilized high-throughput single-cell transcriptomics, like microSPLIT. Therefore, the acquired data were assessed by comparison to (i) population-level bulk transcriptomes (RNA-seq) and (ii) a blind microSPLIT replicate (referred to as the

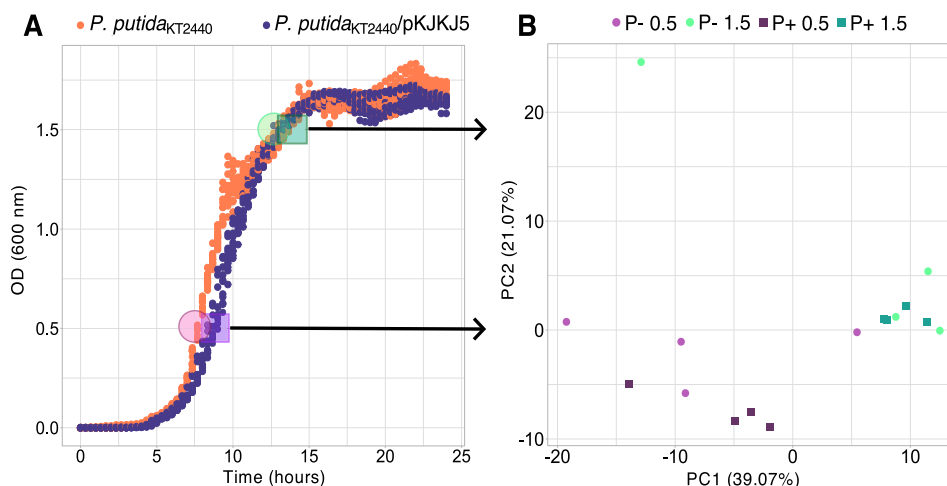


Fig. 1 | Small growth effect of plasmid carriage with little influence on chromosomal gene transcription at the population level. A Growth curve (OD = 600 nm) showing small differences in growth rates of *Pseudomonas putida* KT2440 with (blue) and without (orange) pJKJ5 (Kruskal–Wallis $\chi^2 = 16.516$; $p = 4.823 \times 10^{-5}$) and area under the curve ($t = 10.045$; $p = 1.792 \times 10^{-9}$; $n = 12$ independent experiments, see Table S1). Samples were taken at OD0.5 and OD1.5 for bulk transcriptomics or microSPLIT sample preparation (indicated by transparent colored circles following color code of B). **B** PCA multivariate analysis on

transcripts from chromosome-encoded genes at the population level (bulk; $n = 4$ independent experiments) according to the presence (P+) or absence (P-) of the plasmid and according to ODs (OD0.5 vs. OD1.5). The presence of the plasmid did not significantly discriminate between cell transcriptomes, as confirmed using 10,000 permutations in a two-sided PERmutational Multivariate ANalysis Of Variance (PERMANOVA; $r^2_{\text{plasmid}} = 0.094$; $p_{\text{plasmid}} = 0.0922$; $r^2_{\text{OD}} = 0.253$; $p_{\text{OD}} = 0.001$; $r^2_{\text{plasmid:OD}} = 0.035$; $p_{\text{plasmid:OD}} = 0.7$). Source data and detailed statistical results are displayed in the Source Data file.

microSPLIT control) (Fig. S2). In contrast to the main microSPLIT experiment, the blind microSPLIT control was generated by mixing samples (plasmid presence/absence and growth state) before generating the microSPLIT libraries. The sum of mRNA transcripts per gene in the microSPLIT samples correlated well with those both from population-level (bulk) transcriptomes obtained from total mRNA extractions ($r = 0.793$) and the microSPLIT control ($r = 0.944$), showing that the single-cell data reflected population-level transcription. Importantly, this also indicates that microSPLIT did not introduce a systematic bias for specific transcripts. A total of 1599 single-cell transcriptomes (>85 transcripts/cell, Fig. S3) were recovered, characterized by transcripts from 4432 different genes (over 5698 mRNAs annotated in the genome) over all cells, with an average of 424.36 mRNA transcripts per cell. When performing microSPLIT, a trade-off between cell number (as defined during library preparation) and cell transcriptome coverage was seen: the microSPLIT control generated more single-cell transcriptomes (3165 cells) characterized by transcripts from 4248 genes across all cells with a lower average number of transcripts per cell (181.6 mRNA transcripts/cell) (Fig. S4). In bulk transcriptomes, 5632 different transcripts were obtained.

Among plasmid carrier cells, 91.4% displayed at least 1 plasmid-derived transcript at OD0.5 (average plasmid-derived transcript ratio = 1.24%) and 94.8% at OD1.5 (average plasmid-derived transcript ratio = 1.73%), suggesting that the plasmid is stably maintained in *P. putida*, as shown in previous studies^{33,34}. The few percent of plasmid-encoded transcript free cells among plasmid carriers could result from a lower sequencing coverage per cell, segregational plasmid loss, or both.

For the following analyses, we filtered out transcripts that were present in less than 10% of cells for all growth conditions (P+0.5, P+1.5, P-0.5, P-1.5), resulting in 737 genes across 1486 cells (Table S3, S4, Fig. S5). Single-cell transcriptomics showed the formation of distinct subpopulations constrained by plasmid carriage. Detailed transcript abundance per growth condition and by cluster can be found in Table S5. Unsupervised cluster analyses based on whole single-cell transcriptomes including both plasmid- and chromosome-encoded transcripts (reduced to the 10 top significant principal components, Fig. S6) generated 11 distinct clusters labeled 'W', projected on a uniform manifold approximation and projection (8 dimensions, UMAP) plot (Fig. 2A). Clusters segregated according to growth state and plasmid carriage (Fig. 2A, B). Single cell transcriptomes from samples grown to OD0.5 were mainly distributed in clusters W1-5, while cells from OD1.5 were distributed in clusters W6-11. Clusters W3-5 and W9-11 were dominated by plasmid-free cells, while plasmid carrier cells mainly distributed in clusters W1-2 and W6-8. (Fig. 2A-C). Chromosome-only single-cell transcriptomes formed 9 clusters (Ch1-9 clusters; Fig. S7, Table S5) showing transcriptional subpopulations of cells generally similar to W clusters. This overlay (Fig. S7C) suggests that the segregation between the transcriptome of plasmid-free and -carrier cells is not only due to plasmid-encoded transcripts, but also to changes in the transcription profile of the host chromosome (Fig. S8).

Genes differentially characterizing clusters W1-5 vs. W9-11 (i.e., OD0.5 vs. OD1.5; Fig. 3A, B) as identified with the Seurat function FindMarkers, were housekeeping genes including ribosomal proteins (*rps*, *rpl*, *rpm*), elongation factors (*fusA*, *tufB*), global regulators (*rpoC*), and malate:quinone oxidoreductase (*mgo-II*), whose transcript number was higher at OD0.5. Induced transcription of these genes is typical for early exponential growth³⁵ and *Mgo-II*, for example, is regulated by carbon sources and oxygen availability, which changed over time due to consumption³⁶. At the end of the exponential phase (OD1.5), a high number of transcripts related to respiration in decreasing oxygen conditions characterized the samples: Differential transcribed genes included *ccoNOPQ-I*, the cytochrome b component of the ubiquinol-cytochrome c reductase complex (*petB*), amino acid metabolisms (*arcABCD-1*, *gdhB*), and flagellin (*fliC*). These observations were confirmed with population-level (bulk) transcriptomics comparing OD0.5

and OD1.5. A few genes discriminated between plasmid-free and carrier cells (Fig. 3C, D), such as those involved in rRNA maturation, protein translation and translocation (*rne*, *InfC*, *secD*, *tig*) or the 3-hydroxydecanoyl-dehydratase (*fabA*) whose increased transcript number in plasmid carrier at OD0.5, may have consequence for the membrane fluidity. Indeed, at the outset of growth, flow cytometry analyses revealed a small change in size and cell texture (Fig. S9). These changes may facilitate conjugation and plasmid maintenance.

