A high-throughput approach to the culture-based estimation of plasmid transfer rates

David Kneis [a], Teppo Hiltunen [b,c], Stefanie Heß [b*]

[a] Institute of Hydrobiology, TU Dresden, Germany

[b] Department of Microbiology, University of Helsinki, Finland

[c] Department of Biology, University of Turku, Finland

* Corresponding author: stefanie.hess@helsinki.fi

Abstract

Horizontal gene transfer is an essential component of bacterial evolution. Quantitative information on transfer rates is particularly useful to better understand and possibly predict the spread of antimicrobial resistance. A variety of methods has been proposed to estimate the rates of plasmid-mediated gene transfer all of which require substantial labor input or financial resources. A cheap but reliable method with high-throughput capabilities is yet to be developed in order to better capture the variability of plasmid transfer rates, e.g. among strains or in response to environmental cues. We explored a new approach to the culture-based estimation of plasmid transfer rates in liquid media allowing for a large number of parallel experiments. It deviates from established approaches in the fact that it exploits data on the absence/presence of transconjugant cells in the wells of a well plate observed over time. Specifically, the binary observations are compared to the probability of transconjugant detection as predicted by a dynamic model. The bulk transfer rate is found as the best-fit value of a designated model parameter. The feasibility of the approach is demonstrated on mating experiments where the RP4 plasmid is transfered from *Serratia marcescens* to several *Escherichia coli* recipients. The method's uncertainty is explored via split sampling and virtual experiments.

Keywords: horizontal gene transfer, conjugation, parameter estimation, *Escherichia coli*, *Serratia marcescens*, RP4 plasmid

1 Introduction

Horizontal gene transfer (HGT) is an essential component of bacterial evolution (Ochman et al., 2000). Owing to the fast and global spread of antimicrobial resistance (World Health Organization, 2014), a better quantitative understanding of HGT is needed more than ever. Among the different pathways of HGT, the exchange of plasmids through conjugation is known to be of particular relevance (Bennett, 2008; Wellington et al., 2013; Cabezón et al., 2015). In this process, the plasmid is transfered from cells of a donor strain (D) to cells of a recipient strain (R). Recipients that have acquired the plasmid are classified as transconjugants (T) to distinguish them from the original donors. Conjugation allows genes to be transferred across the boundaries of taxonomic groups (Davies, 1994).

Mating experiments provide a means to study the efficiency of plasmid transfer between donors and recipients under well-defined conditions. Unfortunately, there is no generally accepted measure of the transfer efficiency. Many authors chose to report ratios of cell densities like T/R, T/D or T/RD (Dahlberg et al., 1998; Pinedo and Smets, 2005; Toomey et al., 2009) but such numbers are difficult to compare across studies (Sørensen et al., 2005; Zhong et al., 2012). The plasmid transfer rate constant of a mathematical model, on the other hand, is well defined. A model-based definition of the plasmid transfer efficiency in liquid cultures was made popular by Simonsen et al. (1990) while the underlying differential equations date back to Levin et al. (1979).

In any case, the quantification of plasmid transfer rates requires data on the abundances of the T, D, and R strain. The available methods to separately quantify D, R, and T in the mixed-strain mating cultures include plating on selective media (e.g. Gordon, 1992; Fox et al., 2008), measurement of fluorescence (Christensen et al., 1996; Normander et al., 1998; Pinilla-Redondo et al., 2018), and qPCR (Wan et al., 2011). Any of the approaches is associated with considerable labor input and/or costs if plasmid transfer rates are to be estimated for a larger set of mating pairs and/or variable experimental conditions.

In this study, we demonstrate the feasibility of a low-cost, cultivation-based approach to the estimation of bulk plasmid transfer rates for liquid cultures. In contrast to existing methods it exploits binary information on absence and presence of transconjugants in the wells of a well plate. The rate constant of plasmid transfer is found via inverse modeling using a dynamic numerical solution of the Levin et al. (1979) equations. Specifically, the mismatch between model-simulated probabilities of transconjugant detection and the corresponding binary observations is minimized. To our knowledge, the proposed approach is novel even though the concept of binary observations has been used earlier (Johnsen and Kroer, 2007) and high-throughput setups to study conjugation were developed in the past (see, e.g., Lorenzo-

Diaz and Espinosa, 2009).

We evaluated the method on experimental data for the mating pair *Serratia marcescens* (donor) and *Escherichia coli* (recipient) exchanging the RP4 plasmid. For the purpose of thorough testing, mating experiments were run for 16 recipient strains originating from different environments.

2 Material and methods

2.1 Mathematical approach

Measure of plasmid transfer efficiency

We aim at quantifying the efficiency of plasmid transfer in a well-mixed culture of plasmid donors (D), recipients (R), and transconjugants (T). A common approach to this problem is to fit a mechanistic dynamic model to observed cell counts of D, R, and T. The bulk plasmid transfer efficiency is then represented by a designated parameter of the model: the constant γ.

