nature portfolio

Peer Review File

Plasmid-driven strategies for clone success in Escherichia coli

Corresponding Author: Professor Jukka Corander

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript describes the analysis of a large sample of 2000 E. coli isolates sampled in a longitudinal survey covering two decades. The use of long-read technologies allowed to produce a high-quality set of 4485 plasmid sequences present in these isolates. The manuscript focuses on the study of these plasmids. It has three somewhat disconnected parts. In the first part it describes the collection of plasmids. These are quite similar to previous sequenced ones as indicated by the Rand index (0.96). Hence, the degree of novelty is moderate, but allows to conclude there is already a good sampling of these plasmids in these E. coli. The second part quantifies the age of plasmids in the lineages. This reveals that some plasmids have been in lineages for hundreds of years, which is not uncommon in E. coli or other species (for example the plasmids of Shigella and EIEC remain in lineages for thousands of years), but would not be necessarily expected in ExPEC. Yet, the inference of the plasmid age raises concerns. This section concludes that large plasmids are key resources of bacterial adaptive traits, in contrast to small plasmids. The third part of the manuscript focuses on colicins present in small plasmids, which is puzzling considering the previous statement. This part assesses the function of the microcin V against a range of E. coli (shown before for this microcin). Results suggest the microcin gives some strains lacking antibiotic resistance genes competitive abilities. In summary, the manuscript reflects a lot of work, has relevant information, but lacks focus.

Timing of the acquisition of plasmids. The authors correlate the patristic distances between plasmids and between the core genes of bacteria. When correlation is not significantly different from zero they consider there was transfer of plasmid. When correlation is significantly different from zero they assume there was no transfer. This is incorrect because it will fail to identify transfer whenever there is a some significant joint evolution even in the presence of transfer. Suppose a plasmid was acquired independently in ten lineages and then vertically inherited. It will show a correlation between the two distances because the plasmid and the bacteria evolved together. Yet, there were 10 transfers. The problem in the context of this study is that the assumption leads to longer coalescence times. The consequence is that estimates of the age of the plasmid in the lineage are over-estimated. One needs clear evidence that plasmid and bacteria have similar phylogenies not just that their phylogenies are not completely unrelated.
A few examples. The assumption that pT1-1 was in the MRCA of ST131 is arguable and requires rigorous confirmation when analyzing the patterns of occurrence in figure 3. Same for pT 2-3 in D whose dating depends on a single occurrence in the upper part of the tree and clearly suggests independent acquisition in at least two lineages. Same for pT 13-1 which can only date from 1900 if there was an extraordinary level of loss across the lineage. Given these are conjugative plasmids carrying an adaptive novel trait that increases the likelihood that the transfers will be sampled, multiple transfer events do seem more parsimonious that one single ancestral acquisition.

Sampling affects the conclusions. The analyses were done based on clinical samples. This is a typical problem that leads to certain biases that may affect some analyses. For example, when the authors state that core plasmid functions diversify more than regions involved in virulence or resistance, it should be considered that the latter may have recently integrated different plasmid backbones and therefore that coalescent times are much shorter for the latter than the former. Similarly, when analyzing the evolution of plasmids one must consider they are spreading horizontally and transferring the trait that will lead to isolation of the clones. As such, its not surprising that certain plasmids are much more frequent than others (it's a consequence of sampling for bacteria with plasmid-encoded traits).

Supplementary material. The manuscript mentions the existence of 4 supplementary tables, but none was present in the site

for download.

Other comments:
Line 38. " Our analysis suggests that novel successful plasmid backbones emerge rarely " This needs qualifications. Is it that unexpected considering the period of acquisition of plasmids spanning decades? One can recognize plasmids today that were present in the collections of enterobacteria in the first half of the 20th century. If the current understanding is that plasmid backbones constantly emerge in short periods, then the relevant citations should be given.

I have some concerns regarding the phylogeny in Figure 1A. The scale of the tree indicates that the most distant groups of strains are over 0.2 subs/position away which seems incompatible with the idea that these are all E. coli strains. Maybe the scale is in other units? This should be indicated.

It's difficult to understand the information in Figure 1B.

Line 200. Why is it concluded that small plasmids are not a source of adaptive traits? Just because they only share replication and mobilisation machinery? They may provide adaptive functions that are more diverse (which is the case made by the last section of the manuscript where bacteriocins are in small plasmids). I don't understand how the containment analysis leads to such conclusion.

