Microbiome

METHODOLOGY

Open Access

ProFiT-SPEci-FISH: a novel approach for linking plasmids to hosts in complex microbial communities at the single-cell level

Alvah Zorea[1,2,3], Sarah Moraïs[1,2,3], David Pellow[4], Orly Gershoni-Yahalom[5], Maraike Probst[1,2,3], Sapir Nadler[1,2,3], Ron Shamir[4], Benyamin Rosental[5], Natalie Elia[1,2] and Itzhak Mizrahi[1,2,3*]

Abstract

Background Plasmids are influential drivers of bacterial evolution, facilitating horizontal gene transfer and shaping microbial communities. Current knowledge on plasmid persistence and mobilization in natural environments is derived from community-level studies, neglecting the single-cell level, where these dynamic processes unfold. Pinpointing specific plasmids within their natural environments is essential to unravel the dynamics between plasmids and their bacterial hosts.

Results Here, we overcame the technical hurdle of natural plasmid detectability in single cells by developing SPEci-FISH (Short Probe EffiCIent Fluorescence In Situ Hybridization), a novel molecular method designed to detect and visualize plasmids, regardless of their copy number, directly within bacterial cells, enabling their precise identification at the single-cell level. To complement this method, we created ProFiT (PRObe FInding Tool), a program facilitating the design of sequence-based probes for targeting individual plasmids or plasmid families.

Conclusions We have successfully applied these methods, combined with high-resolution microscopy, to investigate the dispersal and localization of natural plasmids within a clinical isolate, revealing various plasmid spatial patterns within the same bacterial population. Importantly, bridging the technological gap in linking plasmids to hosts in native complex microbial environments, we demonstrated that our method, when combined with fluorescence-activated cell sorting (FACS), can track plasmid-host dynamics in a human fecal sample. This approach identified multiple potential bacterial hosts for a conjugative plasmid that we assembled from this fecal sample's metagenome. Our integrated approach offers a significant advancement toward understanding plasmid ecology in complex microbiomes.

*Correspondence:
Itzhak Mizrahi
imizrahi@bgu.ac.il
[1] National Institute of Biotechnology in the Negev, Ben-Gurion University of the Negev, 8410501 Be'er Sheva, Israel
[2] Department of Life Sciences, Ben-Gurion University of the Negev, 8410501 Be'er Sheva, Israel
[3] The Goldman Sonnenfeldt School of Sustainability and Climate Change, Ben-Gurion University of the Negev, 8410501 Be'er Sheva, Israel
[4] Blavatnik School of Computer Science, Tel Aviv University, 69978 Tel Aviv, Israel
[5] The Shraga Segal Department of Microbiology, Immunology, and Genetics, Faculty of Health Sciences, Regenerative Medicine and Stem Cell Research Center, Ben Gurion University of the Negev, Be'er Sheva, Israel

Introduction

Plasmids play a vital role in bacterial evolution and adaptation, but limited studies have examined their ecology, function, and distribution in natural systems. This is primarily due to several limitations, including the lower quantities of plasmid DNA compared to chromosomal DNA, the low copy numbers of plasmids within microbial cells, and the challenges in accessing diverse bacterial host cells. Consequently, most studies that explore the distribution and function of uncultivated plasmids across and within their ecosystems use metagenomic methods such as plasmidome sequencing [1–3] and bioinformatic tools designed for the assembly of circular plasmid sequences [4–8], or the classification of plasmid contigs among metagenomic reads [9–11], which only approximate the plasmid hosts' range [2, 12–21]. However, examining plasmidomes solely at the metagenomic level neglects fundamental processes occurring at the single-cell level, such as plasmid spatial distribution within individual microbes and transfer between them. These aspects can profoundly influence plasmid maintenance and dispersal within and across the population. For example, the random distribution of high-copy-number plasmids throughout the cell can ensure transfer to daughter cells, while low-copy-number plasmids rely on active partitioning mechanisms to ensure plasmid presence in both daughter cells [22–25]. Moreover, plasmid dispersal across hosts can create source-sink mechanisms, influencing plasmid frequency within and across the population [26]. To study plasmids on the single-cell level, culture-dependent methods or transformation are commonly used [27–29], narrowing the plasmid range to cultivable plasmids. In addition, studies commonly genetically alter cultivated plasmids to facilitate their detection using selective genes or marker genes [30–35], thereby possibly modifying their behaviors. Furthermore, current microscopy-based, single-cell, culture-independent methods are limited to identifying plasmids within known hosts, thereby neglecting the identification of unknown plasmid hosts within complex communities and metagenomes. This leaves the critical task of identifying plasmids' hosts within whole metagenomes unresolved [36]. Other methods that preserve plasmid-host associations in metagenomes prior to sequencing include proximity-ligation techniques like Hi-C [37] and 3C [38], as well as emulsion-based approaches such as EpicPCR [39]. While these methods leverage the physical proximity of plasmid DNA to host DNA, further validation is necessary to confirm the bacterial host and eliminate false positives. Additionally, in the case of Hi-C, the DNA assembly process remains non-trivial and challenging, adding another layer of complexity to the accurate identification of plasmid-host relationships. Hence, to fully understand the ecological forces controlling plasmid effects on ecosystems, it is essential to assess their distribution within individual microbial hosts as well as across populations that form natural communities, providing insights into the community dynamics and evolution. Specifically, this includes understanding how plasmids maintain themselves within the population without selection regimes and how they control their dispersal to avoid extinction from their host populations and communities, and determining the extent to which plasmids are transferred across different microbial cells and hosts in natural systems. To cope with these challenges, our study introduces an innovative bioinformatic and molecular pipeline to target and study plasmids within their hosts in their native environments. ProFiT—SPEci-FISH pipeline stands for PRObe FInding Tool together with Short Probe EffiCIent Fluorescence In-Situ Hybridization. Our pipeline is designed to cope with major hurdles encountered when studying plasmids in their natural bacterial host cell and enables plasmid detection in high specificity within cells using traditional methods such as microscopy and FACS. ProFiT [40] first identifies multiple short probes for multiple plasmids or plasmid genes, while considering all available DNA contigs in the sample to significantly reduce cross-reactivity. This is followed by SPEci-FISH, which leverages these multiple probes and also introduces an enzymatic amplification step boosting the fluorescent signal at the target sequence, a crucial step for pinpointing plasmids within their native environments. After validating the adaptability of our pipeline by optimizing it on two commercial high-copy and low-copy number plasmids, we applied it to four plasmids found in a clinically relevant *Escherichia coli* strain [41], gaining valuable insights into their dispersal and distribution at the population level. By combining our pipeline with high-resolution Stochastic Optical Reconstruction Microscopy (STORM), we achieved an understanding of the spatial distribution of these plasmids within single cells. We were also able to co-localize two plasmids at a time within their bacterial host by dual labeling, shedding light on their interactions and behaviors. Furthermore, we demonstrated that our novel pipeline can specifically detect plasmids of interest in complex samples, identify their bacterial hosts, and recapitulate their dynamics within the communities. The ProFiT and SPEci-FISH methods could be used individually or together as versatile and powerful tools for studying plasmid dynamics, distribution, and interactions. Our approach offers a significant advancement in the detection and analysis of plasmid-host dynamics by allowing high-throughput screening of cells using FACS. This robust approach overcomes the limitations of microscopy-based techniques, which are often constrained by spatial mapping

of known microbial hosts and plasmids and are primarily applicable to biofilms or whole tissues [36]. Unlike existing methodologies that depend heavily on pre-existing sequencing data and prior knowledge of target taxa, we can identify all plasmid-containing taxa, including those with unknown identities, allowing us to detect horizontal gene transfer (HGT) events that might otherwise be overlooked in traditional sequencing workflows. By overcoming the challenges of probe design, optimizing techniques for different plasmid types, and offering insights into single-cell, population and community levels, our methods pave the way for an enhanced understanding of microbial communities and their mobile genetic components [42].

Results

ProFiT-SPEci-FISH: a pipeline for plasmid detection within their bacterial hosts in their natural environment

To address the limitations of studying plasmid ecology in natural systems, we developed a comprehensive pipeline integrating both wet-lab and computational steps. This pipeline consists of SPEci-FISH, a novel molecular technique utilizing multiple short probes to efficiently target plasmids within environmental samples. Additionally, it incorporates ProFiT, a computational tool which designs specific probes by considering the vast diversity of genes in metagenomic data to reduce the likelihood of probe adherence to non-plasmid DNA (Fig. 1). The use of both tools requires the plasmidome of the samples of interest (such as soil, feces, or marine samples, Fig. 1). This could be achieved by direct extraction of the overall plasmidome using molecular techniques [1, 3, 43] and/or by assembling plasmids from whole metagenomic DNA sequencing using specific plasmid assemblers [4, 44]. SPEci-FISH harnesses the combined strengths of single-molecule FISH (smFISH) [45], which uses multiple short probes to achieve specificity and signal amplification, and Catalyzed Reporter Deposition FISH (CARD-FISH) [46], which incorporates an enzymatic amplification step for further signal enhancement.

By utilizing these methods' strengths, we can detect plasmids with exceptional sensitivity and specificity, even in cases where their copy numbers are low, owing to the combined effect of multiple probes and the cascade of enzymatic reactions that amplify the fluorescent signals (as specified in materials and methods and Fig. 1 scheme). Nevertheless, SPEci-FISH maintains a high level of specificity by using multiple short probes binding to many targets and lowering the chance that incidental misbinding to incorrect targets would result in sufficient signal intensity, enabling clear differentiation between true signals and any potential off-target signals.

Fluorescent cells bearing the plasmid or plasmid genes of interest can then be specifically examined or separated using fluorescence-activated cell sorting (FACS). Subsequently, their bacterial hosts can be identified by amplifying and sequencing the 16S rRNA gene of the sorted cells containing the plasmids. To address nonspecific binding in SPEci-FISH, a hurdle often exacerbated by the diverse genetic makeup of environmental samples, we developed ProFiT [40]. This program offers a streamlined yet robust solution for accurate targeting of plasmid DNA. ProFiT conducts a comprehensive screen of all DNA contigs in a sample to identify and select probes that specifically target the desired plasmid DNA, while ensuring high specificity, effectively distinguishing the target DNA from the myriad of other genetic material present in the sample. By employing a greedy algorithm approach, ProFiT optimizes probe selection, prioritizing those with the highest prevalence in the target DNA while minimizing the risk of misbinding to unrelated sequences. This ranking process guarantees that only the most suitable probes are chosen, bolstering the reliability and accuracy of subsequent experimental results. ProFiT also facilitates the creation of probe sets optimized for targeting entire groups of plasmids or plasmid genes, such as those associated with mobilization or antibiotic resistance. Users may customize the number of degenerate nucleotides in the output, enabling the probes to effectively capture a broader range of targets. This feature not only enhances labeling efficiency but also significantly reduces costs associated with probe synthesis and experimentation. Furthermore, ProFiT enables users with other customizable parameters, granting the flexibility to define probe lengths and Limiting the number of probes binding to undesired targets. This versatility enables effective

(See figure on next page.)