Levin et al. (1979) were the first to propose a suitable model accounting for the simultaneous occurrence of horizontal and vertical gene transfer in a liquid culture. It comprises one differential equation for each of the strains D, R, and T. The model is most conveniently expressed in matrix notation (Eqn. 1).

$$\begin{bmatrix} \dot{D} & \dot{R} & \dot{T} \end{bmatrix} = \underbrace{\begin{bmatrix} \beta_D \cdot D \\ \beta_R \cdot R \\ \beta_T \cdot T \\ \gamma \cdot R \cdot (D+T) \end{bmatrix}^T}_{\text{Process rates vector}} \cdot \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}}_{\text{Stoichiometry matrix}} \tag{1}$$

The left hand side of Eqn. 1 represents the derivatives of the individual strains' cell densities with respect to time (cells mL^{-1} hour^{-1}). The process rates vector holds the expressions for growth (rows 1–3) as well as the expression for plasmid transfer involving the constant γ (mL cells^{-1} hour^{-1}). The remaining parameters β_D, β_R, β_T specify the strains' effective growth rate constants (hour^{-1}). Typically, the latter are derived from the strains' intrinsic growth rate constants (μ_D, μ_R, μ_T) via a dimensionless multiplier accounting for resource limitation (e.g. a Monod-type model or a carrying capacity term).

Common approach to the estimation of γ

The common approach to the estimation of γ relies on observed cell densities for the three involved bacterial strains. In its simplest form, the value of γ is estimated from single, synchronous observations of D, R, and T employing an analytical solution of Eqn. 1 (Simonsen et al., 1990).

More generally, Eqn. 1 can be integrated numerically for given initial values and parameters to obtain predictions of D, R, and T at any desired point in time. The value of γ is then found by minimizing the mismatch between predicted and observed cell densities using an optimization scheme. The numerical approach is attractive for its versatility: It can exploit multiple observations over time (including asynchronous ones) and it easily copes with extended model formulations that lack analytic solvability (see, e.g., Zhong et al., 2010).

Novel approach: Inference of γ from binary data

We hypothesize that reliable estimates of γ could also be obtained from a reduced set of information. In contrast to the common approach outlined above, we propose to infer the value of γ from the presence (or absence) of transconjugant cells in small sub-samples drawn from a mating culture at multiple time points. Compared to cell counts, such binary information (0: absence, 1: presence) is much easier to obtain in practice (see Sec. 2.2).

The idea is based on the notion that the probability of detecting a colony forming unit in a homogeneous sample is a function of the cell density. Namely, the probability p_T of seeing at least one transconjugant CFU in a volume v (mL) is related to the abundance of T (cells mL^{-1}) through Eqn. 2

$$p_T = 1 - ppois(k = 0, \lambda = v \cdot T) \tag{2}$$

with *ppois* being the Poisson distribution function. Specifically, the expression $ppois(k = 0, \lambda = v \cdot T)$ represents the probability of not finding any transconjugant cell ($k = 0$) in volume v drawn at cell density T. The probability p_T shows a dependency on T and v as illustrated in Fig. 1.

The best-fit value of γ can be found by minimizing the sum of squared residuals (Eqn. 3), where $\phi_{T,i}$ represents the observation of transconjugants in the i-th time step (absence: 0, presence: 1) and $p_{T,i}$ is the corresponding model prediction based on Eqn. 1 and 2. The predicted dynamics of p_T generally follows a sigmoidal pattern (Fig. 2).

$$SSR = \sum_{i=1}^{k} \left(\phi_{T,i} - p_{T,i}(\gamma)\right)^2 \tag{3}$$

Compared to the common approach outlined in the previous paragraph, the proposed method relies on observations of just one of the three state variables from Eqn. 1 (transconjugants). We hypothesize that the reduced information still allows for the identification of parameter γ. It is quite obvious, however, that it is impossible to simultaneously infer the values of further parameters appearing in Eqn. 1. It is thus necessary to determine

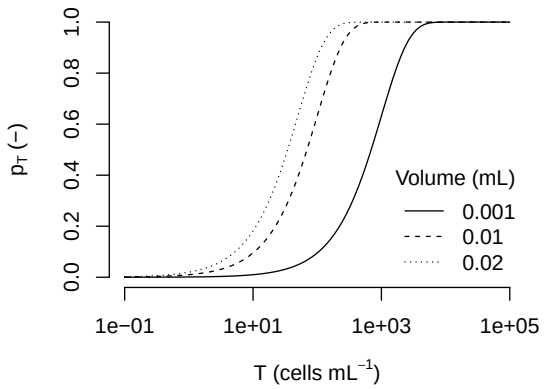

Figure 1: Probability of transconjugant detection (Eqn. 2) as a function of transconjugant density (T) and sampled volume (v).

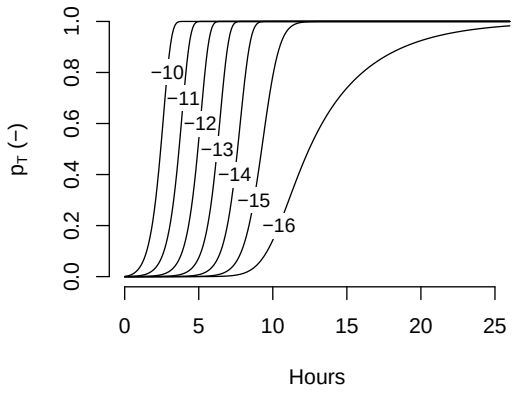

Figure 2: Probability of transconjugant detection over time for typical growth parameters, a sampled volume of 10 μL and various values of the plasmid transfer rate constant γ. Numbers plotted over curves represent $\log_{10}(\gamma)$.

the individual strains' growth rates in separate experiments (e.g. via optical density measurements).