Line 208. I don't understand the argument that the transfer region diversified more than the others. To know they diversified more there must be a phylogenetic frame showing that other regions were co-inherited vertically and evolved slower, but the authors sustain (correctly in my opinion) that homologous functions were acquired independently. Such a statement requires quantification.

Figure S7. Scale unclear. Is it the standard substitutions/position (seems too high)? Maybe something different like the total number of differences?

Line 377. "Our analysis indicated that larger plasmids in E. coli are almost always associated with traits falling under bacterial competition, pathogenesis and antibiotic resistance. " This statement should be moderated. I assume many genes have unknown functions (if not this should be stated).

Reviewer #2

(Remarks to the Author)
In this study, Sergio Arredondo-Alonso and collaborators investigate the genetic basis underlying plasmid success in extra-intestinal pathogenic E. coli. They first obtained high quality closed genomes from 2,000 E. coli isolates recovered over two decades (form the Norwegian epidemiological longitudinal collection, NORM collection). Next, they elucidated and typed the ExPEC plasmidome in detail, and investigated if particular pTs contributed to the success of pandemic ExPEC clones. The authors found strong plasmid-lineage associations, with some bacteriocin-producing plasmids playing a potential key role in the success of specific clones. In general, I think that this is an extremely interesting work shedding new light in the role of plasmids in ExPEC evolution. I only have some minor comments for the authors:

Figure 1, Would it be possible to indicate in the figure (maybe in the upper labels in the plots in panel C), to which Inc / PTU group belongs each pT? that would help the reader to make sense of the type of plasmids in each pT. For example, the label could be “pT2-1 (IncF..)”

Figure 2 and S6, Include a colour legend for gene function?

Figure S3, It would be interesting to also show the results of “plasmid load” by correcting the plasmidome size in bp by the plasmid copy number (average coverage plasmid vs average coverage chromosome), which would give a better idea of the total number of bp devoted to plasmids.

-line 221, what is the reason to exclude plasmids smaller than 10 kb for this analysis?

Figure S9 panel B is partially deleted.

-Line 321 Fig. 45A -> Fig. 5A

-Section “Plasmid-encoded bacteriocins play a pivotal role in bacterial competition”. I may be mistaken, but if I understood correctly, the authors tested the microcin/colicin activity of isolates carrying microcin/colicin-producing plasmids, using as negative controls other isolates that do not carry the microcin/colicin-producing plasmids (or plasmids with deletions of one of the systems, according to figure 5?). It would be nice to have isogenic clones with or without the microcin/colicin-producing plasmids to definitely show that the inhibition is mediated by the plasmid-encoded colicins. The plasmid can be cured using, for example, CRISPR-based methods. Alternatively, the plasmid could be conjugated to a strain which does not carry the colicin-producing plasmid, and then supernatants of the plasmid-carrying/plasmid-free isogenic clones could be tested in parallel. Alternatively, the microcin/colicin clusters could be cloned in a vector. The control right now is using other isolates not harboring a microcin/colicin-producing plasmid, but it is not clear how different the genomes of both isolates are. Do they differ in the composition of colicins in the chromosome or other plasmids? The results are pretty convincing in

general, but it would be nice to have this isogenic control, if possible. Alternatively, the authors could explain a bit better why their control strains are adequate, because I think that it's not completely clear in the results section.

Alvaro San Millan

Reviewer #3

(Remarks to the Author)
The authors present a well written manuscript describing the plasmidome of bloodstream infection-associated E. coli from a Norwegian longitudinal epidemiological study. In-depth genomic analysis identifies successful plasmid-backbones within the ExPEC landscape and dates the acquisition of relevant plasmids in the major ExPEC sequence types ST69, ST73, ST95, and ST131.
This is then complemented by experimental work showcasing the functionality of two colicins carried by plasmids linked to parallel clonal expansion. This result is linked to the maintenance of susceptible lineages in the ExPEC landscape due their association with the colicins and also the efficacy of the colicins against multi-drug resistant lineages.

I recommend this manuscript for acceptance and suggest small changes to help with the accessibility of figures. On line 205 it would be useful to state the relevant colours in Figure 2 that pertain to the operons described as they are difficult to locate in the figure due to small text: "iroBCDEN (green), iutA-iucACBD and sitABCD (turquoise)". Alternatively, these text labels could be highlighted in Figure 2 in some way.
It would also aid the reader to include a ST (including ST131-A, B, C) annotation column in Figure 4B, as these results are described exclusively in the context of specific STs in the manuscript (lines 320-322).