Fig. 1 Illustration of the ProFiT-SPEci-FISH pipeline to explore plasmid dynamics and determine their bacterial host in environmental samples. The sample undergoes sequencing, and plasmids are assembled using SCAPP [4]. ProFiT is then employed to design multiple small, specific digoxigenin-labeled probes, tailored for the plasmids or plasmid genes of interest. These probes are ordered and subsequently hybridized to the target plasmid using the SPEci-FISH method. Probes are then detected by anti-digoxigenin antibodies, conjugated to horseradish peroxidase (HRP). This enzyme catalyzes multiple reactions with tyramide-labeled fluorophores, which serve as the substrate, leading to a localized and amplified fluorescent signal that can be visualized using a fluorescence microscope or analyzed using FACS. Dotted lines represent optional steps in the pipeline

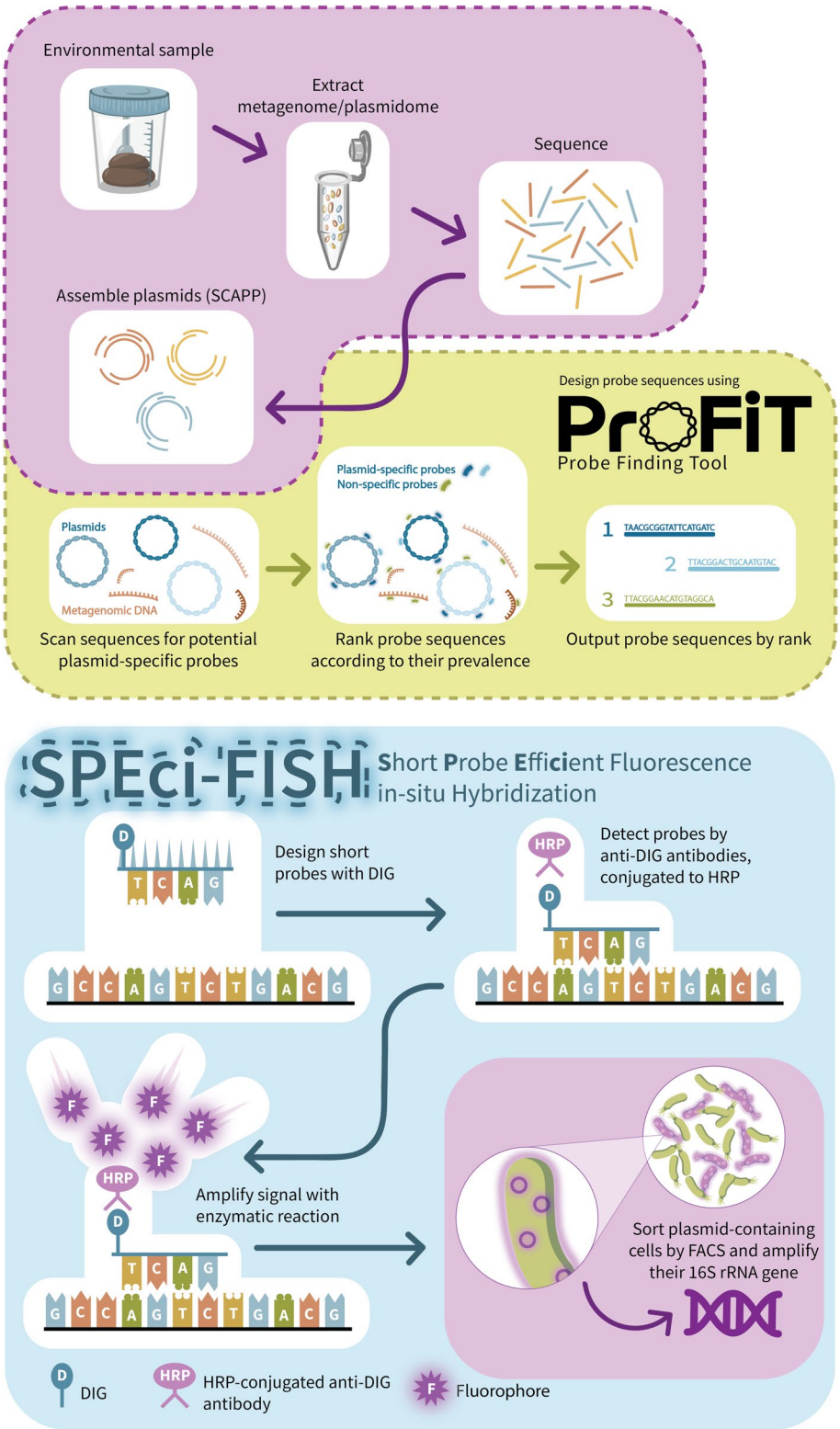

Fig. 1 (See legend on previous page.)

targeting of closely related sequences, ensuring robust and reliable results even in scenarios where sequences are not 100% identical. ProFiT outputs two files, one contains each probe sequence, along with a list of all positions in the target sequence(s) where it matches, and all positions in the negative sequences where it misbinds [40]. The second file lists, for each probe, the new target sequences that are fully covered by adding this probe to the set of previously listed probes, and the new negative sequences that would be incorrectly covered by adding this probe. This setup allows users to determine the desirability of adding each specific probe to their experiment. Probes in each file are ordered based on how few negative sequences and how many new target sequences they match.

Efficient targeting and labeling of multiple plasmids of varying sizes and copy numbers within single cells and populations

To test the SPEci-FISH method on different plasmid copy numbers, we initially implemented the method on commonly used and well-characterized high-copy number (pUC19, > 100 copies/cell) and low-copy number (pSC101, ~ 5 copies/cell) plasmids (Fig. 2Ai). These plasmids were each introduced into the *E. coli* TG1 strain through bacterial transformation. We used ProFiT to design specific probes to target these plasmids and to fluorescently label them using SPEci-FISH. This effort enabled the precise targeting and successful visualization of the plasmids within their bacterial hosts. Labeling efficiency was determined by using FACS to quantify the proportion of fluorescently labeled cells carrying plasmids, compared to control cells lacking the plasmids but subjected to the same labeling procedure. To optimize signal detection, we evaluated sets of 5, 10, 15, and 20 probes targeting the LCN plasmid. Using 20 probes yielded the highest signal-to-noise ratio, calculated as the geometric mean fluorescence intensity of each sample divided by that of its corresponding plasmid-lacking TG1 control (Fig. S1A). While the number of probes can be adjusted depending on the specific plasmid and sample, we used 20 probes across all experiments for consistency.

Our results show that only 0.21% of control cells (lacking the plasmid) exhibited a signal, whereas 99.2% and 98.2% of cells containing the plasmid exhibited a signal for the low-copy number and high-copy number plasmids, respectively (Fig. 2Aii, S1B). These results suggest that the SPEci-FISH method demonstrates a very high signal-to-noise ratio, enabling accurate counting of plasmid-containing cells, while minimizing off-target signals. In these tested populations, most cells carry the plasmids, aligning with the commercial design purposes of these plasmids for industrial and research applications aimed at maximizing the number of plasmid-containing cells to increase gene expression efficiency [47, 48].

Furthermore, to assess the detection sensitivity of the SPEci-FISH approach in native microbial samples, we performed a spiking experiment using a rumen sample (Fig. S1B). A known quantity of *E. coli* cells harboring the high-copy number plasmid pUC19 ($\sim 10^4$ cells) was added to 1 mL of rumen content, which is estimated to contain $\sim 10^9$ total microbial cells per milliliter. This corresponds to a relative abundance of approximately 0.001% for the spiked plasmid-bearing cells. Following SPEci-FISH labeling and FACS analysis, we observed an increase in the fraction of fluorescently labeled cells from 3.36% in the unspiked control to 6.93% in the spiked sample. This ~ 2-fold increase suggests that the method is capable of detecting plasmid-carrying cells at abundances as low as 1 in 10^5 cells, under conditions where the plasmid is present at high copy number and the probes are well optimized. These results demonstrate the method's potential for identifying rare plasmid-host associations in diverse microbial communities.

Next, we explored more complex samples to test the specificity of ProFiT probe design and SPEci-FISH plasmid labeling within single cells and populations. To do so, we used an *E. coli* ST131 clinical isolate population, E2022, previously shown to carry multiple plasmids [41]. We reassembled plasmids from the reads of this strain using SCAPP [4]. Two plasmids matched the two fully assembled plasmids previously submitted to the public database. We also successfully assembled additional plasmids that had not been fully resolved in earlier work [41] (Fig. S2A, S2B). Using ProFiT and SPEci-FISH, we targeted and analyzed four plasmids of various sizes within isolated cells, assessing their distribution across the cellular population and within individual cells (Figs. 2Ai and 3). Probe sets were specifically designed for these plasmids using ProFiT and applied effectively with SPEci-FISH (Fig. 2Ai). FACS was then used to quantify the targeted plasmids (Fig. 2Aii, S1), providing valuable insights into their distribution across the bacterial population. Interestingly, we observed varying prevalence of each plasmid within the same population, indicating that different portions of the isolated population carried different plasmids. This variation suggests a potential distribution of labor, possibly governed by the balance between positive and negative fitness effects associated with each plasmid.