2.2 Experimental work

Bacterial strains and plasmid

Mating experiments were carried out with *Serratia marcescens* (ATCC 13880) as the donor of the RP4 plasmid and *Eschericha coli* as the recipient. The broad host-range RP4 plasmid confers resistance against three antibiotics from different drug classes (ampicillin, kanamycin, tetracyclin). An acquisition of that particular multi-resistance by spontaneous mutation is very unlikely; thus its occurrence in *E. coli* reliably indicates the formation of transconjugants.

The recipient strains originated from two sources: Eight *E. coli* strains labeled E1 to E8 were isolated from river sediments (Heß et al., 2018a) while another eight strains (S1 to S8) were isolated from sewage (Heß et al., 2018b). RP4 is a member of plasmid incompatibility group IncP1. Prior to the mating experiments, all 16 recipient strains were tested for the presence of RP4-incompatible plasmids using the PCR primers published by Bahl et al. (2009).

Mating experiments

Mating took place in the individual wells of a 8×12 deep well plate covered with a sterile, oxygen-permeable foil. Donor and recipient cells were suspended in HT broth (see Supplement); each well contained 600 μL of the suspension with initial donor and recipient densities of 1×10^5 cells mL^{-1}, respectively. The plates were incubated at 37 °C with continuous shaking at 150 rpm. To ensure identical initial conditions in all replicates, experiments were inoculated from cryo stock cultures of donor and recipients containing 6×10^8 cells mL^{-1} each.

The well plates were sampled at 60 minute intervals starting 3 hours after inoculation. Using a 96 pin replicator (EnzyScreen, Netherlands), subsamples of 10 μL from all 8×12 wells were synchronously transfered to Tryptone Bile X-Glucuronide (TBX) agar plates. The replicator was sterilized between samplings using a combination of ethanol and heat.

The TBX agar was supplemented with 150 μg mL^{-1} of ampicillin, 25 μg mL^{-1} of kanamycin and 20 μg mL^{-1} of tetracyclin to select for RP4 plasmid-bearing cells. The plates were incubated at a high temperature of 44 °C to support growth of transconjugants (*E. coli*; blue colonies) while suppressing growth of donor cells (*Serratia marcescens*; white colonies). After 44–48 hours, the agar plates were inspected visually (see Fig. 3 as an example). Spots with blue colonies indicating presence of transconjugants were registered as positive (code 1), others were marked as negative (code 0).

The number of replicate experiments for the individual mating pairs varied between 12 and 24.

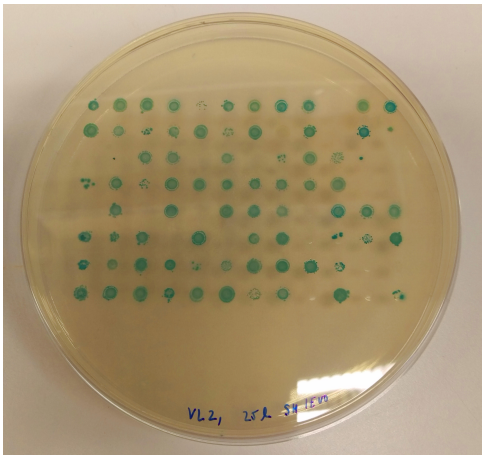

Figure 3: Example of an agar plate with dark spots indicating growth of transconjugants.

Measurement of growth characteristics

The effective growth rates (β_D, β_R, β_T; see Eqn. 1) were determined for all strains from measurements of optical density over 24 h (Tecan plate reader, 600 nm wave length). Growth medium, temperature, and shaking conditions were the same as in the mating experiments outlined above. For each strain (generically denoted X), β was decomposed into three parameters according to Eqn. 4: the growth rate constant μ, a lag time τ and a carrying capacity K.

$$\frac{d}{dt}X = \underbrace{\mu_X \cdot (t > \tau) \cdot \left(1 - \frac{X}{K}\right)}_{\beta \text{ from Eqn. 1}} \cdot X \tag{4}$$

For all strains, μ, τ, and K were determined numerically by fitting Eqn. 4 to the observed optical densities. Raw estimates of K (representing optical densities) were converted to units of cells mL^{-1} by cell count-based calibration curves.

2.3 Employed software

All data processing was done with `R` (R Core Team, 2018). The ODE-based models (Eqn. 1, 4) were implemented using the `rodeo` package (Kneis et al., 2017) relying on `deSolve` (Soetaert et al., 2010) for numerical integration. The Levenberg-Marquard algorithm from package `nls.lm` was used to obtain best-fit values of model parameters.

3 Results

3.1 Growth parameters

As outlined in Section 2.1, knowledge of the growth characteristics of the donor, recipients, and transconjugants (parameters of Eqn. 4) is a prerequisite for estimating the parameter γ. The respective pre-experiments revealed only moderate inter-strain variation of the growth characteristics among recipients (Fig. 4; Table S.1). The growth rate constants, μ, generally range from 0.8 to 1 hour^{-1} (median values) indicating cell division every 42–52 minutes. The carrying capacities K fall in the range from 0.55 to 0.75 (OD_{600}) which is equivalent to 0.82 – 1.02 $\times$ 10^9 cells mL^{-1}, respectively. The fitted lag time τ (hours) was less than 1/2 hour for most of the recipient strains.