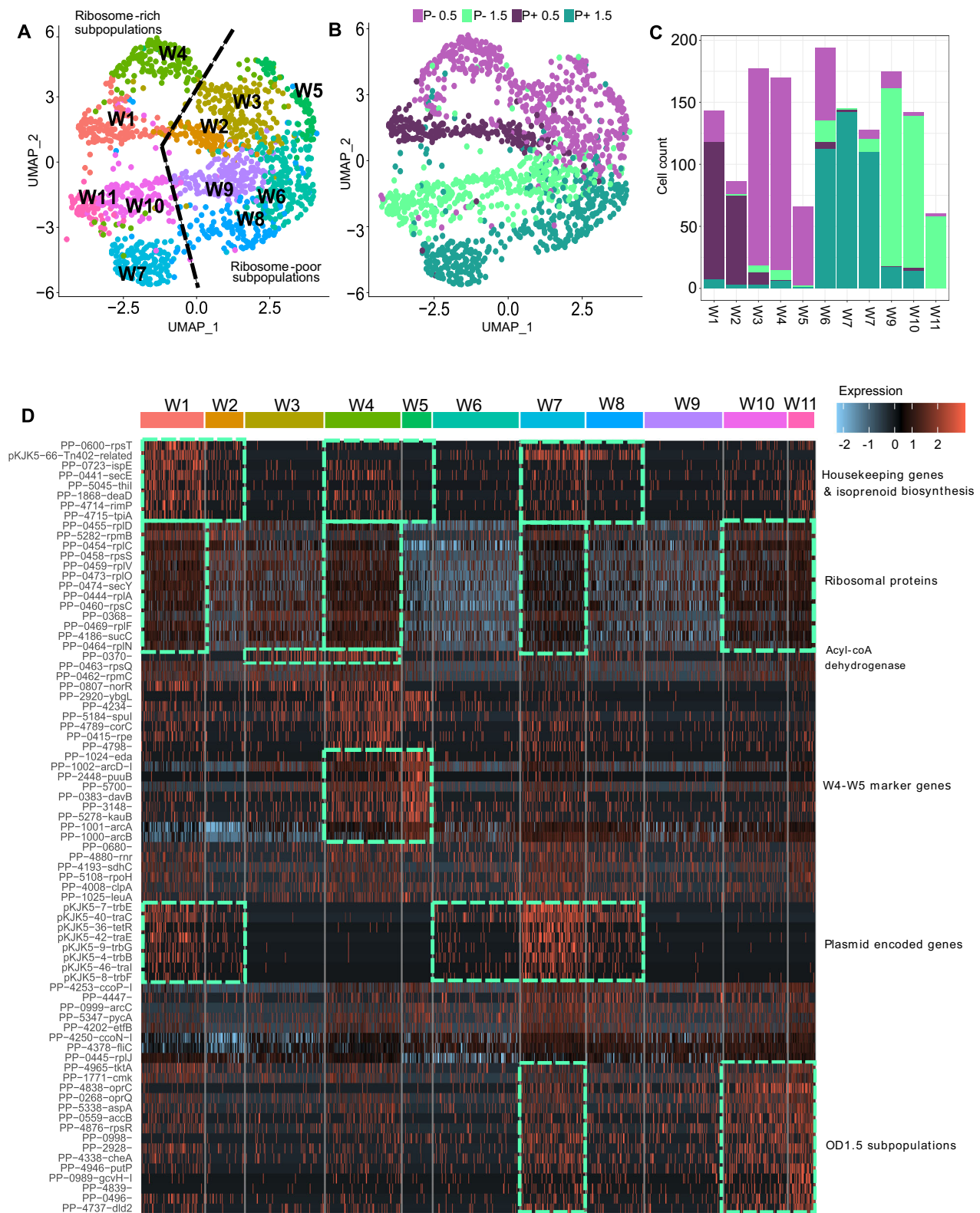
The few differentially transcribed chromosome-encoded genes identified at the population level (bulk transcriptomics), did not discriminate between plasmid-free and carrier cells. Looking at single-cell transcriptomes only, clusters W1, W4, W7, W10 and W11 ('Ribosome-rich subpopulations') were distinct from clusters W2, W3, W5, W6, W8, and W9 ('Ribosome-poor subpopulations') (Fig. 2A, D). Interestingly, the underlier of this separation did not seem to be related to OD or plasmid presence/absence. Instead, subpopulations W1, W4, W7, W10, and W11 were characterized by increased levels of transcripts encoding ribosomal proteins (*rpl*, *rpm* and *rps*). This suggests that subpopulations with distinct translation rates or cell cycle stages co-occurred regardless of the growth state. Furthermore, an increased rate of ribosomal gene transcription co-occurred with an increased abundance of transcripts of house-keeping genes, isoprenoid biosynthesis and, in plasmid carrier cells, to plasmid encoded genes (Fig. 2D).

At the beginning of the exponential phase (OD0.5), plasmid-free cells were evenly distributed among subpopulations W3, W4, and W5 (Fig. 2C) with the latter almost exclusively populated with plasmid-free cells. However, W5 cells were characterized by transcripts that also defined cells at OD1.5. Indeed, when mapping W5 transcriptomes with the transcriptomes from the microSPLIT control (Fig. S10), we see that those cells cluster with cells with OD1.5 transcription profiles (including clusters W6-9). The W5 cluster exhibits increased transcript abundance of genes involved in (i) glycolysis (*eda*), (ii) the arginine deiminase pathway and L-arginine metabolism (*arcABCD*, *kaub*), lysine oxidative decarboxylation (*davB*), (iii) glutamine synthase (PP-3148) and (iv) the oxidation of L-glutamylputrescine (*puuB*) (Fig. 2). This subpopulation of cells was also characterized by having many transcripts involved in motility (*fliC*), and respiration (*ccoP*). The chromosome single-cell transcriptome data suggest that cells in the W5 subpopulation shifted toward transcriptional patterns to an alternative source of energy, and nitrogen through arginine and glutamine metabolisms.

Notably, at the end of the exponential phase (OD1.5), subpopulations of cells displaying an increased number of ribosomal gene transcripts (W7, W10, and W11; Fig. 2D), also displayed more transcripts of efflux pumps (*oprCQ*), for nucleic acid and amino acid synthesis, degradation and import (*tktA*, *cmk*, *putP*, *gchV-I*, PP-0496, *aspA*), regardless of the plasmid (Fig. 2D, "OD1.5 subpopulation").