Intrigued by the single-cell heterogeneity within microbial communities and its potential impact on plasmid dynamics, we harnessed the power of SPEci-FISH to colocalize plasmids within their bacterial hosts. This allowed us to assess the degree of overlap between the

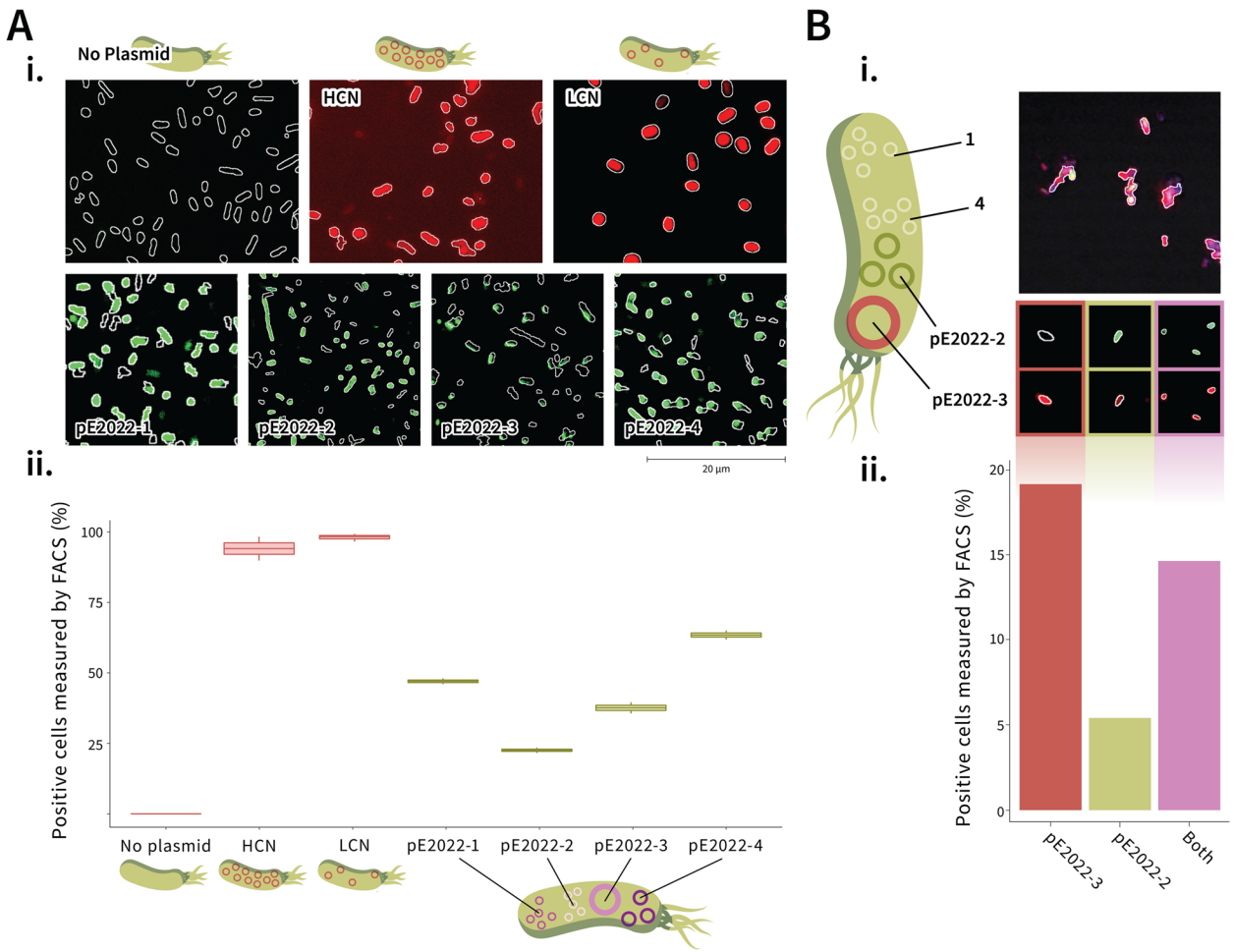

Fig. 2 Microscopic visualization and FACS analysis of SPEci-FISH labeled plasmids. **A** (i) Microscopy images showing SPEci-FISH labeled *E. coli* TG1 cells containing no plasmid, high-copy number (HCN) and low-copy number (LCN) plasmids, as well as *E. coli* E2022 cells containing plasmids pE2022-1 to pE2022-4. Cells were stained with Alexa 647 (red) or 488 (green). (ii) Corresponding FACS analyses for these samples, depicting the proportions of positively labeled cells in each sample. **B** (i) Microscopy images of *E. coli* E2022 cells with plasmids pE2022-3 and pE2022-2 simultaneously labeled using SPEci-FISH, stained with Alexa 647 and 488, respectively. (ii) Corresponding FACS analyses showing the proportions of positively labeled cells containing both plasmids, pE2022-3 only, or pE2022-2 only

different *E. coli* E2022 subpopulations carrying the different plasmids. To apply this application of SPEci-FISH, we employed probe sets that specifically targeted two distinct plasmids within the E2022 strain (Fig. 2Bi, Bii, S1B). To perform dual labeling, the SPEci-FISH protocol was modified to include two consecutive hybridization steps, with each of the plasmid probe sets (see Fig. 5 and materials and methods). Specifically, SPEci-FISH was performed in two consecutive steps once for each plasmid using a specific probe and fluorophore. Using this plasmid dual-labeling approach, we could detect intracellular variability with respect to the plasmid entities they carry. We successfully simultaneously targeted pE2022-2, a 2,080 bp small plasmid without annotated genes and pE2022-3, a larger conjugative plasmid of 33,144 bp (Fig. 2Bi, S2A, S2B). We applied FACS using a specific channel to count instances of cells carrying only one of these plasmids individually vs. co-occurrence of both within single cells, ultimately allowing us to better understand plasmid dispersal at the single-cell level and connect it to plasmid physiology (Fig. 2Bii).

To assess the co-occurrence of plasmids within individual cells, we simultaneously labeled both pE2022-2 and pE2022-3 in the same *E. coli* population. We found that 33.7% of the cells carried pE2022-3, the larger conjugative plasmid, and 20% carried pE2022-2, the smaller non-conjugative plasmid (Fig. 2Bii, S1B). Notably, 14.6% of the cells harbored both plasmids. These results indicate that pE2022-3 appeared independently in 19.1% of the cells, while pE2022-2 was found without the conjugative plasmid in only 5.4% of the cells. Importantly, the individual prevalence rates for each plasmid closely matched

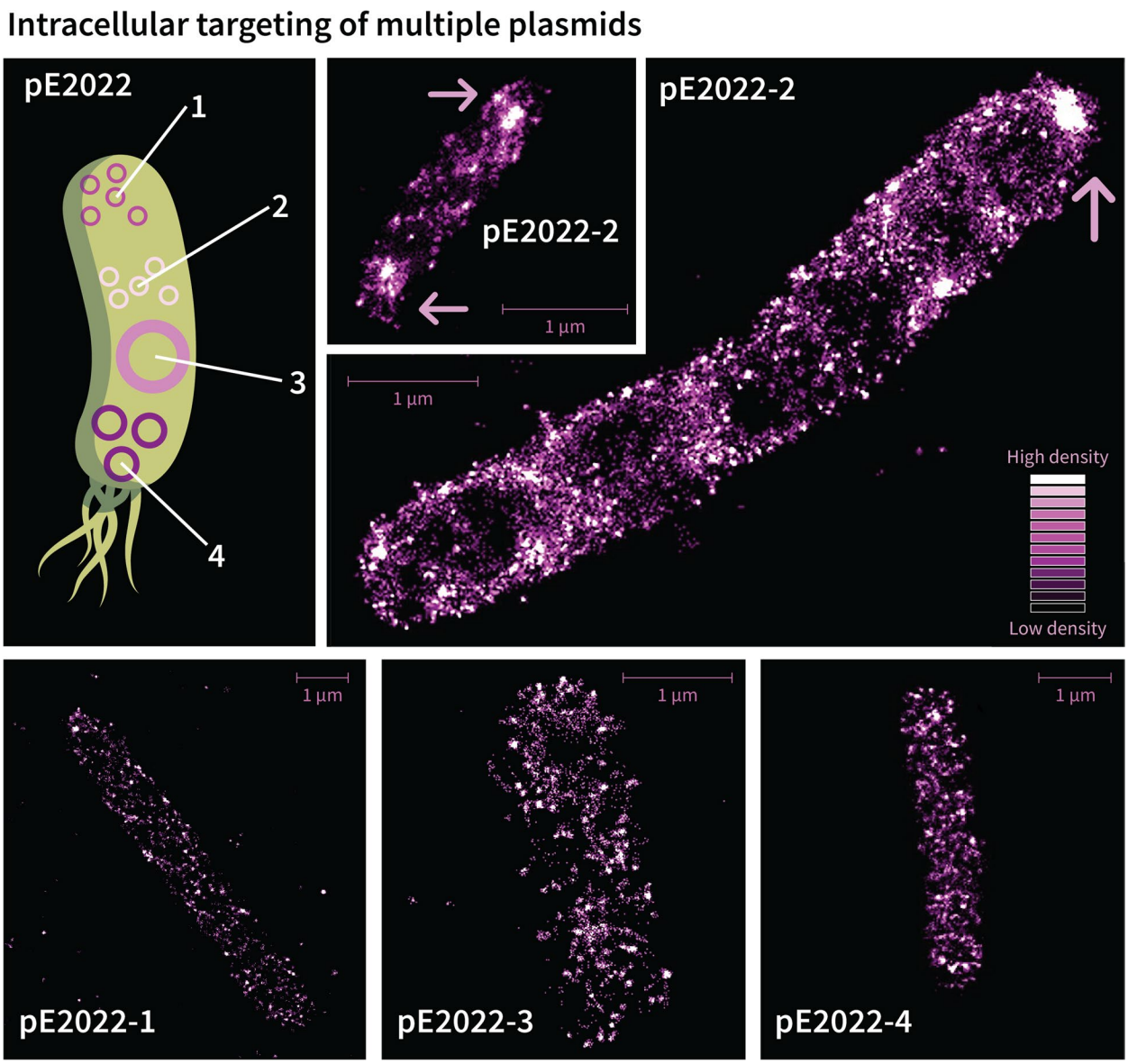

Fig. 3 E2022 plasmids visualized at the single-cell level using STORM. Images of four plasmids present in the isolate E2022, labeled using SPEci-FISH and acquired by Stochastic Optical Reconstruction Microscopy (STORM), see Methods section. Colors reflect the relative normalized density in each image (white, maximum density; black minimum density), see bar in top right image. Arrows indicate pE20222-2 plasmid accumulation at the cell poles. Scale bars are indicated

those observed when we labeled and analyzed each plasmid separately, further validating the method's accuracy (Fig. 2Aii).