The transconjugant strains exhibited a slight increase in the growth rate constants, μ, when compared to the respective recipient strains (Table S.1). Apparently, the acquisition of the RP4 plasmid was not associated with a substantial fitness cost but rather led to a minor stimulation of growth. Such 'negative' fitness costs are the exception instead of the rule (Vogwill and MacLean, 2014). Positive effects of plasmid uptake on the host's fitness were reported, for example, by Enne et al. (2004), Yates et al. (2006), or Wu et al. (2018). The respective mechanisms are currently not well understood but positive epistasis between plasmids provides a possible explanation (San Millan et al., 2014). Whole genome sequencing of selected recipient strains identified plasmids from several incompatibility groups (IncFIB, IncFII, IncI1, IncX3, IncY, Col156, and Col440I) but whether there is a positive interaction with RP4 in the studied *E. coli* strains is currently not known.

The growth rate constant of the *Serratia marcescens* donor strain was considerably lower compared to recipients and transconjugants (μ = 0.68 hour^{-1}). At the same time, the donor exhibited a delayed onset of active growth (τ = 4.3 hours) as well as a reduced carrying capacity of K = 0.32 (0.43 $\times$ 10^9 cells mL^{-1}).

3.2 Plasmid transfer rate constants

The probabilities of transconjugant detection p_T computed from Eqn. 1 and 2 matched the empirical probabilities reasonably well (Fig. 5). The predicted p_T (solid line) mostly falls within the 95% confidence interval of the observations (error bars). However, the example of strain E5 indicates that the fitted model cannot always represent the data with satisfying precision.

It is important to note that error bars are lacking in Fig. 5 when all samples were either negative (empirical $p_T = 0$) or positive (empirical $p_T = 1$). Reasonable bootstrap confidence intervals cannot be computed for those cases. Consequently, slight mismatches between model and observations require a more liberal interpretation if the empirical probabilities are exactly zero or one.

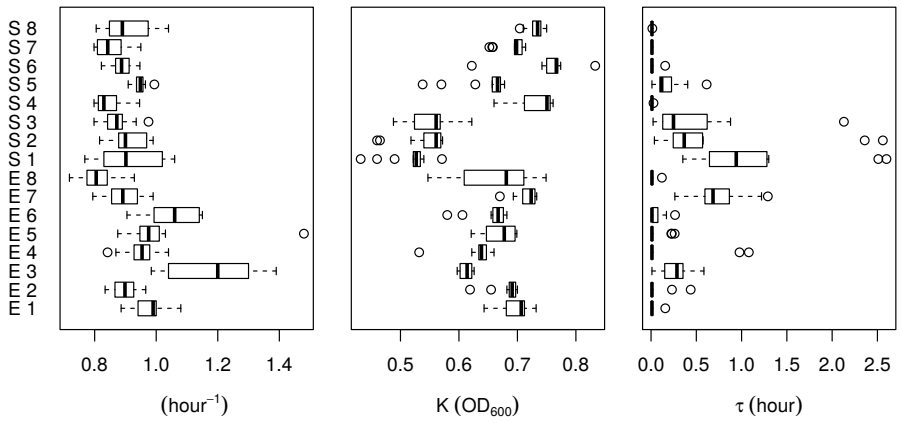

Figure 4: Fitted values of the growth rate constant μ, carrying capacity K, and lag time τ from Eqn. 4 for the sixteen recipient strains E1 to S8.

The best-fit estimates of γ for the 16 mating pairs cover the range from about $10^{-14.5}$ to 10^{-12} mL cells^{-1} hour^{-1} (Fig. 6, Table S.2). Although the values of γ vary over more than two magnitudes, the estimates for the individual strains are fairly robust as suggested by the narrow confidence intervals. This notion is backed-up by the fact that similar values of γ were obtained when either the first or second half of the observation data was ignored (dot symbols in Fig. 6). Overall, the results suggest a substantial variability of γ among the 16 studied recipient strains.

4 Discussion

4.1 Plausibility of the estimates

The conjugation rates for a specific mating pair and plasmid are known to be affected by many biotic and abiotic factors (Fernandez-Astorga et al., 1992; Rysz et al., 2013) including, e.g., resource concentrations, temperature, and turbulence. The obtained estimates of the bulk plasmid transfer rate γ under the given experimental conditions are compatible with previously reported values for mating pairs where *E. coli* acted as a recipient (Simonsen et al., 1990; Zhong et al., 2010). Likewise, the variability of γ among the 16 selected recipient strains (Fig. 6) is in accordance with earlier studies on the inter-strain variability of plasmid transfer rates (see, e.g., Gordon, 1992).