Minor comments:

On lines 253, 278, 285, and 291, 0 is used instead of O when referring to O-types.
Line 296: missing '%'.
Line 303: missing space between % and CI.
Line 321: Is this referencing Fig 4A or 5A or both?
Line 346: include reference for this statement.
Line 367: full stop missing at end of sentence.
Line 380: use the symbol for beta for continuity with the rest of the manuscript.
Line 423-425: is there a suitable reference for this?
For rigour and reproducibility purposes, references should be added in the methods section for Guppy on line 515 and mlplasmid on line 529 (publication or GitHub url as appropriate).
Line 778: full stop missing at end of sentence.
Line 800: comma missing between 'chromosome' and 'plasmid'.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors took my comments in account and have solved all the key issues, including the lack of supplementary material tables, the scales of phylogenetic trees, and some statements that were not very well supported (in my opinion). I thank them for that and have very few remaining comments. This is an interesting and valuable study.
- Regarding the exclusion of the <10 kb plasmids, one should take into account that small plasmids do carry resistance genes and colicins. (This is actually the case of most plasmids with colicin A)
- some references lack page numbers of manuscript ids (45, 49, ...)

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The manuscript describes the analysis of a large sample of 2000 E. coli isolates sampled in a longitudinal survey covering two decades. The use of long-read technologies allowed to produce a high-quality set of 4485 plasmid sequences present in these isolates. The manuscript focuses on the study of these plasmids. It has three somewhat disconnected parts. In the first part it describes the collection of plasmids. These are quite similar to previous sequenced ones as indicated by the Rand index (0.96). Hence, the degree of novelty is moderate, but allows to conclude there is already a good sampling of these plasmids in these E. coli. The second part quantifies the age of plasmids in the lineages. This reveals that some plasmids have been in lineages for hundreds of years, which is not uncommon in E. coli or other species (for example the plasmids of Shigella and EIEC remain in lineages for thousands of years), but would not be necessarily expected in ExPEC. Yet, the inference of the plasmid age raises concerns. This section concludes that large plasmids are key resources of bacterial adaptive traits, in contrast to small plasmids. The third part of the manuscript focuses on colicins present in small plasmids, which is puzzling considering the previous statement. This part assesses the function of the microcin V against a range of E. coli (shown before for this microcin). Results suggest the microcin gives some strains lacking antibiotic resistance genes competitive abilities. In summary, the manuscript reflects a lot of work, has relevant information, but lacks focus.

We thank the reviewer for their detailed comments and for highlighting the amount of computational and experimental work that went into the manuscript.

The experimental work has focused on microcin V which was mainly present in pColV-like plasmids corresponding to pTs 9-1 and 10-1. As shown in Table S2, these are large plasmids (size > 100 kbp) which were clonally expanded by ST95 and clade B from ST131. To emphasize that the microcin V was associated with large plasmids, we have now indicated this in lines 319-320.

Timing of the acquisition of plasmids. The authors correlate the patristic distances between plasmids and between the core genes of bacteria. When correlation is not significantly different from zero they consider there was transfer of plasmid. When correlation is significantly different from zero they assume there was no transfer. This is incorrect because it will fail to identify transfer whenever there is a some significant joint evolution even in the presence of transfer. Suppose a plasmid was acquired independently in ten lineages and then vertically inherited. It will show a correlation between the two distances because the plasmid and the bacteria evolved together. Yet, there were 10 transfers. The problem in the context of this study is that the assumption leads to longer coalescence times. The consequence is that estimates of the age of the plasmid in the lineage are over-estimated. One needs clear evidence that plasmid and bacteria have similar phylogenies not just that their phylogenies are not completely unrelated.
A few examples. The assumption that pT1-1 was in the MRCA of ST131 is arguable and requires rigorous confirmation when analyzing the patterns of occurrence in figure 3. Same for pT 2-3 in D whose dating depends on a single occurrence in the upper part of the tree and clearly suggests independent acquisition in at least two lineages.

Same for pT 13-1 which can only date from 1900 if there was an extraordinary level of loss across the lineage. Given these are conjugative plasmids carrying an adaptive novel trait that increases the likelihood that the transfers will be sampled, multiple transfer events do seem more parsimonious that one single ancestral acquisition.