Plasmid-encoded genes are differentially transcribed among cells

The subpopulation dynamics at OD0.5 seemed to be driven by the presence of the plasmid, which may have led to the decreased growth rate of the plasmid carrier cells, perceived at the outset of growth (Fig. 1A). The burden of the plasmid may also have resulted in the observed decrease of the cluster coefficient in plasmid carrier cells (degree to which genes tend to cluster together; Table S6), in network graphs calculated from spearman correlations between genes of each single-cell at OD0.5 (>85 transcripts/cell; Fig. S11). Reduced co-occurrence correlation between chromosomal genes of plasmid carrier cells suggests that the plasmid interferes with chromosome transcriptional regulation. Furthermore, when calculating the Spearman correlation between plasmid and chromosomal genes (Fig. 4A), we see that plasmid encoded genes, especially *tra* genes, were transcribed concomitantly ($n > 30$, $p < 0.05$) with *mgo-II* and ribosomal protein transcripts (*rpsI*, *rpsP* and *rpmI*, *rpmJ* genes). Additionally, *tra* gene



transcription exhibited significant positive correlations ($r_s > 0.6$, $n > 30$, $p < 0.05$) with transcripts involved in LPS biosynthesis and membrane protein translocase and chaperones (*lptD*, *rfbAD*, *tig*, PP-2304), DNA synthesis (*nrdB*, *purA*), peptide transport, rRNA maturation, protein synthesis and folding (*dppA-II*, *rnr*, *rbbA*, *fkfB*), respiration, energy production, oxidoreductases, oxidative stress (*cyoAC*, *fpr1*, *nuoF*, *atpB*, PP-2010, *sodB*) and carbon metabolism and coordination (*gpm*, *pycAB*, PP-3443, *spoT*), especially after the exponential phase (OD1.5). Intriguingly,

these chromosomal gene transcripts also positively correlated with pKJK5-encoded Tn402-related gene transcription (Fig. 4A).

Transcription of the *traGFEDC* operon specifically, which encodes genes involved in DNA replication and transfer, critical for conjugative plasmid transfer, was further investigated (Fig. 4). The P_{traG} promoter of the *traGFEDC* operon, was fused with the green fluorescent protein sfGFP (P_{traG} -sfGFP) on the pPROBE-NT vector³⁷ to enable quantification of single-cell expression levels by flow cytometry. Both microSPliT

Fig. 2 | Whole single-cell transcriptomes generally cluster in accordance with growth state and the presence/absence of the plasmid. Subpopulation clustering identified by single-cell transcriptomics of *P. putida* carrying a plasmid (P+) or not (P-) at early (OD0.5) and late exponential growth (OD1.5) ($n = 1$ independent experiment; 1486 cells). **A** UMAP obtained by microSPLIT scRNA sequencing identified 11 whole transcriptome clusters (W1-W11). Clustering was performed with an integrative approach combining Euclidean distance-based K-nearest neighbor, refining the edge weights with the Jaccard similarity and the Louvain algorithm. Color of the dot indicate to which cluster the cell belongs as indicated on the graph (W1-11). **B** Same UMAP graph where color of the dot indicate the growth state (OD0.5 and OD1.5) and the presence of the plasmid (P+/P-). **C** Cell

transcriptomes from OD0.5 were mainly distributed in clusters W1-5, while cells from OD1.5 were mainly distributed in clusters W6-W11. Data are presented as absolute cell count. **D** Heatmap of normalized and centered-scaled transcript number per cell as displayed through the expression color code (normalization consist in a \log_{10} transformation of transcript numbers * 10,000 divided by summed transcript number of the cell). Top colors represent the cell cluster (W1-11). The top 8 biomarker genes ($p < 0.05$) identified by a two-sided Wilcoxon signed-rank test are shown for each cluster (common biomarker to multiple groups where displayed only once). The p -value associated with each marker gene for each cluster (cluster of interest vs. all cells) can be found in Supplementary Dataset 1. Source data are provided as a Source Data file.

single cell transcriptomes and flow cytometry of the P_{traG} -*sfGFP* reporter showed, that *traGFEDC* transcription was increased at OD1.5, as more cells in the population had a positive transcription signal (GFP fluorescence and *traGFEDC* transcripts; Fig. 4B, C) and the level of transcription signal in individual cells was higher (fluorescence intensity and transcript number; Fig3D). Furthermore, an unsupervised cluster analysis was performed including only plasmid transcripts from plasmid-carrier cells (>85 transcripts/cell). Transcript numbers were normalized by dividing transcript counts by the summed numbers of plasmid-encoded transcripts and performing $\log+1$ transformation (Fig. 5). Despite a low significance of the principal component, due to low transcript numbers, 6 clusters (P1-6) of plasmid transcriptomes were identified (Fig. 5A), grouping cells independently of growth state. Interestingly, the Tra2 region (*summed trb* genes encoding the mating pair formation including pilus assembly) was homogeneously expressed among plasmid carriers (Fig. 5C). Yet, the transcriptional regulator TrbA that targets P_{trbB} and the Tra1 region, was present in a small number of cells that also had fewer transcripts of other *trb* genes (Fig. 5D). Also, the P4 plasmid cluster which represented about 12% of plasmid carriers, was distinguished by a lack of transcripts from the *traGFEDC* operon (Fig. 5E, F). The occurrence of a subpopulation of cells not expressing the *traGFEDC* was confirmed by the GFP-based reporter plasmid where P_{traG} was inactive in about 13% of the cells (cells with no detectable GFP signal). Nonetheless, the P4 subpopulation did not appear to be inactive or consist of dead cells, as the numbers of *rps* and *rpl* ribosomal protein transcripts were equivalent to those of other plasmid-carrier cells (Fig. S13).

Discussion

Here, we investigated transcript abundance at the single-cell level of *P. putida* with and without the conjugative broad-host-range IncP-1 plasmid pJK5. With microSPLIT, we observed that *P. putida*, despite growing in a continuously homogenized environment, displayed transcriptional heterogeneity. This is in agreement with the few other studies using similar approaches^{7,24,25} on bacteria belonging to different orders. Cells clustered into transcriptional subpopulations, displaying differential expression of ribosomal proteins and alternative amino acid and catabolic, energy-providing pathways. Generally subpopulations may arise due to microenvironmental differences, epigenetic determination, phase variation² and mutations³⁸. The formation of subpopulations has the potential to improve the population-wide fitness of the genotype by facilitating bet-hedging or division of labor strategies. In this study, the subpopulation W5 of plasmid-free cells observed at OD0.5 appeared to undergo a diauxic shift and/or co-utilization of energy sources³⁹, which have been shown to occur through stochastic events leading to bi-phasic shifts within isogenic populations⁴⁰. The use of an alternative energy source by plasmid-free cells may have led to the differences in overall growth rates observed between plasmid-free and carrier cells since, during early exponential growth, no plasmid carriers adopted the transcriptome profile of subpopulation W5. The stability of pJK5 within a bacterial population is ensured by conjugation³³, genes ensuring stable inheritance (partitioning system *parA* and *incC*; *klc* and *kle* genes³³), a putative toxin-antitoxin system⁴¹, and a replication system that enlarges the host

range of the plasmid, especially among proteobacteria⁴¹. However, its stability in a subpopulation of cells seems to also depend on the transcriptomic context. The aforementioned shift to an alternative energy-providing pathway, combined with the cost of plasmid carriage, may have induced a fitness loss, impeding *P. putida* from meeting the burden of plasmid maintenance⁴², by reducing the number of formed subpopulations. The W5 subpopulation was characterized by a higher number of transcripts involved in arginine deiminase pathway leading to the expulsion of ornithine. The plasmid may thus have deprived the population of plasmid carriers of an alternative ATP/ NH_3^+ source in addition to extracellular ornithine, which interestingly, was also shown to act as a public good for *Streptococcus* species^{43,44}.