4.2 Inherent uncertainty of the method

Given that the approach outlined in Sect. 2.1 exploits just binary information on a single state variable (presence of transconjugants), the apparent

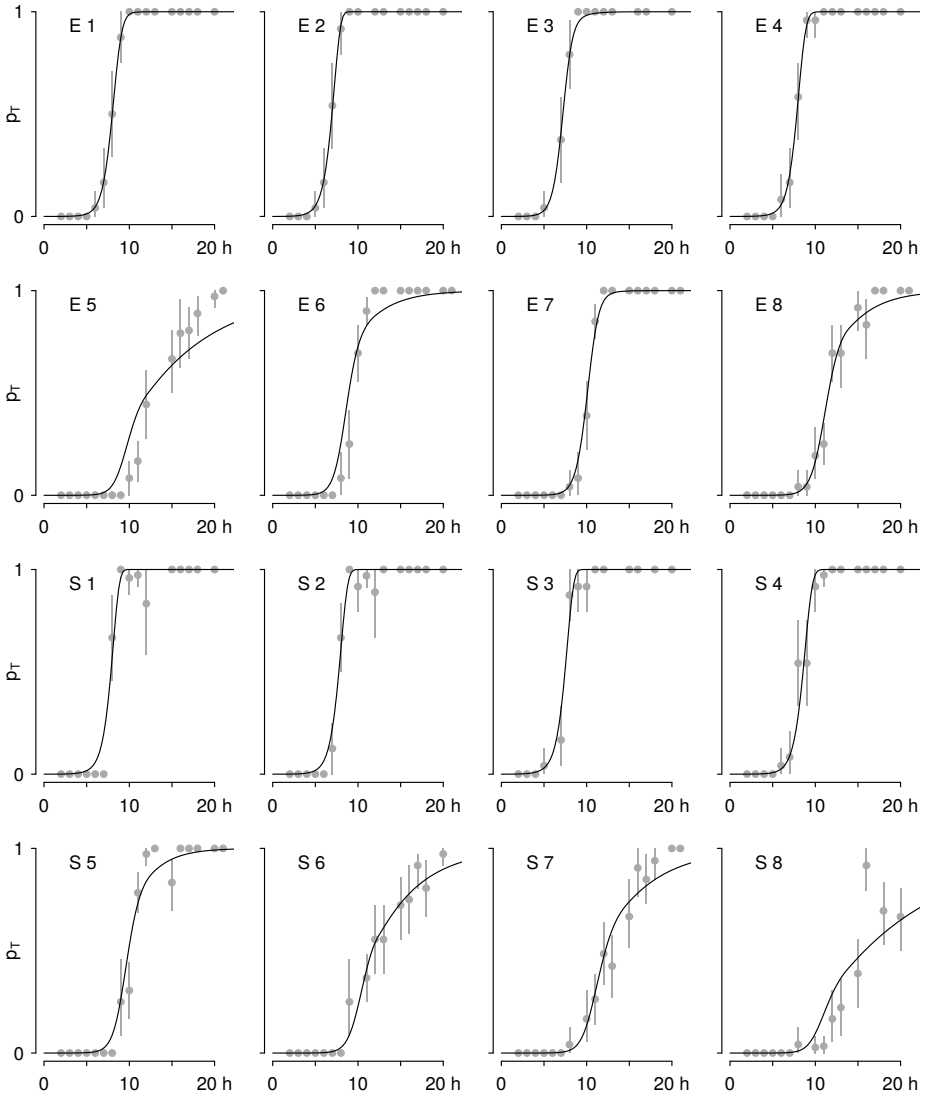

Figure 5: Probability of transconjugant detection over 24 hours. Gray dots indicate the proportion of positive replicate samples (empirical probabilities). The corresponding error bars represent 95% bootstrap confidence intervals. Black curves show predictions by the fitted model (Eqn. 2 in conjunction with Eqn. 1).

uncertainty in the estimates of γ was surprisingly low (Fig. 6, Table S.2). This gave rise to an in-depth exploration of the method's inherent uncertainty by means of virtual experiments (algorithm outlined in Table S.3). As expected, the uncertainty in the estimates of γ is closely linked to the number of replicate experiments (Fig. 7). With four replicates, the estimated

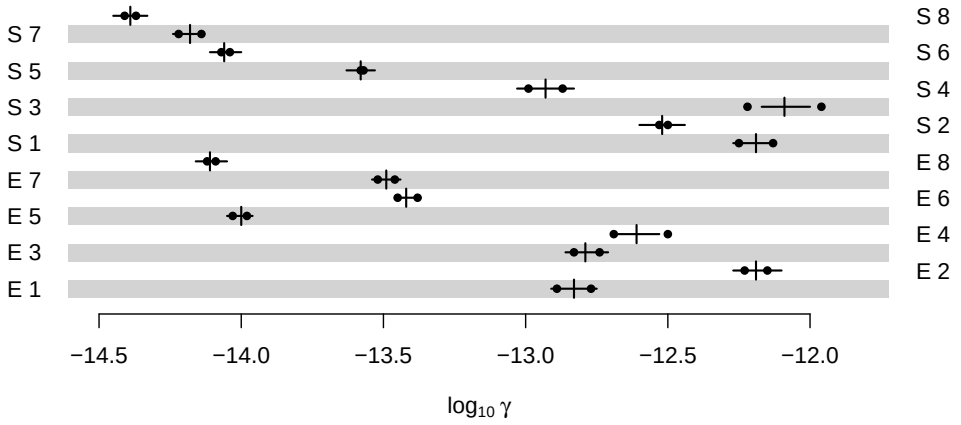

Figure 6: Fitted plasmid transfer rate constants (mL cells^{-1} hour^{-1}; $\log_{10}$) of the 16 recipient strains. Vertical lines represent best-fit estimates based on all experimental data (36 – 60 replicates). Horizontal lines indicate the corresponding 95% confidence intervals derived from Hessian. Estimates of γ being obtained from split samples (50% of replicates) appear as dots. Numbers presented in Table S.2.