These findings may be explained by the mobility potential of the larger plasmid, which, despite its size, can migrate between cells, thereby increasing its prevalence. In contrast, the smaller, non-conjugative plasmid may be more vulnerable to dilution effects, drift, and selective pressures that reduce its prevalence among microbial cells. An alternative explanation is that the smaller plasmid relies on the presence of the conjugative plasmid for its mobilization and long-term maintenance. Further investigation is needed to uncover the underlying mechanisms. Overall, these findings underscore the strength of our pipeline to reveal and study intricate, hidden genetic interactions at the single-cell level within microbial communities.

Determining the intracellular distribution of multiple plasmids using SPEci-FISH

In addition to examining the distribution of plasmids across bacterial populations, SPEci-FISH can also provide insights into their localization within individual cells. Here, we used Stochastic Optical Reconstruction Microscopy (STORM) for the intracellular visualization of four plasmids of the strain E2022 labeled by SPEci-FISH (pE2022-1 to pE2022-4, Fig. 3).

Understanding the spatial distribution of plasmids within bacterial cells could be leveraged for unraveling their impact on microbial physiology and ecology. There are conflicting findings regarding plasmid distribution models, specifically for those present in high-copy numbers, with some studies suggesting a focus of plasmids at the cell poles due to restricted diffusion through the nucleoid, and others reporting a more even distribution [49–56]. Our observations of plasmids within strain E2022 revealed a dynamic interplay between these distribution patterns. Specifically, the fluorescent signals from plasmid pE2022-2 were predominantly localized at the cell poles and outer perimeter of the cell rather than evenly distributed throughout the central region of the cell. This observation aligns with previous findings indicating that plasmids often face limitations in their ability to freely disperse within the nucleoid at the cell's center, resulting in their tendency to aggregate at the cellular poles [49, 51, 56, 57]. Alternatively, we observed that the remaining three plasmids displayed uniform distributions of the fluorescent signal within the cell, without a distinct polar preference. This is consistent with previous studies that have demonstrated that certain plasmids are randomly distributed throughout cells [54, 55]. These plasmid distribution patterns shed light on the multifaceted nature of plasmid dynamics within single cells, raising questions about the underlying mechanisms that govern these diverse behaviors and their implications for microbial population dynamics.

Single-cell analysis using ProFiT and SPEci-FISH reveals broad host range of a promiscuous plasmid in-situ

Identifying plasmids in their native contex is of great importance and a major challenge in the field. To address this, we applied our pipeline and successfully targeted a de novo assembled plasmid of interest (Fig. 4). We detected this plasmid's host in its native environment and found that it is promiscuous, sharing more than one host from different phyla. After assembling the plasmidome from a human gut shotgun sequence, we compiled a list of plasmids for analysis (Fig. 4A). Among them, we identified a ~40 kbp plasmid, which we named pAIZM1, that carries genes such as mobilization proteins and a zeta toxin-antitoxin system. This plasmid had 94.69%% similarity over 84% of its length to the *Bacteroides fragilis* plasmid pBF9343, corroborating the reliability of our assembly pipeline. This similarity presented the opportunity to use an already isolated strain and its plasmid as a control (Figure S3A). We leveraged ProFiT, designed to detect shared probes for multiple plasmids, to create probes shared between these two plasmids. We validated these probes on the isolated pBF9343 and its host before targeting the natural plasmid in its environment (Figure S3A). We applied SPEci-FISH followed by FACS to two human fecal samples: one in which the plasmid was bioinformatically detected and another in which it was not (Fig. 4A, 4B, S3B, S4A). In the sample where the plasmid was not detected, no significant difference was observed in the fluorescent signal between conditions with or without probe addition (Figure S3B). To further confirm probe specificity, we applied nonsense probes designed to have no target within the sample, as verified bioinformatically, and indeed, no signal was detected (Figure S3C). In contrast, in the sample where the plasmid was detected, we observed a significant 3- to fourfold increase in fluorescent signal (Fig. 4B, S4A). We then sorted the plasmid-containing cells and performed 16S rRNA gene amplicon sequencing of the sorted cells, as well as the original sample (Figure S4B). Interestingly, this analysis revealed an enrichment of 20 bacterial species, including

(See figure on next page.)

Fig. 4 Detecting plasmid mobility in gut environments. **A** The plasmidome of a healthy human fecal sample was assembled to identify a specific plasmid of interest. Probes were designed using ProFiT, and the plasmid was labeled in-situ with SPEci-FISH. Fluorescent cells were sorted using FACS and their 16S rRNA gene was amplified and sequenced to identify the bacterial hosts. **B** A boxplot showing the proportions of positively labeled cells harboring pAIZM1 in a fecal sample, analyzed by FACS. The top panel represents the proportions without SPEci-FISH labeling (control), while the bottom panel represents the proportions with SPEci-FISH labeling. **C** Potential bacterial hosts (strain level) of pAIZM1 were determined by sequencing the cells with and without the plasmid, and comparing the abundances. The bar plot shows abundances significantly higher within the fraction containing the plasmid versus the fraction without the plasmid. Colors represent the different phyla. **D** Simultaneous labeling of pAIZM1 with SPEci-FISH and the 16S rRNA gene with standard FISH for two of its potential hosts, *B. fragilis* (top) and *R. bromii* (bottom). pAIZM1 was labeled with Alexa Fluor 647 (right) and the 16S rRNA gene was labeled with Alexa Fluor 488 (left). FACS analyses in the boxplot on the right depict the proportions of positively labeled cells in each bacterial host. Pie charts summarize the proportion of each bacterial host with and without the plasmid

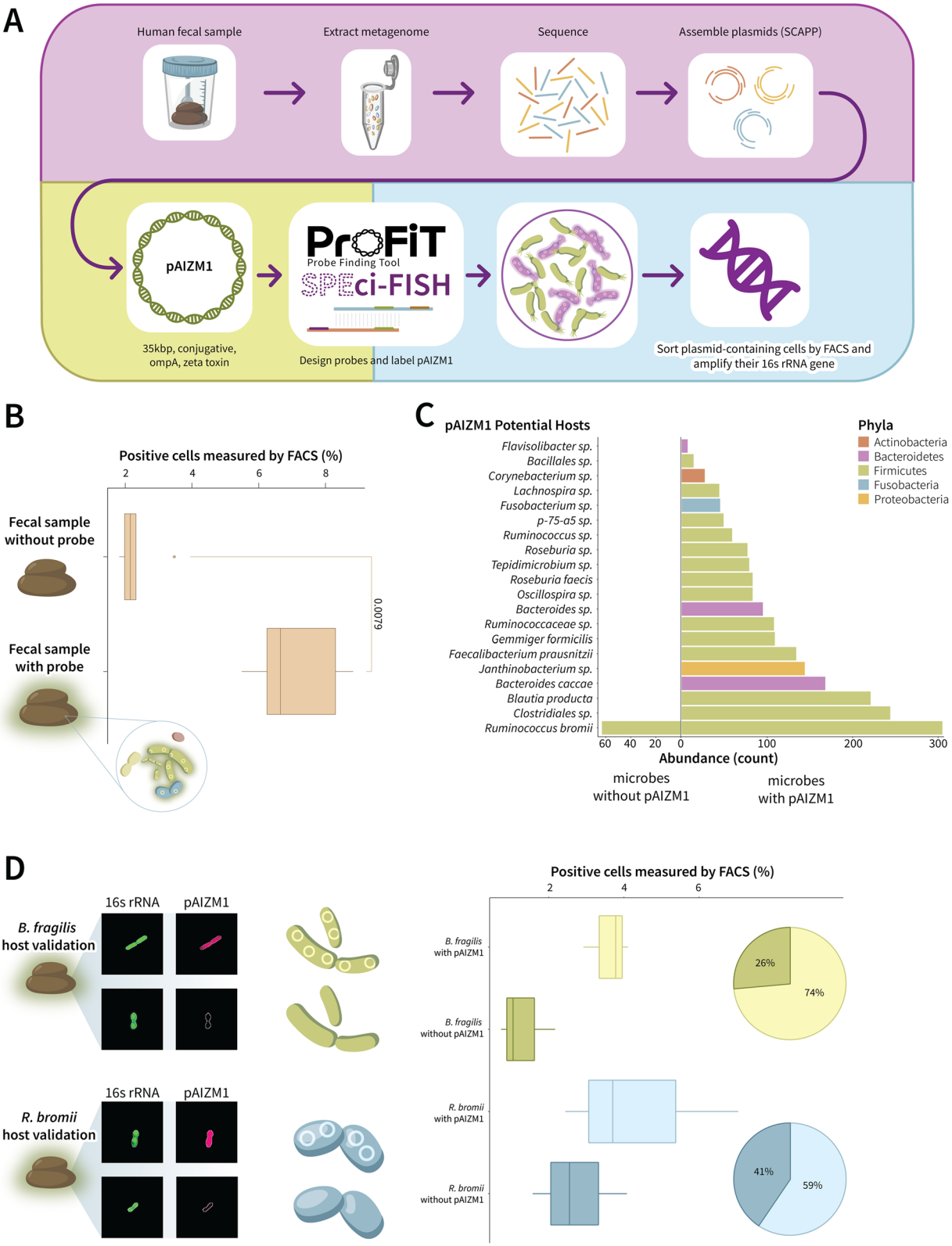

Fig. 4 (See legend on previous page.)

Ruminococcus bromii, a keystone species involved in resistant starch degradation in the human gut, suggesting that plasmid pAIZM1 is associated with multiple hosts from distinct bacterial phyla (Fig. 4C). To test this hypothesis and quantify the plasmid-carrying cells within each microbial species, we chose *R. bromii* and *B. fragilis* as potential hosts to target, due to the availability of validated specific 16S rRNA probes for these microbes [58, 59]. We employed these 16S rRNA probes for dual-fluorescence labeling of *B. fragilis* and *R. bromii*, alongside the AIZM1 plasmid probes, to precisely detect and quantify the proportion of cells carrying the plasmid within their entire specific population (Fig. 4D). Our results showed that, on average, 74% of *B. fragilis* cells and 59% of *R. bromii* cells carry the AIZM1 plasmid. This observation suggests that AIZM1 is compatible with both hosts. However, the higher prevalence of the AIZM1 plasmid within the *B. fragilis* population indicates greater compatibility with this species. These results may suggest a metacommunity patch sorting dynamics scenario from the plasmid's perspective and a source-sink dynamic between the two species' populations, with *B. fragilis* serving as the source and *R. bromii* as the sink as previously shown for multiple plasmids [26]. The plasmid exhibits higher fitness in one patch (*B. fragilis*) compared to the other (*R. bromii*). However, these dynamics potentially increase the probability of plasmid maintenance at the community level, ensuring its persistence even if one of the species is excluded. Moreover, this demonstrates how our pipeline can be leveraged to study plasmid host ranges in their natural environments and corroborates the methods' robustness and specificity.