When investigating correlations between plasmid- and chromosome-encoded transcript abundances among plasmid-encoded genes, we found that *tra* genes involved in DNA replication for transfer, were co-transcribed with housekeeping genes, which imply that the pJK5 plasmid decreases its burden by aligning with its host's activity. Likewise, we show that the abundance of transcripts associated with the Tn-402-related transposon encoded by pJK5 was also correlated to the same housekeeping genes. Similar observations were made in another single-cell transcriptome study²⁵ and we speculate that many MGEs may, at least in part, rely on, or synchronize with, the cell's transcriptional activity for the induction of transposition or horizontal mobilization, as observed with other transposable elements (e.g.,⁴⁵), which could be advantageous for the fitness of the host and the success of the MGE itself. Lastly, we show with both microSPLIT and the P_{traG} -*sfGFP* promoter activity reporter, that a defined subpopulation of cells did not transcribe the *traGFEDC* operon and thus did not engage in conjugation. Interestingly, this subpopulation did transcribe ribosomal proteins and did, therefore, not appear to be either inactive or dead. It may seem counterintuitive that the plasmid does not maximize its potential for horizontal transfer; however, decreasing the expression of the large conjugative machinery in a subpopulation of cells, and thus its burden on the host cell, may represent a bet-hedging strategy by which the plasmid promotes its association with the genotype, in a specific subset of cells.

This study is among the first to discern the role of bacterial subpopulations in plasmid biology. Using microSPLIT, we illustrate that transcriptional heterogeneity among cells can indeed be central to plasmid-host genome interactions. Plasmid carriage impacted the subpopulation dynamics of the host, and the transcript abundance of plasmid-encoded genes was highly heterogeneous. Central plasmid functions such as replication and maintenance seem to fluctuate, probably in sync with the cell cycle, while transcript abundance of genes critical for conjugation was biphasic, resulting in a smaller subpopulation that likely does not engage in conjugation. Studying subpopulation dynamics holds great promise for the understanding of bacterial ecology and evolution and for biotechnological applications. An important perspective of the findings presented here is that we need to further understand single-cell heterogeneity in plasmid biology at a broader scale and in natural environments where microbes experience a large diversity of environmental factors, and face a large variety of MGEs, like plasmids, ICEs, phages, PICIs and transposons,

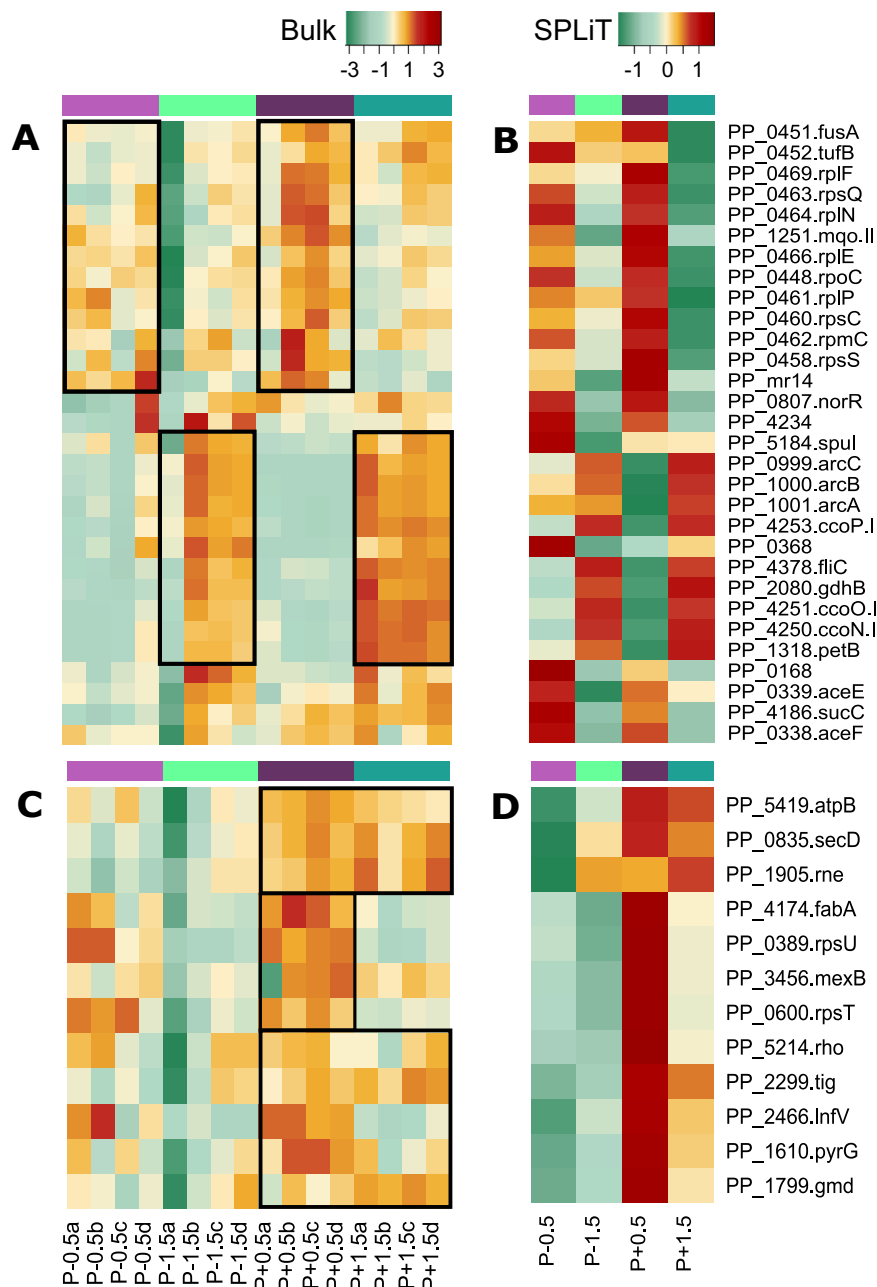


Fig. 3 | A few genes discriminate between samples according to OD and the presence of the plasmid. Heatmaps obtained from (A, C) the normalized population (bulk; $n = 4$ biological replicates) and (B, D) the summed normalized microSPLiT (scRNA seq; $n = 1$ independent experiment) normalized transcription rate of genes characterizing cells categorized by growth stage (OD0.5 or OD1.5) and plasmid presence (P-/P+) showing the 30 first marker genes (two-sided non-parametric Wilcoxon rank sum test; $p < 0.01$) obtained with Seurat separating

(A, B) clusters “W1”, “W2”, “W3”, “W4”, “W5” vs. “W6”, “W7”, “W8”, “W9”, “W10”, “W11” mainly composed of OD0.5 and OD 1.5 cells, respectively; and (C, D) separating clusters “W1”, “W2” vs. “W6”, “W7”, “W8”, and “W6”, “W7”, “W8” vs. “W9”, “W10”, “W11” mainly composed of plasmid carrier- or free- cells at each OD, respectively. Normalization was operated by dividing genes counts by the total count of the sample *10,000 as displayed on the color key. Source data are provided as a Source Data file.

enhancing genetic variability among populations (e.g.,^{46,47}). Therefore, and in the context of the ongoing antibiotic resistance crisis, a deeper understanding of the interactions between plasmids and other MGEs with their host will be critical.