γ is accurate up to $\pm 1/2$ log unit with a chance of 95% if samples are taken every hour. If the number of replicates is increased to 12, the error is limited to about $\pm 1/4$ log unit with 95% confidence.

The uncertainty in the estimates of γ also exhibits a dependence on the value of γ itself. Larger values of γ are associated with slightly wider confidence intervals (Fig. 7). This trend was not just observed in data from virtual experiments but it is also visible in Fig. 6 showing the uncertainty of the parameter estimates for the 16 mating pairs. The explanation is essentially provided by Fig. 2. While the curves for high values of γ lie close together (e.g. 10^{-10} and 10^{-11}), the distance between neighboring curves increases as γ becomes smaller. Hence, given a fixed sampling interval, the method's precision improves the lower the plasmid transfer rates are.

The confidence intervals shown in Fig. 7 are essentially symmetric with regard to the dashed line indicating a 1:1 match between estimate and truth. Consequently, in the absence of other sources of error, the parameter estimates are essentially unbiased.

4.3 Estimation of high plasmid transfer rates

Transfer rate constants reported in the literature for various plasmids and bacterial hosts range from about 10^{-8} to 10^{-15} mL cells^{-1} hour^{-1} (Zhong et al., 2010). The rate constants for the studied mating pairs (Fig. 6) are in the

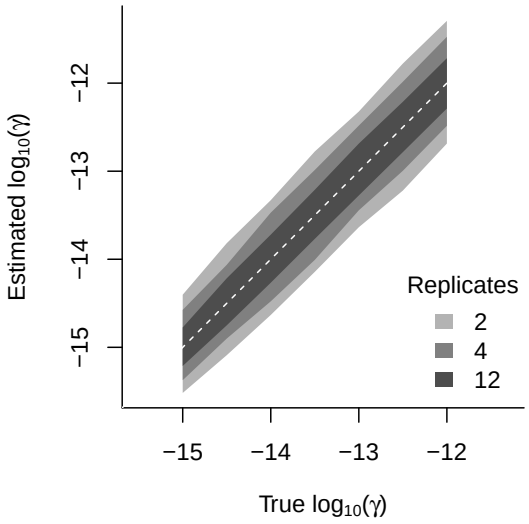

Figure 7: 95% confidence intervals of the estimated plasmid transfer rate constant for typical growth parameters, sampling of 10 μL at hourly intervals, and different numbers of replicates (cf. Table S.3). The dashed line indicates a perfect match between true and estimated parameter values.

middle to lower end of that range. It thus remains an open question whether the proposed method yields reliable parameter estimates even if the bulk transfer rates are particularly high. For large γ, the transition from $p_T = 0$ to $p_T = 1$ occurs in the very early phase of experiments (Fig. 2) and adequate sampling might become difficult.

In order to test the method's applicability to cases where γ is high (e.g. 10^{-8}), virtual experiments were used. Specifically, the algorithm from Table S.3 was run until step 2 after choosing a large γ in step 0. The advantage of virtual experiments lies in the fact that the method's performance is easily tested for a whole range of γ and various specific setups (e.g. sampling frequencies).

As expected, the estimation of γ failed if the transition from $p_T = 0$ to $p_T = 1$ occurred very early, namely before the first sampling. The objective function (Eqn. 3) then returns identical results for arbitrarily large values of γ. Generally, the higher the chosen plasmid transfer rate was, the more likely was numerical optimization to fail.

The virtual experiments showed that convergence can always be ensured by sampling at short enough time intervals. However, a high-frequency sampling may not be desired since the experiment is constantly disturbed unless the procedure is fully automatized. Another effective yet more appropriate strategy is to start the mating experiments with lower initial concentrations.

Reducing the inocula of donors and recipients leads to a delayed transition from $p_T = 0$ to $p_T = 1$ which is better observable, e.g. by hourly sampling. Assuming the actual plasmid transfer rate constants of the studied mating pairs to be as high as 10^{-8} mL cells^{-1} hour^{-1}, inocula of 10^5 cells mL^{-1} would have resulted in convergence failure. A simple reduction of the initial donor and recipient concentration to 10^4 cells mL^{-1} each (dilution by factor 10) would have been sufficient to make parameter estimation successful.

Apart from choosing smaller inocula, the transition between $p_T = 0$ and $p_T = 1$ could be delayed by further experimental modifications. For example, one could reduce the sampling volume (see Fig. 1) accepting a more fragile pin replicator and smaller marks on the agar plate (cf. Fig. 3). Alternatively, the experiments could be run at lower temperatures to decelerate both horizontal and vertical gene transfer. However, this might conflict with the experiment's objective. All in all, adjusting the inocula of donors and recipients seems to be the best option to ensure the method's applicability to mating pairs with very high (or low) efficiencies of plasmid transfer.