Discussion

Here, we developed a comprehensive pipeline for plasmid targeting and visualization within single cells, allowing us to study plasmid dynamics within their microbial hosts in natural environments at various levels of resolution, from the single-cell to the community level. At the single cell level, when applying ProFiT and SPEci-FISH in combination with high-resolution microscopy to visualize the spatial arrangement of the *E. coli* clinical isolate plasmid pE2022-2 within individual cells, we observed plasmid clusters at the cell poles. This observation aligns with findings from previous research indicating that plasmids exhibit limited mobility within the central region of the cell, potentially due to nucleoid hindrance, leading to their accumulation at the cell poles [49, 51, 56, 57]. Alternatively, we observed a more uniform distribution of plasmids from the *E. coli* clinical isolate, pE2022-1, pE2022-3, and pE2022-4 throughout the cells, indicative of random dispersion throughout the cellular cytoplasm, a pattern which has also been identified for other plasmids in previous works [54, 55]. These observations suggest that different plasmids employ distinct distribution dynamics within the same population's cells, underscoring the importance of investigating the underlying mechanisms governing these distribution behaviors and their ecological implications. At the population level, our pipeline enabled the analysis of the prevalence of 4 plasmids of *E. coli* isolate E2022, as well as simultaneous targeting and colocalization of plasmids in the population. Unlike the commercial plasmids, none of the examined plasmids were detected across all cells in the population, suggesting a division of labor in their maintenance. This observation has far-reaching implications for microbial ecology and evolution, indicating plasmid-driven intrapopulation genetic heterogeneity that might have significant phenotypic consequences. For example, such genetic heterogeneity could enable specific subpopulations carrying ecologically relevant plasmid genes to invade and adapt to new environments inaccessible to the rest of the population. This could include the ability to utilize specific materials or withstand toxins, thereby creating evolutionary bottlenecks, promoting divergence and diversification within the different systems, and potentially leading to speciation. Additionally, our single-cell co-occurrence analysis revealed distinct dispersal patterns throughout the population that could be connected to plasmid physiology where the conjugative plasmid, pE2022-3, showed higher prevalence across the cells compared to the non-conjugative pE2022-2. This observation could be attributed to the plasmid lifestyle, where the conjugative plasmid is more prone to migration. Additionally, it is tempting to speculate that the non-conjugative plasmid pE2022-2 may rely on the conjugative plasmid pE2022-3 for mobilization, given that pE2022-2 is often found in conjunction with pE2022-3. This could be due to potential hitchhiking of pE2022-2 on the mobilization machinery of pE2022-3, even in the absence of a mob gene on pE2022-2, a notion recently suggested [60]. Using the ProFiT-SPEci-FISH pipeline, potential dependencies between natural plasmids could be studied to detect plasmid population dynamics and potential interactions.

At the community level, we examined plasmid distribution patterns in situ within a complex human fecal sample. This allowed us to identify potential hosts for a plasmid detected bioinformatically, further narrowing the technological gap in identifying uncultured plasmids and their hosts in native microbial ecosystems. We assembled plasmids from a human metagenome sample and identified a conjugative plasmid with high similarity to the *B. fragilis* pBF9343 plasmid, based on BLAST results. This plasmid, which we termed pAIZM1, was selected for further analysis of its host range. We targeted

this plasmid using a set of 20 short probes efficiently designed with ProFiT, and labeled it through SPEci-FISH. This procedure allowed us to identify 20 potential hosts of plasmid AIZM1 by sorting plasmid-containing cells and using 16S rRNA amplicon sequencing to annotate these hosts. Moreover, we confirmed the presence of the plasmid within two potential hosts, *B. fragilis* and *R. bromii*, using dual labeling of their 16S rRNA genes alongside the plasmid, followed by FACS analysis. Interestingly, these two hosts, as well as the other putative ones, belong to very divergent phylogenies of different phyla. This finding aligns with previous research indicating that plasmid transfer can naturally cross large phylogenetic barriers [13, 61]. Moreover, the existence of natural broad host plasmids has been examined in the past, with cases reporting over a hundred predicted bacterial hosts per plasmid [14, 32]. Additionally, in experimental setups introducing broad host plasmids to microbially rich habitats such as soil, these plasmids were found to invade unexpectedly diverse phylogenies [32]. Collectively, these observations suggest that plasmids naturally tend to show higher promiscuity with regard to their hosts, and might also explain the high ratio of HGT often observed in nature [62]. Our findings also suggest a source-sink dynamic between the potential hosts, which could contribute to the maintenance and prevalence of the plasmid, in agreement with previous research [26, 63].

While the ProFiT–SPEci-FISH pipeline offers a novel and powerful framework for detecting plasmids and linking them to their bacterial hosts in their native ecosystems, several limitations should be acknowledged. First, although the method does not rely on prior taxonomic knowledge, probe design through ProFiT requires plasmid sequence assemblies. In environments where assembly is particularly challenging (e.g., due to low coverage or high complexity), designing specific probes may be more difficult. Second, applying the method to different plasmids or plasmid-encoded genes may require recalibration of the hybridization and labeling conditions, as variations in sequence composition, secondary structure, and plasmid copy number can affect probe binding efficiency and signal intensity. Additionally, fixation and permeabilization protocols may need to be tailored to specific sample types to ensure effective probe entry and target accessibility, and the number of probes may also require calibration accordingly. Finally, in natural samples, heterogeneity in plasmid carriage [64] within a specific target host cell population may limit our ability to optimize the labeling conditions, as we cannot readily determine whether the plasmid has been sufficiently labeled without additional experimental validation, such as single-cell real-time PCR. Despite these challenges, the ProFiT–SPEci-FISH pipeline remains a versatile and powerful approach for investigating plasmid dynamics across single-cell, population, and community levels within native microbial ecosystems. Its modular design allows for flexible adaptation at each step, enabling broad application to different samples. In addition to its simplicity, this approach is advantageous as it does not require culturing or plasmid isolation processes, making it convenient for unsupervised targeting and identification of plasmid hosts. This enhances our understanding of plasmid functionality, localization, and distribution within the specific environment by enabling the detection of horizontal gene transfer events via plasmids, which is pivotal for unraveling the intricate workings of microbial communities.

Materials and methods

Strains and plasmids

Plasmids pUC19 and pSC101 were acquired from A2S technologies Ltd (Yavne, IL) and Biological Industries Ltd. (Beit-Haemek, IL), respectively, and subsequently transformed into chemocompetent cells of the *E. coli* strain TG1. An *E. coli* isolate (E2022) from a clinical study of a patient with a urinary tract infection [41] was kindly sent to us by Prof. Fernando de la Cruz. We ordered two 1 kbp DNA sequences from A2S technologies Ltd (Yavne, IL), which each correspond to specific DNA sequences present on two different plasmids from E2022. These sequences were subsequently inserted into pUC19 and pSC101 plasmids. This strategic approach allowed us to use the same probes to target multiple plasmids in all these strains, reducing overall costs significantly. Additionally, we ordered the *B. fragilis* strain DSMZ 2151, which harbors a plasmid (pBF9343, accession number: CR626928.1), from the Deutsche Sammlungvon Mikroorganismen und Zellkulturen (DSMZ).

TG1 plasmid-containing cells were grown overnight at 37 °C in LB medium, supplemented with the corresponding antibiotics: ampicillin for pUC19 and tetracycline for pSC101. Similarly, E2022 was grown under similar conditions, with ampicillin. *B. fragilis* was grown anaerobically overnight at 37 °C in YCFA medium.

Plasmid assembly

Paired-end reads of the E2022 bacterial strain were downloaded from the European Nucleotide Archive (ENA), accession number: PRJEB6262. These were trimmed and cleaned using Trim Galore v2.6 [65] and assembled into contigs by Megahit v1.0.3 [66]. These were then assembled into plasmids by SCAPP v0.1.4 [4], resulting in 4 plasmids. These plasmid sequences were deposited in GenBank (accession numbers: OR345921 to OR345924). The same process was done on a fecal sample which was sequenced locally. The plasmidome was assembled, and

a plasmid was chosen for targeting within the fecal sample. This assembled plasmid, which we named pAIZM1, exhibited a high similarity (95% identity over 84% of the assembled plasmid's length) to plasmid pBF9343.

Probe design using ProFiT: PRObe FInding tool

To simplify the process of choosing multiple probe sets to target many plasmids, we developed ProFiT (PRObe FInding Tool). This program accepts a List of target plasmid or plasmid gene sequences as the positive set, along with other genetic sequences found in the sample as the negative set and attempts to cover the positive set without misbindings in the negative set, in a greedy fashion. The program scans the sequences to count the number of positive and negative sequences hit by all possible probes of a given probe length with a user-defined maximum number of degenerate nucleotides. The probes are then ranked by their prevalence in the target set, with the probes with misbinding in more than a maximal number of negative sequences excluded. Probes are added to the probe set in ranked order, after filtering candidates that overlap within 3 nt of already placed probes or exceed the maximal allowed misbindings in the negative set. Upon adding each probe, the number of probes matching each target sequence is updated, and the program terminates when all target sequences are covered by a given minimum number of probes. If the algorithm exhausts the set of possible probes without covering all positive sequences, it begins adding probes with misbinding in the negative set using a similar greedy process to ensure few false positives until the entire positive set is covered. For each probe in the created probe set, the positions of hits in the positive and negative sequences are reported for further tailoring by the user. ProFiT provides customizable parameters, allowing users to specify the probe length, the maximum number of probes that can misbind in a negative sequence, the minimum number of probes needed to cover each positive sequence, and the maximum number of degenerate nucleotides allowed, which effectively enables the targeting of similar sequences, even if they are not 100% identical.

A set of 20 short DNA probes (20 nt) was designed for each plasmid using ProFiT. These were ordered (HyLabs, Rehovot, Israel) with digoxigenin (DIG) modifications added to the 5' end of the probes (Table S1).