Methods

Bacterial strains, cultures, growth curve, and competition experiment

Pseudomonas putida KT2440 (ATCCTM 47054) growth was performed on Lysogeny-Broth medium (LB) at 30 °C and 250 RPM. *Pseudomonas putida* KT2440/pKJK5 was obtained by electroporating 1 µg of pKJK5

plasmid (private collection) into 50 µL of competent *P. putida* KT2440. Competency was achieved at an optical density (600 nm) of 0.8. A culture of 50 mL was washed twice with 50 mL of 10% ice-cold glycerol and finally resuspended in 800 µL of 10% ice-cold glycerol. Cells were aliquoted (50 µL) and electroporated in a 2 mm cuvette at a voltage of 2400 V, 1 pulse. For the following experiments, *P. putida* KT2440 and *P. putida* KT2440/pKJK5 were pre-cultured separately overnight in 5 mL of LB medium. *Pseudomonas putida* KT2440/pKJK5 was cultivated with tetracycline (50 µg/mL). For the growth curves (3 biological replicates, 4 technical replicates), pre-cultures were washed in LB medium and diluted 10⁴ times before distributing 200 µL in a 96-well plate

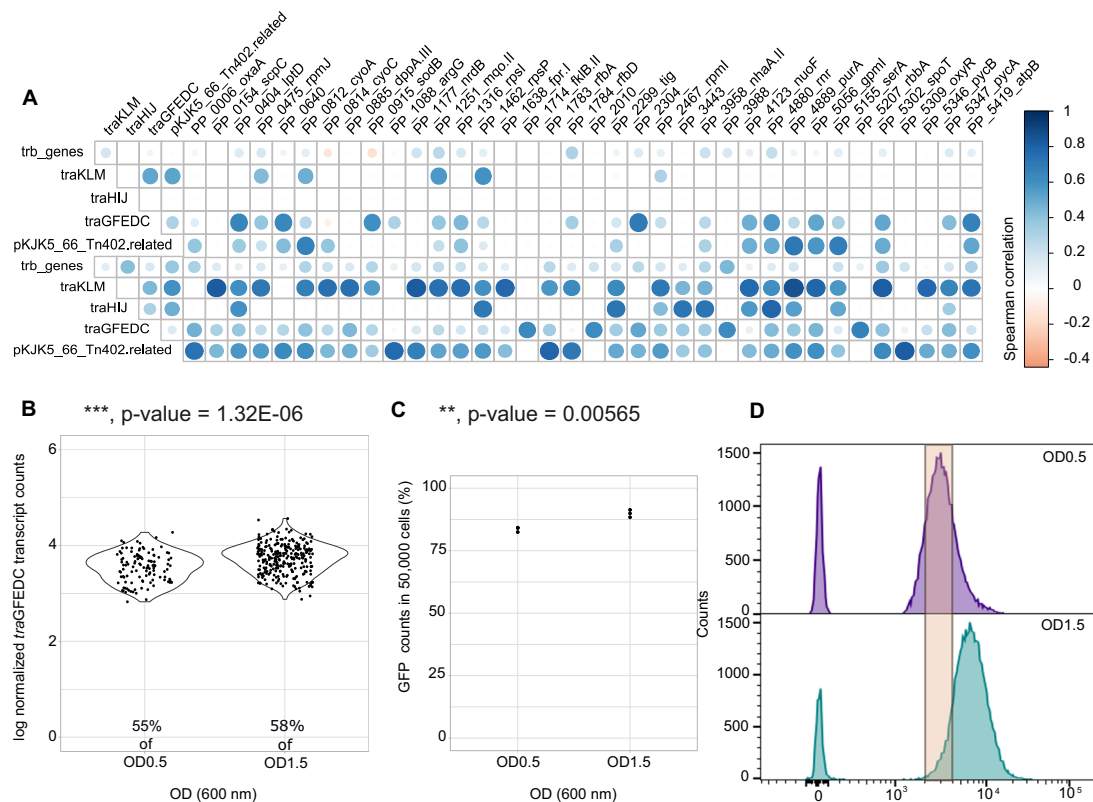


Fig. 4 | Subpopulation transcriptional heterogeneity of the *traGFEDC* operon, which is essential for conjugation, validated by flow cytometry of the promoter-*sfGFP* fusion reporter. A Spearman correlation matrix between plasmid genes with several chromosome-encoded genes. Significance was determined with a two-sided *t*-test by the *rcorr*-5.1-2.R package ($n > 30$ cells, $p < 0.05$). The size of dot is proportional to the correlation **(B)** Normalized summed transcript counts of *traGFEDC* genes from individual cells on a violin plot according to OD ($n = 1$

independent experiment; $n = 384$ *traGFEDC* +/P+ cells). **C** Percentage (mean \pm SD; $n = 3$ independent experiments; $n = 50,000$ cells) of *P*_{traG} promoter reporting system activating cell fluorescence (FITC-A) as determined by GFP fluorescent cells counted using flow cytometry **(D)** Flow cytometry histogram for *PtraG-sfGFP* reporter signal from samples taken at OD0.5 and OD1.5 representing cell count in FITC-A expression level (experiment was repeated 3 times with similar results). Source data and detailed statistical results are displayed in Source Data file.

incubated at 30 °C and shaken every 20 min. Optical density (600 nm) was recorded after shaking using a Synergy HI microplate reader (BioTek, Winooski, VT). For the competition experiment (4 biological replicates), overnight cultures were washed in LB medium and mixed (*Pseudomonas putida* KT2440 + *Pseudomonas putida* KT2440/pKJK5) equally. The mixes were diluted 10^4 times, and 800 μ L was distributed in a 12-well plate. Total cell number and plasmid carriers were counted before and after competition using a 10^5 -fold dilution plated on LB plates and LB and tetracycline 20 μ g/mL, respectively. The fitness of the plasmid was calculated as $w = \ln(P_{\text{carrier}_f}/P_{\text{carrier}_i})/\ln(P_{\text{free}_f}/P_{\text{free}_i})$. For the transcriptomics analyses, 100 mL cultures were incubated in flasks from a 200-fold dilution and sampled at 600 nm optical densities of 0.5 and 1.5.

GFP-based reporter plasmid and flow cytometry analysis

The promoter *PtraG* was cloned using HiFi Gibson Assembly into a pPROBE-NT backbone with primers traGp_Fw (5'-gttagttagggaa-taagccgagttttaaggagcctcgcg-3') and traGp_Rv (5'-aggctgactc-taggatcgccaggaagaggcctaag-3') to amplify the *PtraG* region from pKJK5 with overhangs to pPROBE-NT (Addgene Plasmid #37818) and primers MFHO38 (5'-gatcctctagagtcgacctgc-3') and MFHO37 (5'-tcggcttattccctaactaactaaag-3') for pPROBE-NT. The GFP reporter plasmids were transformed into *P. putida* KT2440 (negative control) and *P. putida* KT2440/pKJK5 following the electroporation procedure. Overnight cultures were diluted 1/50 and normalized to an OD of 0.1. These cultures ($n = 3$) were then grown at 30 °C and 250 RPM. For cell size and granulometry, samples were taken at OD0.5 and 1.5 and examined by flow cytometry (FACS Aria III, Becton Dickinson Biosciences, San Jose,

CA, USA) using the 488 nm laser and the FSC and SSC detection channels, respectively ($n = 3$ independent experiments). To measure GFP intensity and abundance, samples were taken at OD0.5 and OD1.5, washed twice with PBS and examined by flow cytometry using a 488 nm excitation laser and the FITC (530/30 nm bandpass filter) detector channel. The closing gate was set at 50,000 counts using the FSC and SSC detection channels. A culture of *P. putida* KT2440/pKJK5 containing the pPROBE-NT plasmid was used as a control to set up the gate for the GFP signal on the FITC detection channel. The data in Fig. 4 were analyzed using FlowJo software (Tree Star Inc., USA).

Transcriptomic data acquisition

After an overnight culture of 16 h in 5 mL of selective LB (250 RPM), *Pseudomonas putida* KT2440 and *Pseudomonas putida* KT2440/pKJK5 were washed and diluted 500 times in 100 mL of LB. Flasks were incubated at 30 °C (250 RPM), and cultures were sampled at turbidity (600 nm) of 0.5 and 1.5.