We would like to thank the reviewer for this comment. We agree that a shared evolution event in which there are nested horizontal and vertical transmission of the plasmids is plausible and would reflect a correlation significantly different from zero. We would like to remark that this correlation analysis was only conducted for the association between ST95 and pT 10-1. In this case, this correlation test was done as an additional layer to support the hypothesis that the plasmid could have been introduced by the MRCA of ST95 as strongly suggested by the presence/absence of the plasmid in the phylogeny and the CaveDive analysis. It is true that a correlation significantly different from zero cannot rule out the possibility of a nested evolution event but still makes plausible the hypothesis that pT 10-1 was introduced by ST95 MRCA. However, we agree with the reviewer in the interpretation of this test and we have now indicated this in the manuscript in lines 238-240.
The examples highlighted above by the reviewer all correspond to the association between pTs and particular clades or subclades of the different STs. In this case, the age given corresponds to the acquisition date of the plasmids by those particular (sub)clades. We agree with the reviewer that there are multiple acquisition events taking place in each of the phylogenies. However, we only highlighted in the text and in Table S3 single acquisition events (from the multiple ones plausible given the phylogeny) for which we had strong evidence and were coupled with a CaveDive expansion indicating a possible clonal expansion associated to the acquisition of the plasmids. This is true for all the plasmid and clade/sub-clade associations except for ST95 and pT 10-1 whose results and evidence were explained above. Some detailed information of the highlighted examples below.

- pT 2-3 in D was highlighted in the text as the association between pT 2-3 and a particular subclade of the ST131 B lineage. Here, the given plasmid acquisition date (95% CI, 1946-1986) (Table S3) corresponds to when we estimated that the pT was introduced (one of the possible introductions) in this subclade of ST131 (internal nodes 232-236) .
- pT 13-1 was highlighted in the text as the association between pT 13-1 and a particular subclade of ST73 (fimH10:O6:H1) (internal nodes 311-312 in the phylogeny, Table S3). We agree with the reviewer that based on the phylogeny presented in Figure 3C, there are other acquisition events of the plasmid in other lineages but we did not date those acquisitions.

Sampling affects the conclusions. The analyses were done based on clinical samples. This is a typical problem that leads to certain biases that may affect some analyses. For example, when the authors state that core plasmid functions diversify more than regions involved in virulence or resistance, it should be considered that the latter may have recently integrated different plasmid backbones and therefore that coalescent times are much shorter for the latter than the former. Similarly, when analyzing the evolution of plasmids one must consider they are spreading horizontally and transferring the trait that will lead to isolation of the clones. As such, its not surprising that certain plasmids are much more frequent than others (it's a consequence of sampling for bacteria with plasmid-encoded traits).

We thank the reviewer for their comment. We agree with the observation made regarding that some plasmids are much more frequent than others due to their plasmid-encoded traits. We believe this is an important and remarkable observation and we have now stated this in the Discussion in lines (471-474).

Supplementary material. The manuscript mentions the existence of 4 supplementary tables, but none was present in the site for download.

We thank the reviewer for highlighting this. All the tables are now uploaded with the revision.

Other comments:
Line 38. " Our analysis suggests that novel successful plasmid backbones emerge rarely " This needs qualifications. Is it that unexpected considering the period of acquisition of plasmids spanning decades? One can recognize plasmids today that were present in the collections of enterobacteria in the first half of the 20th century. If the current understanding is that plasmid backbones constantly emerge in short periods, then the relevant citations should be given.

We thank the reviewer for highlighting this. We have removed this statement in the abstract since we agree that this requires some detailed quantification analyses which are out of the scope of the current manuscript.

I have some concerns regarding the phylogeny in Figure 1A. The scale of the tree indicates that the most distant groups of strains are over 0.2 subs/position away which seems incompatible with the idea that these are all E. coli strains. Maybe the scale is in other units? This should be indicated.

We thank the reviewer for spotting out this problem. Figure 1A is a core-SNP species wide tree of ~1,000,000 variable sites when mapped to EC958. Previously the scale was 0.02 SNPs per SNP-site which is not meaningful, thank you for pointing this out. We have now corrected it to 0.0004 substitutions per site as the genome size is approximately 5MB. All the Figures using the same core-SNP species wide tree have been corrected and the legend now shows the correct unit. However, we would like to point out as this is core-SNPs, like with a core genome tree, that it contains substitutions introduced via both mutation and recombination. In Figure 1A we try to illustrate both the complete sampling of major clones and representatives of the other lineages across the whole species found in BSIs for which a recombination free phylogeny is not computationally feasible to calculate within a reasonable time frame using existing methods. Such a phylogeny is further not necessary for the current scope.