4.4 Additional sources of uncertainly

The proposed method requires that the growth characteristics (i.e. rate constants, carrying capacities, lag times) observed in cultures of the individual strains are applicable to the mating cultures as well. This assumption is shared with the end-point method of Simonsen et al. (1990). It is nevertheless advisable to validate the transferability of growth parameters between single-strain and mixed-strain cultures. With respect to the mating pairs used in this study, the transferability was confirmed by quantifying donors and recipients in mixed cultures through plating of serial dilutions on selective agar (spot checks). If the growth parameters differ between single- and mixed-strain cultures, the method does not necessarily loose its applicability. However, the differences must be known and the model (Eqn. 1) needs to be extended accordingly.

Furthermore, the method relies on the plating of very small sample volumes obtained with a pin replicator. Although the manufacturer specifies a design value for the sample volume (10 μL in this study), a certain degree of imprecision has to be expected. However, as long as the sampled volumes are normally distributed around the design value, the estimates of γ are not affected. Only systematic errors in the sampled volumes would lead to biased parameter estimates.

By design, the replicator should avoid cross-contamination when lifted and moved over the well plate. Nevertheless, a careful handling that avoids shocks is certainly advisable to minimize the risk of a carry-over of sample material from one well to another. As a check, the agar plates (Fig. 3) should be inspected for spots of transconjugant growth outside of the grid layout. If those spots occur in greater numbers, the risk of cross-contamination of the

wells is high and the cause must be found and corrected. In general, a larger number of replicate experiments should guard against deleterious effects of random cross-contamination.

The quality of the parameter estimates is also affected by the frequency of sampling. If sub-samples of the mating culture are inspected at hourly intervals, the transition between $p_T \approx 0$ and $p_T \approx 1$ is reflected by a few observations only (Fig. 5). The number of observations is particularly limited for mating pairs with high plasmid transfer rates (e.g. $\log_{10}(\gamma) >$ -12.5). Obviously, sampling at a higher frequency could reduce the uncertainty in parameter estimates through a better coverage of the transition phase. On the other hand, every sampling is associated with a disturbance of the experiment due to a short disruption of shaking and possible temperature anomalies. Last but not least, the frequency of sampling is also determined by practical constraints unless the process is fully automatized.

Finally, the strategy used to obtain replicate data needs attention. In order to fulfill the requirement of independence, replicate mating experiments of the same strain should be run on separate well plates whenever possible. Otherwise, abnormalities in the incubation of a particular well plate or problems with the cultivation of transconjugants sampled at a specific time step are likely to cause biased results.

4.5 Limitations and potential

The used model (Eqn. 1) does not account for individual steps of the conjugation process like the maturation of transconjugants and the recovery of donors after a transfer event (Andrup and Andersen, 1999) or the formation of mating pairs (Zhong et al., 2010). Moreover, a possible loss of the plasmid by cells of the D or T strain is not explicitly modeled. The parameter γ is thus an integrative measure of the plasmid transfer efficiency. Technically, the numerical model is easily extended to allow for a more detailed description of the governing biological processes. However, a further decomposition of the currently lumped model requires prior knowledge on any newly introduced parameters. While data on the presence of transconjugants allowed for a reliable identification of γ, multiple unknowns can hardly be estimated using just binary observations of a single state variable.

The major advantage of the presented method lies in its potential to perform and analyze a large number of mating experiments in parallel at comparatively low cost. The approach is particularly useful to explore variabilities in the effective plasmid transfer rates, e.g. among mating pairs or in response to different stressors. The obtained parameter estimates are valuable inputs to parsimonious mechanistic models relying on a lumped representation of gene transfer processes. Such models are urgently required to predict the spread of antibiotic resistance in environmental systems.

Acknowledgements

S. H. received funding from Deutsche Forschungsgemeinschaft [grant number HE8047]. T. H. was supported by the Finnish Academy [Grant number 294666] and by the Helsinki Institute of Life Science (HiLIFE).

Competing interests

The authors declare no competing interests.

References

Andrup, L. and Andersen, K. (1999). A comparison of the kinetics of plasmid transfer in the conjugation systems encoded by the F plasmid from Escherichia coli and plasmid pCF10 from Enterococcus faecalis. *Microbiology*, 145(8):2001–2009.

Bahl, M. I., Burmølle, M., Meisner, A., Hansen, L. H., and Sørensen, S. J. (2009). All IncP-1 plasmid subgroups, including the novel ϵ subgroup, are prevalent in the influent of a danish wastewater treatment plant. *Plasmid*, 62(2):134–139.

Bennett, P. M. (2008). Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *British Journal of Pharmacology*, 153:S347–S357.

Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F., and Arechaga, I. (2015). Towards an integrated model of bacterial conjugation. *FEMS Microbiol. Rev.*, 39(1):81–95.

Christensen, B. B., Sternberg, C., and Molin, S. (1996). Bacterial plasmid conjugation on semi-solid surfaces monitored with the green fluorescent protein (gfp) from aequorea victoria as a marker. *Gene*, 173 (1 Spec No):59–65.

Dahlberg, C., Bergström, M., and Hermansson, M. (1998). In situ detection of high levels of horizontal plasmid transfer in marine bacterial communities. *Appl Environ Microbiol*, 64(7):2670–2675.

Davies, J. (1994). Inactivation of antibiotics and the dissemination of resistance genes. *Science*, 264(5157):375–382.

Enne, V. I., Bennett, P. M., Livermore, D. M., and Hall, L. M. C. (2004). Enhancement of host fitness by the sul2-coding plasmid p9123 in the absence of selective pressure. *Journal of Antimicrobial Chemotherapy*, 53(6):958–963.