SPEci-FISH: Short Probe EffiCIent Fluorescence in-situ Hybridization

SPEci-FISH includes fixation, permeabilization, hybridization, and enzymatic signal amplification steps (Figs. 1 and 5). In this method, cells were harvested from isolates or fecal samples (~10^9 cells) and fixed with 4% PFA by incubation at room temperature for 15 min with a tube rotator. Fixed cells were then collected by centrifugation, washed once with PBS, and stored at 4 °C until further use. All centrifugations were carried out at room temperature, for 5 min at 8000 g. To prevent cell clumping during PBS washes, 0.1% pluronic [Thermo Fisher Scientific, Massachusetts, United States] was added (final concentration: 0.001%).

For *E. coli* isolates, the fixed cells were resuspended in water, and ethanol was added gradually for cell permeabilization, up to a final concentration of 70%. The cells were then incubated at room temperature for 2 h. Subsequently, the cells were centrifuged, and the resulting pellet was resuspended in 50 µL of 30% formamide hybridization buffer, prepared based on a previously published protocol [67] (900 mM NaCl, 20 mM Tris–HCl [pH 8], 0.01% [wt/vol] sodium dodecyl sulfate [SDS], 30% [vol/vol] formamide [Sigma-Aldrich, Missouri, United States]). Next, 10 µL of the probe mix (100 µM, final concentration: 20 µM) was added to all reactions, except for the negative control, and incubated overnight at 35 °C for hybridization. The optimal concentrations of formamide and probes, as well as hybridization temperatures, were determined through calibration experiments (data not shown).

On the following day, 800 µL PBS was added to the cells, which were then centrifuged. The pellet was washed twice with a washing solution (WS) prepared as described previously [68] (95 mM NaCl, 20 mM Tris–HCl [pH 8], 5 mM EDTA [pH 8], 0.01% [wt/vol] sodium dodecyl sulfate [SDS]). The cells were resuspended in 1 mL WS for each wash, incubated at 37 °C for 10 min, and pelleted. Pellets were then resuspended in 1 mL antibody binding solution (10% [vol/vol] Western blocking reagent [Roche Diagnostics, Indiana, United States] in PBS) and incubated at room temperature for 45 min, with a tube rotator. Subsequently, 2 µL Anti-Digoxigenin antibody [Sigma-Aldrich, Missouri, United States] was added and cells were incubated for another 1.5 h. Cells were centrifuged and washed twice in PBS at room temperature for 10 min each, with a tube rotator.

Cell pellets were centrifuged and resuspended in 100 µL amplification solution prepared based on a previously published protocol [66] (2 M NaCl, 10% [wt/vol] dextran sulfate, and 0.1% [wt/vol] blocking reagent [Roche Diagnostics, Indiana, United States] in PBS). Then, 1 µL 0.15% H_2O_2 was added (freshly diluted) to a final concentration of 0.0015%, together with 0.25 µL Alexa-labeled tyramide (Alexa Fluor 647 or 488, resuspended in 150 µL DMSO) [Thermo Fisher Scientific, Massachusetts, United States]. Cells were incubated at 46 °C for 45 min in the dark, then 800 µL PBS was added, and cells were centrifuged. Two additional washing rounds with PBS were performed and the cells were counterstained with

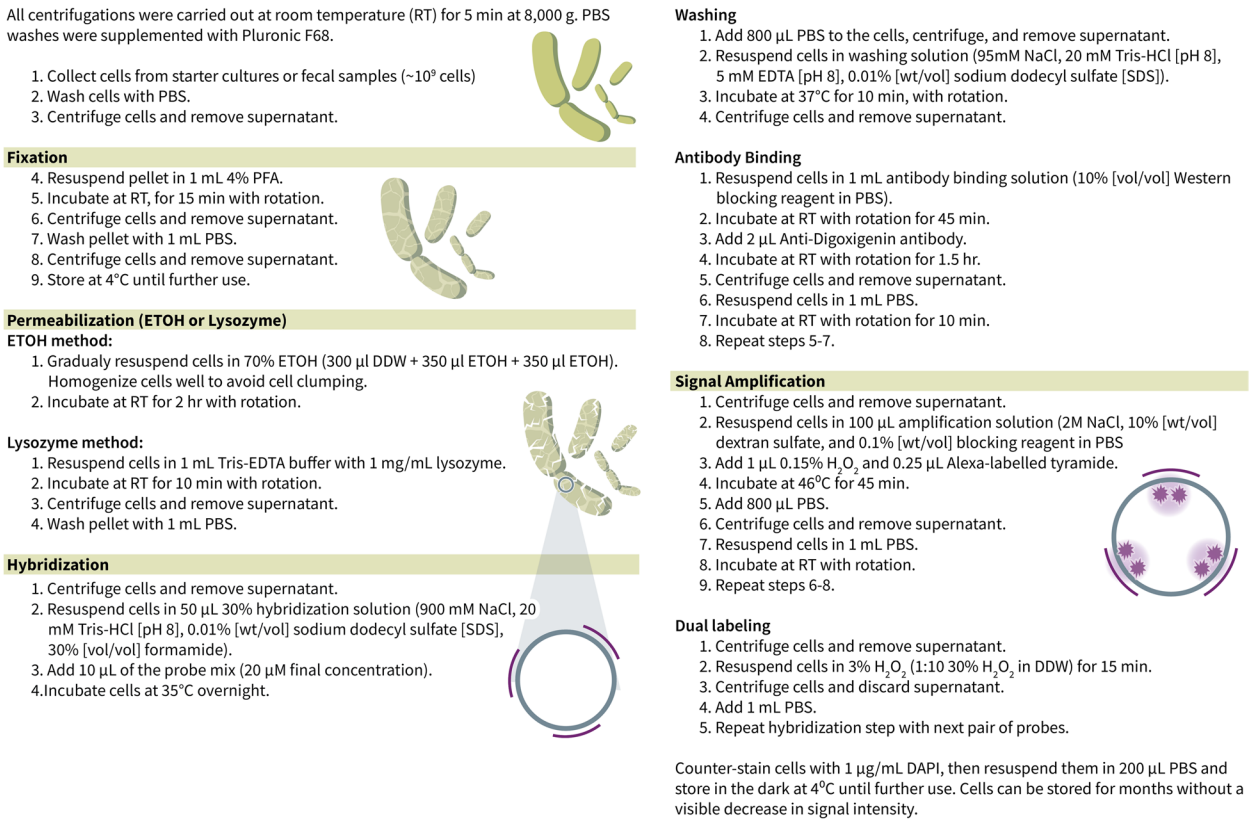

Fig. 5 SPEci-FISH protocol. The workflow of SPEci-FISH includes fixation, permeabilization, hybridization, and enzymatic signal amplification

4, 6-diamidino-2-phenylindole stain (DAPI, 1 µg/mL). Then, the cell pellets were directly resuspended in 200 µL PBS and stored at 4 °C before imaging.

In the case of dual labeling, DAPI was not added after the first hybridization round. Instead, cells were washed in 3% H_2O_2 for 15 min with a tube rotator, to deactivate excess HRPs. Cells were then centrifuged and washed once in PBS. After washing the cells, the procedure was repeated starting from the hybridization step, using the second set of probes and a different Alexa-labeled tyramide to achieve two fluorescent signals simultaneously. After the second hybridization round, cells were stained with DAPI.

Stochastic Optical Reconstruction Microscopy (STORM)

Twenty microliters of labeled E2022 *E. coli* cells were immobilized on poly-L-lysine-treated STORM high glass bottom 35 mm µ-Dish (ibidi, 81,158). To induce photoswitching of Alexa-647, the cells were immersed in STORM imaging buffer. The buffer was prepared by combining a solution of 50 µL MEA (77 mg Cysteamine in 1 mL of Buffer A), 20 µL Gloxy (13.39 mg glucose oxidase in 1 mL of buffer A), and 930 µL Buffer B (10% glucose in Buffer A). Buffer A stock solution contains 1 M TRIS at pH 8 and 5 M NaCl. Images were collected on the Zeiss Elyra PS1 inverted wide field fluorescence microscope using a Zeiss 100X N.A 1.46 in epifluorescence mode. 641 nm laser was used at 100% power, and 405 nm irradiation was applied as needed. 10,000 images were collected in each experiment with a frame-rate of 55 fps using irradiation intensities of ~1–3 kW cm^{-2}. Single particle localization, fitting, drift correction and reconstruction was done in Zen (Zeiss).

Fluorescence-activated cell sorting

For sorting and flow cytometry analyses we used a Sony MA900 FACS machine. The determination of bacterial cells in each sample was established by assessing size, granularity, and DAPI intensities positive. The sorting gates were defined based on the fluorescence signals emitted by the probes (Alexa Fluor 488 and 647), allowing the separation of cells containing the target plasmids. The sorted cells were collected in sterile tubes for sequencing analyses. Subsequent FACS analyses were conducted using FlowJo v10.

Fecal sample pretreatment and DNA extraction

To target plasmids in a complex environment, two human fecal samples were pre-treated using the protocol below. Bioinformatic analysis identified one sample as containing the plasmid pAIZM1, while the other did not. Both samples underwent the pretreatment steps, but DNA extraction and sequencing were performed only for the pAIZM1-positive sample. For pretreatment, feces in PBS 1:1 [w/vol] were homogenized by vortex, the homogenate was centrifuged at 4 °C for 20 min at 10,000 g, the pellet was dissolved 1:4 [w/vol] in extraction buffer (100 mM Tris–HCl, 10 mM EDTA, 0.15 M NaCl, 0.15% Tween 80 [v/v], pH 8.0), and incubated at 4 °C for 1 h. The suspension was then centrifuged at 500 g for 15 min at 4 °C to remove ruptured plant particles while keeping the bacterial cells in suspension. The supernatant was then passed through four layers of cheesecloth, centrifuged (10,000 g, 20 min, 4 °C), and the pellet was kept in TE 1:1 [w/v] at − 20 °C until DNA extraction. For the pAIZM1-positive sample, DNA was extracted as previously described [69]. Briefly, cells were lysed by bead disruption with phenol followed by phenol/chloroform DNA extraction. The final supernatant was precipitated with 0.6 volume of isopropanol and resuspended overnight in 50–100 µL TE (10 mM Tris–HCl, 1 mM EDTA), then stored at -20 °C. DNA was sequenced on the Illumina MiSeq platform [70], yielding approximately 9.5 million reads. Plasmids were assembled from the resulting paired-end reads as described above.