We have now detailed this in the Methods section (lines 527-531).

It's difficult to understand the information in Figure 1B.

In line with another suggestion from Reviewer #2, we have included in the Inc information in Figures 1B and 1C which will help the readers to contextualize each of the pTs shown.

Line 200. Why is it concluded that small plasmids are not a source of adaptive traits? Just because they only share replication and mobilisation machinery? They may provide adaptive functions that are more diverse (which is the case made by the last section of the manuscript where bacteriocins are in small plasmids). I don't understand how the containment analysis leads to such conclusion.

We thank the reviewer for pointing this out. We agree with the reviewer that the containment analysis cannot support the statement and we have decided to remove it from the manuscript (line 198 in the revised document).

Line 208. I don't understand the argument that the transfer region diversified more than the others. To know they diversified more there must be a phylogenetic frame showing that other regions were co-inherited vertically and evolved slower, but the authors sustain (correctly in my opinion) that homologous functions were acquired independently. Such a statement requires quantification.

We thank the reviewer for highlighting this. In this section, we just wanted to show that the transfer region among pTs 8-2/9-1/10- and pTs 2-1/2-2/2-3 was highly conserved. This suggested that these pTs could have evolved from the same ancestral pT, based on this we further did some plasmid phylogenies to confirm this.

Accordingly, we have rephrased this section (lines 205-208).

Figure S7. Scale unclear. Is it the standard substitutions/position (seems too high)? Maybe something different like the total number of differences?

We thank the reviewer for pointing this out. This phylogenetic tree was computed by Gubbins using Snippy output, the scale is in substitutions per site multiplied by the number of recombination-filtered SNPs.

We have indicated now in the Methods section (lines 640-643) and in the legend of Figure S7.

Line 377. "Our analysis indicated that larger plasmids in E. coli are almost always associated with traits falling under bacterial competition, pathogenesis and antibiotic resistance. " This statement should be moderated. I assume many genes have unknown functions (if not this should be stated).

We agree with the reviewer and we have rephrased the sentence accordingly (lines 374-376).

Reviewer #2 (Remarks to the Author):

In this study, Sergio Arredondo-Alonso and collaborators investigate the genetic basis underlying plasmid success in extra-intestinal pathogenic E. coli. They first obtained high quality closed genomes from 2,000 E. coli isolates recovered over two decades (form the Norwegian epidemiological longitudinal collection, NORM collection). Next, they elucidated and typed the ExPEC plasmidome in detail, and investigated if particular pTs contributed to the success of pandemic ExPEC clones. The authors found strong plasmid-lineage associations, with some bacteriocin-

producing plasmids playing a potential key role in the success of specific clones. In general, I think that this is an extremely interesting work shedding new light in the role of plasmids in ExPEC evolution. I only have some minor comments for the authors:

We thank the reviewer for his positive and kind words regarding our manuscript.

Figure 1, Would it be possible to indicate in the figure (maybe in the upper labels in the plots in panel C), to which Inc / PTU group belongs each pT? that would help the reader to make sense of the type of plasmids in each pT. For example, the label could be “pT2-1 (IncF..)”

We thank the reviewer for this suggestion, We agree on it, and we have now included the IncF information available in Table S2 in Figure 1 (for both panels B and C). For the pTs corresponding to multireplicon sequences (Table S2), we have indicated the major Inc present in the group.

Figure 2 and S6, Include a colour legend for gene function?

We thank the reviewer for this suggestion. In agreement with Reviewer #3, we have added a legend with the most important operons shared across distinct pTs to facilitate the interpretation of the figure. In this case, the colour does not correspond to gene function but gene cluster homology. This is generated by default by Clinker and permits to easily scan which regions are shared across pTs.

Figure S3, It would be interesting to also show the results of “plasmid load” by correcting the plasmidome size in bp by the plasmid copy number (average coverage plasmid vs average coverage chromosome), which would give a better idea of the total number of bp devoted to plasmids.