Population-level (bulk) RNA-seq. Samples ($n = 4$ independent experiments) were centrifuged (3 min, 7000 $\times g$), RNA content was directly extracted using the Quick-RNA Fungal/Bacterial Miniprep Kit (Zymo Research, CA, USA) following the manufacturer's instructions, and libraries were prepared using the Zymo-Seq RiboFree Total RNA Library Kit (Zymo Research, CA, USA). Fragment integrity and size were assessed using a fragment analyzer (Agilent, CA, USA) and quantified with a Qubit dsDNA HS assay kit (Thermo Fisher Scientific, Waltham, MA, USA) before sequencing on a Novaseq 6000 (Illumina, CA, USA) at 2 \times 150 bp performed by Novogene Co., Ltd. (Cambridge,

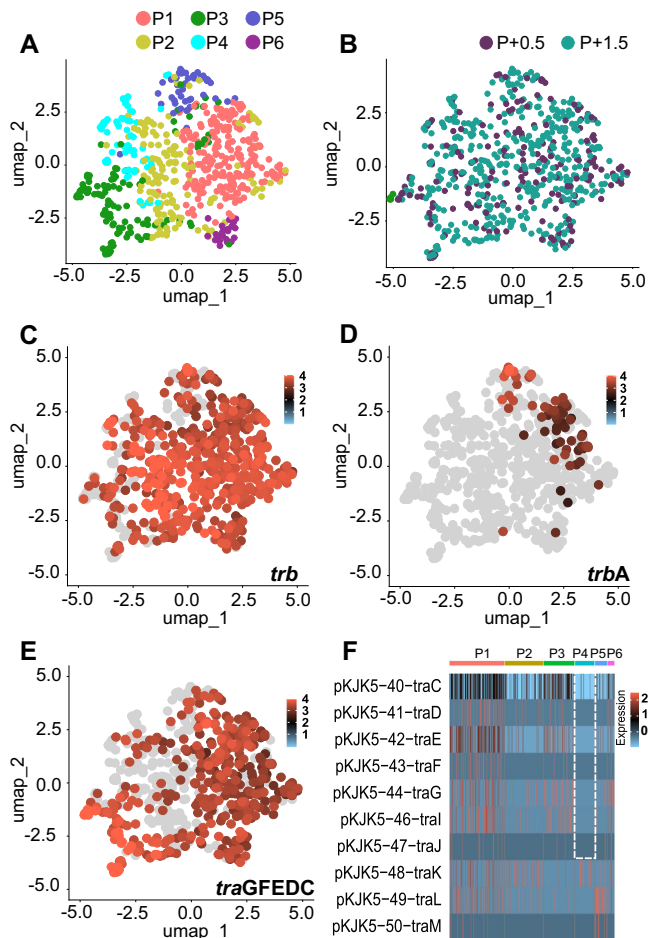


Fig. 5 | Plasmid single-cell transcriptome clustering suggests population heterogeneity in plasmid gene transcription, with differential expression of *tra* genes. Subpopulation clustering identified by single-cell transcriptomics of *P. putida* plasmid carriers at early (OD0.5) and late exponential growth (OD1.5) ($n = 621$ P+ cells, Table S3). **A** UMAP obtained by microSPLiT scRNA sequencing identified 6 plasmid transcriptome clusters (P1–P6) when only plasmid transcripts were analyzed with 6 dimensions (see Fig. S12). Clustering was performed with an integrative approach combining Euclidean distance-based K-nearest neighbor, refining the edge weights with the Jaccard similarity and the Louvain algorithm. Same UMAP, where dots were colored according to **(B)** the OD of the cell or **(C–E)** to normalized transcript number of **(C)** all *trb* genes, **(D)** *trbA* genes, **(E)** *traGFEDC* genes. Normalization consist in a \log_{10} transformation of transcript numbers $+1 * 10,000$ divided by summed plasmid transcript number of the cell. **F** Heatmap of the normalized centered-scaled number of *tra* transcripts per cell ordered according to clusters P1–6. Source data are provided as a Source Data file.

United Kingdom). Sequence files can be found at NCBI under Bioproject ID PRJNA1019643.

microSPLiT scRNA-seq. The experiment was repeated 2 times ($n = 2$ independent experiments) following the microSPLiT protocol⁷. Samples were centrifuged (3 min, 7000 $\times g$), resuspended in the same volume of cold paraformaldehyde 4% (pH 7.2) for fixation and kept at 4 °C overnight. Fixed samples were permeabilized by lysozyme, and a polyA tail was added to mRNA using *E. coli* Poly(A) Polymerase (New England Biolabs, Ipswich, MA). Cell subsamples were stained with SYBRTM Green I (Thermo Fisher Scientific, Waltham, MA, USA), and their concentration in each sample was measured using a FACS Aria III (Becton Dickinson Biosciences, San Jose, CA, USA). A total of 250,000 cells were distributed in 48 wells where reverse transcription occurred (Maxima H Minus Reverse Transcriptase, Thermo ScientificTM, Thermo Fisher Scientific, Villebon sur Yvette, France), each well containing

primers with a specific barcode. Cell distribution was carried out blindly during the first experiment, while P-0.5, P-1.5, P+0.5 and P+1.5 samples were knowingly distributed among the 48 wells. After reverse transcription, the cells were pooled together and physically separated (vortex and filtration steps at 10 and 1 μm , pluriSelect Life Science UG (haftungsbs.) & Co. KG, Leipzig, Germany) and randomly split into 96 wells where a well-specific oligo was added to the cDNA construct by ligation with a T4 DNA Ligase (New England Biolabs, Ipswich, MA). Cell pooling, random splitting and ligation were repeated for a third barcode addition so that, statistically, cell-hosted cDNAs carry a unique 3-barcode combination. Cells were pooled, divided into sublibraries of 10,000 (replicate 1) or 6000 cells (replicate 2) and lysed after measuring their concentration using a FACS Aria III (Becton Dickinson Biosciences, San Jose, CA, USA). The second strand synthesis was performed using template switching oligos, and cDNAs were amplified using a KAPA HiFi HotStart ReadyMix PCR kit (Roche Sequencing Solutions, Pleasanton, CA) supplemented with EvaGreen[®] Dye (Biotium, San Francisco, CA) for 14 cycles in total. Sublibraries were fragmented using Enzymatics Fragmentase (QIAGEN Beverly, Beverly, CA) and amplified. Sublibrary-specific adapters were ligated using Enzymatics ligase (QIAGEN Beverly, Beverly, CA), and sublibraries of cDNA were amplified using a KAPA HiFi HotStart ReadyMix PCR kit (Roche Sequencing Solutions, Pleasanton, CA).

Fragment integrity and size were assessed using a Fragment Analyzer (Agilent, CA, USA) using an NGS Fragment Kit (1–6000 bp) and quantified with a Qubit dsDNA HS assay kit (Thermo Fisher Scientific, Waltham, MA, USA) before sequencing on a Novaseq 6000 (Illumina, CA, USA) at 2×150 bp performed by Novogene Co., Ltd. (Cambridge, United Kingdom). A list of primers and associated barcodes can be found in Kuchina et al.⁷. Sequence files can be found at NCBI under Bioproject ID: PRJNA1019643. The number of cells and transcripts per cell were equivalent in the different generated sublibraries (Fig. S4B).

Computational method

The reference genome was generated using STAR 2.7.9a (<https://github.com/alexdobin/STAR>) combining the pKJK5 plasmid and *Pseudomonas putida* KT2440 genome using, respectively, GenBank record AM261282.1 and assembly GCA_000007565 from EnsemblBacteria. The reference genome was indexed for STAR using parameters (`--genomeSAindexNbases 10`) specific for small genomes using the formula $\min(14, \log_2(\text{Genome Length}/2 - 1))$. The created genome was used as a reference for both approaches (bulk & microSPLiT).