Plasmid-host association

To determine the bacterial hosts of pAIZM1 within a fecal sample, we used the sorted cells from the FACS analyses, with and without pAIZM1. The V4 region of the 16S ribosomal RNA gene subunit of these cells was amplified using 16S universal primers (ACTCCTACG GGAGGCAGCAGT and GTATTACCGCGGCTGCTG GCAC) with custom barcodes, and sequenced on the Illumina MiSeq platform [69], yielding ~ 7 k reads (subsampled to an even read depth of 6,631 reads). Using Qiime2 [71], the paired-end reads were then demultiplexed, trimmed using dada2 denoise-paired, and assigned to taxonomy using the Greengenes database v13-8 [72]. The constructed relative abundance barplot of OTUs was then viewed and exported using the online Qiime2 viewer.

Identifying pAIZM1 HGT events in a fecal sample

Once the predicted hosts of pAIZM1 were identified, we labeled two of these hosts along with the plasmid in situ. To accomplish this, we first labeled pAIZM1 using SPEci-FISH (with Alexa 647), as described previously, but with a modified permeabilization step [59]. Specifically, fixed cells were permeabilized using a Tris–EDTA buffer containing 1 mg/mL lysozyme [Sigma-Aldrich, Missouri, United States] for 10 min at room temperature when targeting *B. fragilis* or 10 mg/mL lysozyme for 1 h at room temperature when targeting *R. bromii*. This was followed by an additional washing step in PBS.

Subsequently, we utilized a previously described [73] 16 s rRNA FISH protocol to target either *B. fragilis* or *R. bromii* (using probes tagged with Alexa Fluor 488 5': Bfra 998 [58], GTTTCCACATCATTCCACTG, and Rbro 730 [59], TAAAGCCCAGYAGGCCGC). Briefly, the plasmid-lebelled cells were not counterstained with DAPI, and 50 µL 30% hybridization solution were added, as described above. Then, 4 µL of the relevant probe was added to all reactions aside from the negative control (50 ng/µL final concentration: 4 ng/µL), and incubated overnight at 35 °C for hybridization. On the following day, 800 µL PBS was added to the cells, which were then centrifuged. The pellet was washed once in 1 mL washing solution (WS) prepared as described previously [67] (65 mM NaCl, 20 mM Tris–HCl [pH 8], 5 mM EDTA [pH 8], 0.01% [wt/vol] sodium dodecyl sulfate [SDS]), incubated at 37 °C for 20 min, and pelleted. The cells were counterstained with 4,6-diamidino-2-phenylindole stain (DAPI, 1 µg/mL). Then, the cell pellets were directly resuspended in 200 µL PBS and stored at 4 °C before imaging.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40168-025-02238-z.

Supplementary Material 1: Figure S1. Evaluation of SPEci-FISH probe performance and representative FACS analyses. (A) Boxplot showing the signal-to-noise ratio of the geometric mean fluorescence intensity, comparing the use of 5, 10, 15 and 20 probes in the SPEci-FISH method. (B) FACS analyses of SPEci-FISH-labeled *E. coli* TG1 cells containing no plasmids, high-copy number (HCN) plasmids, or low-copy number (LCN) plasmids. Also shown are *E. coli* E2022 cells harboring plasmids pE2022-1 through pE2022-4, as well as simultaneously labeled cells for plasmids pE2022-2 and pE2022-3 using SPEci-FISH. Rumen samples labeled with SPEci-FISH- with and without spiking with 10^4 of the HCN plasmid- are also included. Cells were labeled with Alexa Fluor 647 (top three panels on the left and bottom two), Alexa Fluor 488, (top four panels on the right), or both dyes for dual labeling. Figure S2. pE2022 plasmid maps and genes. (A) Maps of plasmids pE2022-1 to pE2022-4 assembled by SCAPP, found in *E. coli* strain E2022. (B) Common plasmid encoded features and their presence on plasmids of strain E2022. Figure S3. Validation of SPEci-FISH specificity using isolates and fecal samples. (A) Microscopy images and FACS plots of *Bacteroides fragilis* harboring the naturally occurring plasmid pBF9343, labeled using SPEci-FISH with and without probes that are shared between pBF9343 and the pAIZM1 plasmid. Alexa Fluor 488 was used for detection. (B) A fecal sample that, based on bioinformatic metagenomic analysis, lacks the pAIZM1 plasmid shows comparable Alexa Fluor 488 and 647 intensities regardless of probe presence, indicating no off-target binding. (C) A fecal sample labeled with nonsense probes exhibits minimal fluorescence intensity, highlighting the specificity of the

SPEci-FISH method. Figure S4. Identification of potential hosts of pAIZM1 in a human fecal sample using SPEci-FISH. (A) FACS plots corresponding to Figure 4B showing a fecal sample, identified by metagenomic analysis to harbor plasmid pAIZM1, which was targeted using SPEci-FISH with Alexa Fluor 647. (B) Relative abundance of bacterial species based on 16S rRNA amplicon sequencing of the fecal sample (left) and sorted cells labeled for pAIZM1, showing only species identified as potential hosts of the plasmid (right). (C) FACS plots of dual labeling targeting both pAIZM1 and the 16S rRNA gene of two putative host species: *B. fragilis* (left) and *R. bromii* (right). The 16S rRNA gene was labeled with Alexa Fluor 488, and pAIZM1 was labeled with Alexa Fluor 647. Table S1 DNA sequences of probe sets, designed using ProFiT to target plasmids using SPEci-FISH.

Acknowledgements
The authors thank the scientific artist Daphne Perlman for the scientific illustrations and the assistance with graphic design in all figures of this paper and Dr Arik Segal for providing human fecal samples. The authors also thank Shani Talice for her help with the images taken in Figure 2Bi.

Authors' contributions
AZ and IM conceived the study and designed the experiments. AZ performed ProFiT-SPEci-FISH exeperiments with technical help from MP and SN. DP and RN developed ProFiT, OGY and BR conducted fluorence-cell sorting experiments, and NE conducted STORM experiments and analysed the data. AZ, IM and SM analysed the data and wrote the paper. IM secured funding supervised the work and synthesized the results. All authors read and approved the final manuscript.

Funding
This research was supported by the German-Israeli Project Cooperation (DIP) [Grant No. 2476/2–1], the European Research Council (ERC) [Grant No. 866530], and the Israel Science Foundation (ISF) [Grants No. 1947/19 and 979/25].

Data availability
Plasmids of strain E2022 that were assembled in this study are publicly available in the European Nucleotide Archive (ENA) database under accession numbers: OR345921 to OR345924, and plasmid pAIZM1 is available under accession PV929812. The ProFiT program can be downloaded from github [40]. The genome sequencing data of the fecal microbiome sample, as well as the 16S rRNA sequences of plasmid-hosting bacteria identified via SPEci-FISH and sorted using FACS, have been deposited in GenBank under accession number (PRJNA1161407).

Declarations

Ethics approval and consent to participate
The collected fecal sample was obtained from a study approved by the ethics committee of Soroka University Medical Center (SUMC, 0266–15). The participant provided written informed consent to participate in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 3 February 2025 Accepted: 19 August 2025
Published online: 02 December 2025

References

1. Brown Kav A, Benhar I, Mizrahi I. A method for purifying high quality and high yield plasmid DNA for metagenomic and deep sequencing approaches. J Microbiol Methods. 2013;95:272–9.
2. Jørgensen TS, Xu Z, Hansen MA, Sørensen SJ, Hansen LH. Hundreds of circular novel plasmids and DNA elements identified in a rat cecum metamobilome. PLoS One. 2014;9:e87924.
3. Jørgensen TS, Kiil AS, Hansen MA, Sørensen SJ, Hansen LH. Current strategies for mobilome research. Front Microbiol. 2014;5:750.
4. Pellow D, Zorea A, Probst M, Furman O, Segal A, Mizrahi I, et al. Scapp: an algorithm for improved plasmid assembly in metagenomes. Microbiome. 2021;9:144.
5. Rozov R, Brown Kav A, Bogumil D, Shterzer N, Halperin E, Mizrahi I, et al. Recycler: an algorithm for detecting plasmids from de novo assembly graphs. Bioinformatics. 2017;33:475–82.
6. Antipov D, Raiko M, Lapidus A, Pevzner PA. Plasmid detection and assembly in genomic and metagenomic data sets. Genome Res. 2019;29:961–8.
7. Müller R, Chauve C. HyAsP, a greedy tool for plasmids identification. Bioinformatics. 2019;35:4436–9.
8. Arredondo-Alonso S, Bootsma M, Hein Y, Rogers MRC, Corander J, Willems RJL, et al. Gplas: a comprehensive tool for plasmid analysis using short-read graphs. Bioinformatics. 2020;36:3874–6.
9. Pellow D, Mizrahi I, Shamir R. Plasclass improves plasmid sequence classification. PLoS Comput Biol. 2020;16:e1007781.
10. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb Genom. 2018;4. Available from: https://doi.org/10.1099/mgen.0.000206.
11. Schwengers O, Barth P, Falgenhauer L, Hain T, Chakraborty T, Goesmann A. Platon: identification and characterization of bacterial plasmid contigs in short-read draft assemblies exploiting protein sequence-based replicon distribution scores. Microb Genom. 2020;6. Available from: https://doi.org/10.1099/mgen.0.000398.
12. Shapiro JT, Zorea A, Kav AB, Ontiveros VJ, Mizrahi I, Pilosof S. Multilayer networks of plasmid genetic similarity reveal potential pathways of gene transmission. ISME J. 2023. Available from: https://doi.org/10.1038/s41396-023-01373-5.
13. Brown Kav A, Rozov R, Bogumil D, Sørensen SJ, Hansen LH, Benhar I, et al. Unravelling plasmidome distribution and interaction with its hosting microbiome. Environ Microbiol. 2020;22:32–44.
14. He W, Russel J, Klincke F, Nesme J, Sørensen SJ. The role of plasmids in the gut microbiome during the first year of life. bioRxiv. 2023. Available from: http://biorxiv.org/lookup/doi/https://doi.org/10.1101/2023.04.05.535656.
15. Zorea A, Pellow D, Levin L, Pilosof S, Friedman J, Shamir R, et al. Plasmids in the human gut reveal neutral dispersal and recombination that is overpowered by inflammatory diseases. Nat Commun. 2024;15:3147.
16. Ji Y, Shang J, Tang X, Sun Y. HOTSPOT: hierarchical host prediction for assembled plasmid contigs with transformer. Bioinformatics. 2023;39. Available from: https://doi.org/10.1093/bioinformatics/btad283.
17. Robertson J, Bessonov K, Schonfeld J, Nash JHE. Universal whole-sequence-based plasmid typing and its utility to prediction of host range and epidemiological surveillance. Microb Genom. 2020;6. Available from: https://doi.org/10.1099/mgen.0.000435.
18. Redondo-Salvo S, Fernández-López R, Ruiz R, Vielva L, de Toro M, Rocha EPC, et al. Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. Nat Commun. 2020;11:3602.
19. Acman M, van Dorp L, Santini JM, Balloux F. Large-scale network analysis captures biological features of bacterial plasmids. Nat Commun. 2020;11:2452.
20. Aytan-Aktug D, Clausen PTLC, Szarvas J, Munk P, Otani S, Nguyen M, et al. PlasmidHostFinder: Prediction of Plasmid Hosts Using Random Forest. mSystems. 2022;7:e0118021.
21. Krawczyk PS, Lipinski L, Dziembowski A. Plasflow: predicting plasmid sequences in metagenomic data using genome signatures. Nucleic Acids Res. 2018;46:e35.
22. Nordström K, Austin SJ. Mechanisms that contribute to the stable segregation of plasmids. Annu Rev Genet. 1989;23:37–69.
23. Williams DR, Thomas CM. Active partitioning of bacterial plasmids. J Gen Microbiol. 1992;138:1–16.
24. Baxter JC, Funnell BE. Plasmid Partition Mechanisms. Microbiol Spectr. 2014;2. Available from: https://doi.org/10.1128/microbiolspec.PLAS-0023-2014.
25. Wang Y. Spatial distribution of high copy number plasmids in bacteria. Plasmid. 2017;91:2–8.