We thank the reviewer for this excellent suggestion. We have generated the results of the plasmid load by multiplying the length of the plasmid sequences by their depth as estimated by Unicycler. In this case, Unicycler computes a depth which normalizes the depth of contigs in the graph to the median value. Following this, most of the contigs in the assembly graph correspond to the chromosome and consequently chromosome-derived contigs tend to have a depth near 1x and plasmid-derived contigs can have a higher copy number. We have now added this information in the Methods section (lines 733-736)

We have now included this information in a new panel (E) in Figure S3. Overall, the median plasmid size without considering the plasmid load was 122.3 kbp (mean 125.6 kbp). However, after the inclusion of the copy number to estimate how many bp are devoted to plasmids, this median number increases to 182.9 kbp (mean 262.3 kbp).

-line 221, what is the reason to exclude plasmids smaller than 10 kb for this analysis?

We thank the reviewer for pointing this out. We focused on plasmids higher than 10 kbp to streamline our analysis focusing only on plasmids with the potential to carry one or multiple operons involving iron-acquisition, immune evasion, antimicrobial resistance and clone competition (bacteriocins) among other functions.

Figure S9 panel B is partially deleted.

We thank the reviewer for spotting this issue. We have resized the figure to fit into the manuscript.

-Line 321 Fig. 45A -> Fig. 5A

We thank the reviewer for spotting this typo. We have corrected this in the text.

-Section “Plasmid-encoded bacteriocins play a pivotal role in bacterial competition”. I may be mistaken, but if I understood correctly, the authors tested the microcin/colicin activity of isolates carrying microcin/colicin-producing plasmids, using as negative controls other isolates that do not carry the microcin/colicin-producing plasmids (or plasmids with deletions of one of the systems, according to figure 5?). It would be nice to have isogenic clones with or without the microcin/colicin-producing plasmids to definitely show that the inhibition is mediated by the plasmid-encoded colicins. The plasmid can be cured using, for example, CRISPR-based methods. Alternatively, the plasmid could be conjugated to a strain which does not carry the colicin-producing plasmid, and then supernatants of the plasmid-carrying/plasmid-free isogenic clones could be tested in parallel. Alternatively, the microcin/colicin clusters could be cloned in a vector. The control right now is using other isolates not harboring a microcin/colicin-producing plasmid, but it is not clear how different the genomes of both isolates are. Do they differ in the composition of colicins in the chromosome or other plasmids? The results are pretty convincing in general, but it would be nice to have this isogenic control, if possible. Alternatively, the authors could explain a bit better why their control strains are adequate, because I think that it’s not completely clear in the results section.

We understand the Reviewer's perspective, and clarify the strain selection. When selecting the isolates for experimental work we aimed to minimize potential confounding factors (e.g., presence of additional bacteriocins), and include strains that could serve as controls as well. A full characterization (accessory genes and plasmid types - pTs) is provided in table S1, but to facilitate comprehension (without the need to cross-check tables) we now edited Table S4 for clarity. In this edited version of Table S4, we indicate more clearly which bacteriocins the isolates encode (and their location: plasmid vs chromosome), and specify the "version" (i.e., pT) of pColV plasmid they carry.
Our objective was to investigate whether pColV plasmids had a role in mediating bacterial competition. Thus, we tested three ST95 isolates encoding microcin V and colicin Ia (in the related pT9-1 or pT10-1 plasmids), but no other bacteriocins. We verified that these isolates were able to inhibit the growth of the lab strain *E. coli* MG1655.
We further tested two ST131-B isolates, one that does not encode any bacteriocins, and another that carries the pT9-1 pColV (but does not encode any other bacteriocin). These work as negative and positive controls (in a nearly similar way the reviewer suggests with isogenic clones). As expected, inhibition of *E. coli* MG1655 was only observed with the isolate carrying the pT9-1 pColV plasmid, but not with the other.
It could be argued that the isolates tested could encode other inhibitory factors that we failed to predict when conducting our genomic analyses. Microcin V and colicin Ia are relatively well-studied and their expression is known to be induced through iron limitation and SOS response, respectively. Our results with the pColV carrying isolate 27-61 (ST95) demonstrate that the inhibitory effect is not observed in the absence of the specific induction. The inhibition assays were conducted in iron-limiting conditions, to stimulate the expression of Cir (the common receptor for