Population-level (bulk) RNA-seq. The sequenced reads were trimmed of the remaining adapter sequences and low quality based using bbduk (BBMap 38.90–Bushnell B. – sourceforge.net/projects/bbmap/), with right-side trimming using parameters $\{k=23 \text{ mink}=11 \text{ hdist}=1 \text{ tpe} \text{ tbo} \text{ trimq}=10\}$. Trimmed reads were mapped against the reference genome, and the per-cell gene counts were quantified using STAR 2.7.9a (<https://github.com/alexdobin/STAR>). Transcripts identified as rRNA, tRNA and plasmid encoded were removed from the data. For heatmaps, feature counts from the generated contingency table ($\text{gene} \times \text{sample}$) were normalized (divided by the total counts, multiplied by 10,000 + 1, and \log_{10} transformed) and centered-scaled with the R scale function.

microSPLiT scRNA-seq. Sequenced reads were demultiplexed (barcode list can be found in ref. 7) and mapped against the reference genome, and the per-cell gene expression was quantified using STARsolo (STAR 2.7.9a; <https://github.com/alexdobin/STAR/>). A matrix of unique molecular identifier (UMIs, unique gene transcript-associated barcode) counts for each cell (N-by-K matrix, with N cells and K genes) was generated by gathering cells from all sublibraries. UMIs identified as rRNA or tRNA were removed from the dataset, and cells with less than 85 UMIs were sorted out. This threshold was chosen

to be above the sharp drop of UMIs⁴⁸ as determined with the kneegraph (UMI count per barcode rank; Fig. S3). This left samples sizes of 3165 cells (Control E1 experiment) and 1486 cells (E2 experiment). We consider that this sample size was sufficient since cell identity could be retraced from the clustering⁴⁸. UMIs identified as plasmid genes were counted and added as metadata to characterize cells, as well as the sublibrary, the associated experiment and the sample type. For each subset of data (whole-genome encoded genes; chromosome encoded genes or plasmid encoded genes), a SeuratObject was created using the R package Seurat 5.0.1. (<https://cran.r-project.org/web/packages/Seurat/index.html>) with R 4.3.2. Genes presents in less than 10% of cells in all cell types (P-0.5; P-1.5; P+0.5; P+1.5) were sorted out for subsets of data made of whole-genome encoded genes and chromosome encoded genes in order to ensure a robust analysis⁴⁹. Feature counts per cell were divided by the total counts of the cell, normalized (divided by the total counts, multiplied by 10,000 + 1, and log10 transformed), scaled and centered using the scale R function (heatmaps) or linear model with the ScaleData function from Seurat R package for cell clustering. Variables were individually regressed against each feature, and the resulting residuals are then scaled and centered. Linear dimensional reduction was performed with a PCA, and clustering parameters were selected by combining a JackStraw resampling test and an elbow plot (Fig. S7). Cells were clustered with an integrative approach combining Euclidean distance-based K-nearest neighbor, refining the edge weights with the Jaccard similarity and the Louvain algorithm. A nonlinear dimensionality reduction was then applied with UMAP. Inter-gene Spearman correlation and associated *p* values were calculated with the *rcorr-5.1-2R* package (<https://search.r-project.org/CRAN/refmans/Hmisc/html/rcorr.html>).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sequence data generated in this study have been deposited in the National Center for Biotechnology Information's SRA repository under accession code [PRJNA1019643](https://www.ncbi.nlm.nih.gov/sra/PRJNA1019643). The reference genome data used in this study were generated combining the pKJK5 plasmid sequence available in the NCBI database under accession code [AM261282.1](https://www.ncbi.nlm.nih.gov/nuccore/AM261282.1) and *Pseudomonas putida* KT2440 genome available in the Ensembl Bacteria database under accession code [GCA_000007565](https://www.ensembl.org/Genome/View/Sequence?g=1000000000&h=1000000000&v=1000000000). The mapping, demultiplexing and quantification for bulk and single cell RNA-seq data generated in this study have been deposited in the Zenodo database under accession code [1135666](https://zenodo.org/record/1135666). Source data are provided with this paper.

Code availability

Statistical analyses were performed with R 4.3.2 (packages *vegan*-2.6-4, *mvabund*-4.2.1, *Seurat*-5.0.1, *patchwork*-1.2.0, *dplyr*-2.4.0, *scCustomize*-2.0.1, *ggplot2*-3.4.4, *rcorr* -5.1-2 and *corrplot*-0.92). The sequenced reads were trimmed of the remaining adapter sequences and low quality based using *bbduk* (BBMap 38-90-Bushnell B. - sourceforge.net/projects/bbmap/). Reads were mapped against the reference genome, and the per-cell gene counts were quantified using *STAR*-2.7.9a (bulk transcriptomics) or *STARsolo* from *STAR*-2.7.9a (single-cell transcriptomics) (<https://github.com/alexdobin/STAR>). Flow Cytometry analyses were obtained with *FlowJo* software v10 (Tree Star Inc., USA).

References

- Han, Y. & Zhang, F. Heterogeneity coordinates bacterial multi-gene expression in single cells. *PLoS Comput. Biol.* **16**, 1–17 (2020).
- van Vliet, S. et al. Spatially correlated gene expression in bacterial groups: the role of lineage history, spatial gradients, and cell-cell interactions. *Cell Syst.* **6**, 496–507.e6 (2018).
- Rosenthal, A. Z. et al. Metabolic interactions between dynamic bacterial subpopulations. *Elife* **7**, 1–18 (2018).
- Grimbergen, A. J., Siebring, J., Solopova, A. & Kuipers, O. P. Microbial bet-hedging: the power of being different. *Curr. Opin. Microbiol.* **25**, 67–72 (2015).
- Lim, HanN. & Van Oudenaarden, Alexander “A multistep epigenetic switch enables the stable inheritance of DNA methylation states.”. *Nat. Genet.* **39**, 269–275 (2007).
- Christen, M. et al. Asymmetrical distribution of the second messenger c-di-GMP upon bacterial cell division. *Science* **328**, 1295–1297 (2010).
- Kuchina, A. et al. Microbial single-cell RNA sequencing by split-pool barcoding. *Science* **371**, 1–9 (2021).
- Dubnau, D. & Blokesch, M. Mechanisms of DNA uptake by naturally competent bacteria. *Annu. Rev. Genet.* **53**, 217–237 (2019).
- Vial, L. & Hommais, F. Plasmid-chromosome cross-talks. *Environ. Microbiol.* **22**, 540–556 (2020).
- San Millan, A. & MacLean, R. C. Fitness costs of plasmids: a limit to plasmid transmission. *Microbiol. Spectr.* **5**, 1–12 (2017).
- Hall, J. P. J. et al. Plasmid fitness costs are caused by specific genetic conflicts enabling resolution by compensatory mutation. *PLoS Biol.* **19**, e3001225 (2021).
- Thompson, C. M. A. et al. Plasmid manipulation of bacterial behaviour through translational regulatory crosstalk. *PLoS Biol.* **21**, e3001988 (2023).
- Ni, Songwei et al. “Conjugative plasmid-encoded toxin-antitoxin system PrpT/PrpA directly controls plasmid copy number.”. *Proc. Natl Acad. Sci.* **118**, e2011577118 (2021).
- Wang, L. et al. A bacterial Argonaute with efficient DNA and RNA cleavage activity guided by small DNA and RNA. *Cell Rep.* **41**, 111533 (2022).
- San Millan, A. et al. Integrative analysis of fitness and metabolic effects of plasmids in *Pseudomonas aeruginosa* PAO1. *ISME J.* **12**, 3014–3024 (2018).
- Takahashi, Y. et al. Modulation of primary cell function of host *Pseudomonas bacteria* by the conjugative plasmid pCAR1. *Environ. Microbiol.* **17**, 134–155 (2015).
- San Millan, A., Toll-Riera, M., Qi, Q. & MacLean, R. C. Interactions between horizontally acquired genes create a fitness cost in *Pseudomonas aeruginosa*. *Nat. Commun.* **6**, 1–8 (2015).
- Kawano, H. et al. A novel small RNA on the *Pseudomonas putida* KT2440 chromosome is involved in the fitness cost imposed by IncP-1 Plasmid RP4. *Front. Microbiol.* **11**, 1328 (2020).
- Dar, D., Dar, N., Cai, L. & Newman, D. K. Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution. *Science* **373**, eabi4882 (2021).
- Sheng, K., Cao, W., Niu, Y., Deng, Q. & Zong, C. Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat. Methods* **14**, 267–270 (2017).
- Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A. & Vogel, J. Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat. Microbiol. Lett.* **5**, 1202–1206 (2020).
- Xu, Z. et al. Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq. *Nat. Commun.* **14**, 5130 (2023).
- McNulty, R. et al. Probe-based bacterial single-cell RNA sequencing predicts toxin regulation. *Nat. Microbiol.* **8**, 934–945 (2023).
- Blattman, S. B., Jiang, W., Oikonomou, P. & Tavazoie, S. Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. *Nat. Microbiol.* **5**, 1192–1201 (2020).
- Ma, P. et al. Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. *Cell*, **186**, 1–15 (2023).