26. Hall JPJ, Wood AJ, Harrison E, Brockhurst MA. Source-sink plasmid transfer dynamics maintain gene mobility in soil bacterial communities. Proc Natl Acad Sci U S A. 2016;113:8260–5.
27. Shintani M, Sanchez ZK, Kimbara K. Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. Front Microbiol. 2015;6:242.
28. Pukall R, Tschäpe H, Smalla K. Monitoring the spread of broad host and narrow host range plasmids in soil microcosms. FEMS Microbiol Ecol. 1996;20:53–66.
29. San Millan A, Heilbron K, MacLean RC. Positive epistasis between co-infecting plasmids promotes plasmid survival in bacterial populations. ISME J. 2014;8:601–12.
30. Arends K, Schiwon K, Sakinc T, Hübner J, Grohmann E. Green fluorescent protein-labeled monitoring tool to quantify conjugative plasmid transfer between gram-positive and gram-negative bacteria. Appl Environ Microbiol. 2012;78:895–9.
31. Klümper U, Droumpali A, Dechesne A, Smets BF. Novel assay to measure the plasmid mobilizing potential of mixed microbial communities. Front Microbiol. 2014;5:730.
32. Klümper U, Riber L, Dechesne A, Sannazzarro A, Hansen LH, Sørensen SJ, et al. Broad host range plasmids can invade an unexpectedly diverse fraction of a soil bacterial community. ISME J. 2015;9(4):934–45.
33. Tamhane T, Wolters BK, Illukkumbura R, Maelandsmo GM, Haugen MH, Brix K. Construction of a plasmid coding for green fluorescent protein tagged cathepsin L and data on expression in colorectal carcinoma cells. Data Brief. 2015;5:468–75.
34. Simonsen L, Gordon DM, Stewart FM, Levin BR. Estimating the rate of plasmid transfer: an end-point method. J Gen Microbiol. 1990;136:2319–25.
35. Fang P, Elena AX, Kunath MA, Berendonk TU, Klümper U. Reduced selection for antibiotic resistance in community context is maintained despite pressure by additional antibiotics. ISME Commun. 2023;3:52.
36. Grodner B, Shi H, Farchione O, Vill AC, Ntekas I, Diebold PJ, et al. Spatial mapping of mobile genetic elements and their bacterial hosts in complex microbiomes. Nat Microbiol. 2024;9:2262–77.
37. Belton J-M, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. Hi-C: a comprehensive technique to capture the conformation of genomes. Methods. 2012;58:268–76.
38. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002;295:1306–11.
39. Spencer SJ, Tamminen MV, Preheim SP, Guo MT, Briggs AW, Brito IL, et al. Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. ISME J. 2016;10:427–36.
40. ProFiT. Github; Available from: https://github.com/labmizrahi/ProFiT.
41. Lanza VF, de Toro M, Garcillán-Barcia MP, Mora A, Blanco J, Coque TM, et al. Plasmid flux in *Escherichia coli* ST131 sublineages, analyzed by plasmid constellation network (PLACNET), a new method for plasmid reconstruction from whole genome sequences. PLoS Genet. 2014;10:e1004766.
42. Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EPC, de la Cruz F. Mobility of plasmids. Microbiol Mol Biol Rev. 2010;74(3):434–52.
43. Norman A, Riber L, Luo W, Li LL, Hansen LH, Sørensen SJ. An improved method for including upper size range plasmids in metamobilomes. PLoS One. 2014;9:e104405.
44. Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A, Pevzner PA. Plasmidspades: assembling plasmids from whole genome sequencing data. Bioinformatics. 2016;32:3380–7.
45. Femino AM, Fay FS, Fogarty K, Singer RH. Visualization of single RNA transcripts in situ. Science. 1998;280:585–90.
46. Schönhuber W, Fuchs B, Juretschko S, Amann R. Improved sensitivity of whole-cell hybridization by the combination of horseradish peroxidase-labeled oligonucleotides and tyramide signal amplification. Appl Environ Microbiol. 1997;63:3268–73.
47. Wahle E, Kornberg A. The partition locus of plasmid pSC101 is a specific binding site for DNA gyrase. EMBO J. 1988;7:1889–95.
48. Lin-Chao S, Chen WT, Wong TT. High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. Mol Microbiol. 1992;6:3385–93.
49. Reyes-Lamothe R, Tran T, Meas D, Lee L, Li AM, Sherratt DJ, et al. High-copy bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are randomly partitioned at cell division. Nucleic Acids Res. 2014;42:1042–51.
50. Hsu T-M, Chang Y-R. High-copy-number plasmid segregation-single-molecule dynamics in single cells. Biophys J. 2019;116:772–80.
51. Yao S, Helinski DR, Toukdarian A. Localization of the naturally occurring plasmid ColE1 at the cell pole. J Bacteriol. 2007;189:1946–53.
52. Bouet J-Y, Funnell BE. Plasmid Localization and Partition in Enterobacteriaceae. EcoSal Plus. 2019;8. Available from: https://doi.org/10.1128/ecosalplus.ESP-0003-2019.
53. Wang Y, Penkul P, Milstein JN. Quantitative localization microscopy reveals a novel organization of a high-copy number plasmid. Biophys J. 2016;111:467–79.
54. Summers D. Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. Mol Microbiol. 1998;29:1137–45.
55. Summers DK, Sherratt DJ. Multimerization of high copy number plasmids causes instability: cole 1 encodes a determinant essential for plasmid monomerization and stability. Cell. 1984;36:1097–103.
56. Pogliano J, Ho TQ, Zhong Z, Helinski DR. Multicopy plasmids are clustered and localized in *Escherichia coli*. Proc Natl Acad Sci U S A. 2001;98:4486–91.
57. van Gijtenbeek LA, Robinson A, van Oijen AM, Poolman B, Kok J. On the spatial organization of mRNA, plasmids, and ribosomes in a bacterial host overexpressing membrane proteins. PLoS Genet. 2016;12:e1006523.
58. Rigottier-Gois L, Rochet V, Garrec N, Suau A, Doré J. Enumeration of *Bacteroides* species in human faeces by fluorescent in situ hybridisation combined with flow cytometry using 16S rRNA probes. Syst Appl Microbiol. 2003;26:110–8.
59. Harmsen HJM, Raangs GC, He T, Degener JE, Welling GW. Extensive set of 16S rRNA-based probes for detection of bacteria in human feces. Appl Environ Microbiol. 2002;68:2982–90.
60. Ares-Arroyo M, Coluzzi C, Rocha EPC. Origins of transfer establish networks of functional dependencies for plasmid transfer by conjugation. Nucleic Acids Res. 2023;51:3001–16.
61. Brown Kav A, Sasson G, Jami E, Doron-Faigenboim A, Benhar I, Mizrahi I. Insights into the bovine rumen plasmidome. Proc Natl Acad Sci U S A. 2012;109:5452–7.
62. Aminov RI. Horizontal gene exchange in environmental microbiota. Front Microbiol. 2011;2:158.
63. León-Sampedro R, DelaFuente J, Díaz-Agero C, Crellen T, Musicha P, Rodríguez-Beltrán J, et al. Pervasive transmission of a carbapenem resistance plasmid in the gut microbiota of hospitalized patients. Nat Microbiol. 2021;6:606–16.
64. Jahn M, Vorpahl C, Hübschmann T, Harms H, Müller S. Copy number variability of expression plasmids determined by cell sorting and droplet digital PCR. Microb Cell Fact. 2016;15:211.
65. Krueger F. Trim galore. Wrapper Tool Cutadapt FastQC Consistently Apply Qual Adapt Trimming FastQ Files. 2015;516. https://github.com/FelixKrueger/TrimGalore.
66. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31:1674–6.
67. Pernthaler A, Pernthaler J, Amann R. Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. Appl Environ Microbiol. 2002;68:3094–101.
68. Thiele S, Fuchs BM, Amann RI. Identification of microorganisms using the ribosomal RNA approach and fluorescence in situ hybridization. In: Wilderer P, editor. Treatise on Water Science. Burlington (MA): Academic Press; 2011. p.171–189.
69. Stevenson DM, Weimer PJ. Dominance of *Prevotella* and low abundance of classical ruminal bacterial species in the bovine rumen revealed by relative quantification real-time PCR. Appl Microbiol Biotechnol. 2007;75:165–74.
70. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012;6:1621–4.
71. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:852–7.

72. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590–6.
73. Rigottier-Gois L, Bourhis A-G, Gramet G, Rochet V, Doré J. Fluorescent hybridisation combined with flow cytometry and hybridisation of total RNA to analyse the composition of microbial communities in human faeces using 16S rRNA probes. FEMS Microbiol Ecol. 2003;43:237–45.

Publisher's Note