microcin V and colicin Ia). During pilot experiments to optimize the conditions for these assays we observed a lack of (or at best, very faint) inhibitory effect during the assay if the indicator strain was not growing in iron limitation conditions (data not shown). We argue that the probability of the pColV-carrying isolates to encode undetected inhibitory factors that behave exactly (e.g., same induction conditions and required cell receptor) like the predicted bacteriocins would be quite rare. Especially considering the conserved patterns observed with both ST95 and ST131-B isolates. Since pColV encodes two bacteriocins simultaneously, we included (as further controls) strains encoding only one of these two bacteriocins. Again, their inhibitory patterns were as expected. Moreover, we crucially show with the inhibition pictures in Fig 5 inhibition to be due to a diffusing substance and not to bacteriophages. Altogether, we believe that we provide compelling evidence that the inhibitory effects observed are mediated by the pColV-encoded bacteriocins, and not by other confounding/unpredicted factors.

Alvaro San Millan

We thank the reviewer for positive remarks and very useful suggestions which have been taken carefully into account when preparing the revision.

Reviewer #3 (Remarks to the Author):

The authors present a well written manuscript describing the plasmidome of bloodstream infection-associated E. coli from a Norwegian longitudinal epidemiological study. In-depth genomic analysis identifies successful plasmid-backbones within the ExPEC landscape and dates the acquisition of relevant plasmids in the major ExPEC sequence types ST69, ST73, ST95, and ST131.
This is then complemented by experimental work showcasing the functionality of two colicins carried by plasmids linked to parallel clonal expansion. This result is linked to the maintenance of susceptible lineages in the ExPEC landscape due their association with the colicins and also the efficacy of the colicins against multi-drug resistant lineages.

I recommend this manuscript for acceptance and suggest small changes to help with the accessibility of figures.

We thank the reviewer for their positive feedback and enthusiasm regarding our manuscript.

On line 205 it would be useful to state the relevant colours in Figure 2 that pertain to the operons described as they are difficult to locate in the figure due to small text: “iroBCDEN (green), iutA-iucACBD and sitABCD (turquoise)”. Alternatively, these text labels could be highlighted in Figure 2 in some way.

We thank the reviewer for highlighting this. We have followed the suggestion and included a colour legend with the most important operons shared across the pTs and with a larger font size. We have done this for both Figures 2 and S6.

It would also aid the reader to include a ST (including ST131-A, B, C) annotation column in Figure 4B, as these results are described exclusively in the context of specific STs in the manuscript (lines 320-322).

We thank the reviewer for this suggestion. We have now included the main STs in Figure 4B.

Minor comments:
On lines 253, 278, 285, and 291, 0 is used instead of O when referring to O-types.

Line 296: missing '%'.
Line 303: missing space between % and CI.
Line 321: Is this referencing Fig 4A or 5A or both?
Line 346: include reference for this statement.
Line 367: full stop missing at end of sentence.
Line 380: use the symbol for beta for continuity with the rest of the manuscript.
Line 423-425: is there a suitable reference for this?
For rigour and reproducibility purposes, references should be added in the methods section for Guppy on line 515 and mlplasmid on line 529 (publication or GitHub url as appropriate).
Line 778: full stop missing at end of sentence.
Line 800: comma missing between 'chromosome' and 'plasmid'.

We thank the reviewer for pointing out the above issues in the text. We have now corrected all of them in the manuscript.

- In Line 321, we aimed to reference Figure 4A.
- In line 346, we have added the following two references with PMIDs 21982038 and 35000591.
- For Lines 423-245, we have added the following two references with PMIDs 27390780 and 35076267

We would like to thank the reviewer again for spotting those errors and for their positive words on the manuscript.

Point-by-point response

Reviewer #1 (Remarks to the Author):

The authors took my comments in account and have solved all the key issues, including the lack of supplementary material tables, the scales of phylogenetic trees, and some statements that were not very well supported (in my opinion). I thank them for that and have very few remaining comments. This is an interesting and valuable study.
- Regarding the exclusion of the <10 kb plasmids, one should take into account that small plasmids do carry resistance genes and colicins. (This is actually the case of most plasmids with colicin A)

Response: We thank the reviewer for positive comments and note that all small plasmids (<10kb) with colicin genes have been included in the analysis presented in Fig 4. The size restriction was used only in the first part of the Results to present results related to a number of well-known plasmids from the literature.

- some references lack page numbers of manuscript ids (45, 49, ...)

Response: These are articles in electronic journals typically lacking page numbers. However, we have added page numbers where it was appropriate.