Overview:
Research Article. Cross-sectional (technically patients were seen at multiple time points which would suggest a longitudinal study, but only one culture was obtained and it was only compared to other cultures, so I think a cross-section better describes this study.)

Abstract:
The abstract accurately summarizes the article, with the exception that there is a vague reference to an outbreak investigation that was an undisclosed objective until one reaches the discussion section of the paper.

Intro:
I found the introduction to be properly and sufficiently sourced (15 references). The authors did an excellent job of explaining the prevalence and severity of PA infections in this population. They moved through the literature with increasing specificity, ultimately ending with a justification for determining the predominant mechanism of transmission within this population. The two goals seem to be 1) what is the main mode of infection (E-to-P, or P-to-P) and 2) is this evidenced by the data collected from genotyping/phenotyping the samples.

The second to last paragraph was a bit confusing. (Page 2, ¶2) “Few of these P aeruginosa phenotypes... have been associated with the more severe lung infection.” How does this support the content of the paragraph that begins with “the pathogenicity of P aeruginosa in CF is promoted by the diversification of the bacterial population and the presence of multiple phenotypes.”? It’s certainly interesting, but since this paper is not looking specifically at severe lung infections, it is unclear how this is relevant here. Additionally, if the point of the paper is to uncover “the link between specific P aeruginosa phenotypic traits and genotypic prevalence” doesn’t a statement that suggests this has already been done in severe lung infections undermine the value of this research?

The final ¶ sums up the research question, and suggests “possible transmission may have occurred” but overall the results are ambiguous.

Methods:
Why is there such a large gap between when the data was collected (concluded in April 2009) and this paper (~2016)? Is this a typical duration for this number of samples / use of PFGE?

A more quantitative description of what constitutes “intermittently or chronically colonized” (regarding the patient population) would be desirable.

The choice to “blindly choose” an isolate sounds desirable, but a description of why you think it is would be even better. “blindly chosen” so as to prevent...what? Researchers who really enjoy green isolates instead of pink from biasing the study?

The methods section also lacks any description of when demographic / clinical characteristics of the patients were obtained. If a patient has been there for a number of years, was FEV obtained when
cultures were obtained or is this done at the first visit years ago? It seems unlikely that all of these patients would have similar lung infection severities which could influence colony composition.

**Results:**

It is odd that the first acknowledgment of the methodology for collecting FEV and description of what constitutes “chronically colonized” appear in the results section. This partially addresses the questions I had above. There is, however, still no explanation for which cultures (from chronically or intermittently colonized) were used in the analysis: the first ones, the last ones, etc.

There is a vague mention of a “contacts investigation” on Page 3.

(phenotype characterization) The most prevalent type that was shared by the most subjects (P14) was also lacking in almost all of the phenotypes that are suggested to cause pathogenicity... And P6 which was found in half as many subs as P14 had many of the phenotype characteristics under investigation.

**Discussion**

(P4, ¶ 1) “the objective was to establish the presence of a high-risk outbreak...” I thought the point of this paper was to identify the phenotypes associated with p-to-p transmission? I didn’t know the authors assumed there was an active outbreak occurring at this facility. I would revise that statement or eliminate it entirely. It seems to suggest they failed to find the thing they didn’t acknowledge or justify they were investigating.

(P4, ¶ 2) “It has been reported that different management...[of patients treatment]... is probably responsible for the discordant epidemiology”. Why wouldn’t you control for the different treatment regimens as potential confounders in your analysis? If you had access to their personal information via the clinic, I imagine you have access to this info as well.

(P4, ¶ 2) “the structure and organization of this centre changed in 2012...” How would changes that occurred 3 years after you collected data affect your results? If they do, the authors haven’t properly explained the longitudinal components of this study.

(P6, ¶1) “p-to-p transmission cannot be due to a specific... phenotypic trait analysed in this study.” I think this is extremely over-reaching. Just because there was heterogeneity between all samples doesn’t mean that one specific trait is not responsible since you failed to demonstrate p-to-p transmission within this study to begin with. At best you’ve indicated p-to-p transmission cannot be ruled out, not that it definitively occurred and in which samples.

I’m glad you mentioned co-infections and the single sampling as limitations.

**Overall recommendation:** Minor Revision (needs to be re-reviewed after suggested changes are made)

**Future research / Big picture question:** Some phenotypic variation (like motility, twitching, etc.) are adaptations designed to induce infections, as the authors acknowledge. Obviously the bacteria no longer needs to move once it has colonized. Why in attempting to identify these phenotypes would you actively recruit patients who had intermittent/chronic infections? Isn’t the fact that you saw huge variation in phenotypes / lack of phenotypic presentation confounded by the patient’s late stage of infection?