26. Wang, B. et al. Single-cell massively-parallel multiplexed microbial sequencing (M3-seq) identifies rare bacterial populations and profiles phage infection. *Nat. Microbiol.* **8**, 1846–1862 (2023).
27. Ankenbauer, A. et al. *Pseudomonas putida* KT2440 is naturally endowed to withstand industrial-scale stress conditions. *Microb. Biotechnol.* **13**, 1145–1161 (2020).
28. Gong, T. et al. Engineering *Pseudomonas putida* KT2440 for simultaneous degradation of carbofuran and chlorpyrifos. *Microb. Biotechnol.* **9**, 792–800 (2016).
29. Cyriaque, V., Jacquiod, S., Riber, L., Abu Al-soud, W. A., Gillan, D. C. & Sørensen, S. J. et al. Selection and propagation of IncP conjugative plasmids following long-term anthropogenic metal pollution in river sediments. *J. Hazard. Mater.* **382**, 121173 (2020).
30. Jacquiod, S. et al. Deciphering conjugative plasmid permissiveness in wastewater microbiomes. *Mol. Ecol.* **26**, 3556–3571 (2017).
31. Hall, J. P. J., Wood, A. J., Harrison, E. & Brockhurst, M. A. Source-sink plasmid transfer dynamics maintain gene mobility in soil bacterial communities. *Proc. Natl Acad. Sci.* **113**, 8260–8265 (2016).
32. Cyriaque, V. et al. Lead drives complex dynamics of a conjugative plasmid in a bacterial community. *Front. Microbiol.* **12**, 1–14 (2021).
33. Bahl, M. I., Hansen, L. H. & Sørensen, S. J. Impact of conjugal transfer on the stability of IncP-1 plasmid pKJK5 in bacterial populations. *FEMS Microbiol. Lett.* **266**, 250–256 (2007).
34. Røder, HenrietteLyng et al. Biofilms can act as plasmid reserves in the absence of plasmid specific selection. *npj Biofilms Microbiomes* **7**, 78 (2021).
35. Yuste, L. et al. Growth phase-dependent expression of the *Pseudomonas putida* KT2440 transcriptional machinery analysed with a genome-wide DNA microarray. *Environ. Microbiol.* **8**, 165–177 (2006).
36. Van der Rest, M. E., Frank, C. & Molenaar, D. Functions of the membrane-associated and cytoplasmic malate dehydrogenases in the citric acid cycle of *Escherichia coli*. *J. Bacteriol.* **182**, 6892–6899 (2000).
37. Miller, W. G., Leveau, J. H. & Lindow, S. E. Improved gfp and inaZ broad-host-range promoter-probe vectors. *Mol. Plant Microbe Interact.* **13**, 1243–50 (2000).
38. Vincent, M. S. & Uphoff, S. Bacterial phenotypic heterogeneity in DNA repair and mutagenesis. *Biochem. Soc. Trans.* **48**, 451–462 (2020).
39. Perrin, E. et al. Diauxie and cointegration of carbon sources can coexist during bacterial growth in nutritionally complex environments. *Nat. Commun.* **11**, 3135 (2020).
40. Solopova, A. et al. during bacterial diauxic shift. *Proc. Natl Acad. Sci.* **111**, 7427–7432 (2014).
41. Li, L. et al. Plasmids persist in a microbial community by providing fitness benefit to multiple phylotypes. *ISME J.* **14**, 1170–1181 (2020).
42. Rysz, M., Mansfield, W. R., Fortner, J. D. & Alvarez, P. J. J. Tetracycline resistance gene maintenance under varying bacterial growth rate, substrate and oxygen availability, and tetracycline concentration. *Environ. Sci. Technol.* **47**, 6995–7001 (2013).
43. Fulde, M. et al. The arginine-ornithine antiporter arcD contributes to biological fitness of streptococcus suis. *Front. Cell. Infect. Microbiol.* **4**, 1–10 (2014).
44. Sakanaka, A., Kuboniwa, M., Takeuchi, H., Hashino, E. & Amano, A. Arginine-ornithine antiporter ArcD controls arginine metabolism and interspecies biofilm development of *Streptococcus gordonii*. *J. Biol. Chem.* **290**, 21185–21198 (2015).
45. Kim, N. H. et al. Real-time transposable element activity in individual live cells. *Proc. Natl Acad. Sci. USA* **113**, 7278–7283 (2016).
46. Humphrey, S. et al. Staphylococcal phages and pathogenicity islands drive plasmid evolution. *Nat. Commun.* **12**, 5845 (2021).
47. Ibarra-Chávez, R., Brady, A., Chen, J., Penadés, J. R. & Haag, A. F. Phage-inducible chromosomal islands promote genetic variability by blocking phage reproduction and protecting transductants from phage lysis. *PLoS Genet.* **18**, e1010146 (2022).
48. Luecken, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* **15**, e8746 (2019).
49. Sonesson, C. & Robinson, M. Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods* **15**, 255–261 (2018).

Acknowledgements

This work was supported by Villum Fonden (BioRep-HGT, Grant No. 00028304) and by the Fond National de Recherche Scientifique (F.R.S.-FNRS). A.K. and G.S. acknowledge support from the Department of Energy Office of Science, Biological and Environmental Research (BER) Program, Grant DE-SC0023091. We acknowledge the use of computing resources at the core facility for biocomputing at the Department of Biology, University of Copenhagen. We would like to thank Prof. Lars Hestbjerg Hansen and Assistant Prof. Witold Piotr Kot for help in sequencing to verify the quality of the initial microSPLi-seq libraries.

Author contributions

V.C. and J.S.-M. conceived the study and designed the laboratory experiments. V.C. and R.I.-C. carried out the laboratory experiments and conducted the analysis of the experimental data. J.S.-M. and J.N. contributed to materials and analysis tools. A.K. and G.S. provided protocols and expertise. All authors discussed the theory and wrote the manuscript.

Competing interests

G.S. is a co-founder and shareholder of Parse Biosciences, a scRNA-seq company. The rest of the authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-49793-x>.

Correspondence and requests for materials should be addressed to Valentine Cyriaque or Jonas Stenlørkke Madsen.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://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) 2024