Fernandez-Astorga, A., Muela, A., Cisterna, R., Iriberri, J., and Barcina, I. (1992). Biotic and abiotic factors affecting plasmid transfer in Escherichia coli strains. *Applied and Environmental Microbiology*, 58(1):392–398.

Fox, R. E., Zhong, X., Krone, S. M., and Top, E. M. (2008). Spatial structure and nutrients promote invasion of IncP-1 plasmids in bacterial populations. *ISME*, 2:1024–1039.

Gordon, D. M. (1992). Rate of plasmid transfer among escherichia coli strains isolated from natural populations. *Journal of General Microbiology*, 138:17–21.

Heß, S., Berendonk, T. U., and Kneis, D. (2018a). Antibiotic resistant bacteria and resistance genes in the bottom sediment of a small stream and the potential impact of remobilization. *FEMS Microbiology Ecology*, 94:fiy128.

Heß, S., Kneis, D., Österlund, T., Li, B., Kristiansson, E., and Berendonk, T. U. (2018b). Sewage from airplanes exhibits high abundance and diversity of antibiotic resistance genes. Submitted to Environmental Science and Technology.

Johnsen, A. R. and Kroer, N. (2007). Effects of stress and other environmental factors on horizontal plasmid transfer assessed by direct quantification of discrete transfer events. *FEMS Microbiology Ecology*, 59(3):718–728.

Kneis, D., Petzoldt, T., and Berendonk, T. U. (2017). An R-package to boost fitness and life expectancy of environmental models. *Environmental Modelling and Software*, 96:123–127.

Levin, B. R., Stewart, F. M., and Rice, V. A. (1979). The kinetics of conjugative plasmid transmission: Fit of a simple mass action model. *Plasmid*, 2:247–260.

Lorenzo-Diaz, F. and Espinosa, M. (2009). Large-scale filter mating assay for intra- and inter-specific conjugal transfer of the promiscuous plasmid pmv158 in gram-positive bacteria. *Plasmid*, 61(1):65–70.

Normander, B., Christensen, B. B., Molin, S., and Kroer, N. (1998). Effect of bacterial distribution and activity on conjugal gene transfer on the phylloplane of the bush bean (phaseolus vulgaris). *Applied and Environmental Microbiology*, 64(5):1902–1909.

Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial evolution. *Nature*, 405:299–304.

Pinedo, C. A. and Smets, B. F. (2005). Conjugal TOL Transfer from Pseudomonas putida to Pseudomonas aeruginosa: Effects of Restriction Proficiency, Toxicant Exposure, Cell Density Ratios, and Conjugation Detection Method on Observed Transfer Efficiencies. *Appl Environ Microbiol*, 71(1):51–57.

Pinilla-Redondo, R., Cyriaque, V., Jacquiod, S., Sørensen, S. J., and Riber, L. (2018). Monitoring plasmid-mediated horizontal gene transfer in microbiomes: recent advances and future perspectives. *Plasmid*, Electronic publication ahead of print, DOI: 10.1016/j.plasmid.2018.08.002.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rysz, M., Mansfield, W. R., Fortner, J. D., and Alvarez, P. J. J. (2013). Tetracycline resistance gene maintenance under varying bacterial growth rate, substrate and oxygen availability, and tetracycline concentration. *Environmental Science & Technology*, 47(13):6995–7001. PMID: 23383991.

San Millan, A., Heilbron, K., and MacLean, R. C. (2014). Positive epistasis between co-infecting plasmids promotes plasmid survival in bacterial populations. *ISME Journal*, 8(3):601–612.

Simonsen, L., Gordon, D. M., Stewart, F. M., and Levin, B. R. (1990). Estimating the rate of plasmid transfer: an end-point method. *J. Gen. Microbiol.*, 136(11):2319–2325.

Soetaert, K., Petzoldt, T., and Setzer, R. W. (2010). Solving differential equations in R: Package deSolve. *Journal of Statistical Software*, 33(9):1–25.

Sørensen, S. J., Bailey, M., Hansen, L. H., Kroer, N., and Wuertz, S. (2005). Studying plasmid horizontal transfer in situ: A critical review. *Nature Reviews Microbiology*, 3:700–710.

Toomey, N., Monaghan, A., Fanning, S., and Bolton, D. J. (2009). Assessment of antimicrobial resistance transfer between lactic acid bacteria and potential foodborne pathogens using in vitro methods and mating in a food matrix. *Foodborne Pathog Dis*, 6(8):925–933.

Vogwill, T. and MacLean, R. C. (2014). The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evolutionary Applications*, 8(3):284–295.

Wan, Z., Varshavsky, J., Teegala, S., McLawrence, J., and Goddard, N. L. (2011). Measuring the rate of conjugal plasmid transfer in a bacterial population using quantitative PCR. *Biophysical Journal*, 101(1):237–244.

Wellington, E. M., Boxall, A. B., Cross, P., Feil, E. J., Gaze, W. H., Hawkey, P. M., Johnson-Rollings, A. S., Jones, D. L., Lee, N. M., Otten, W., Thomas, C. M., and Williams, A. P. (2013). The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. *Lancet Infect. Dis.*, 13:155–165.

World Health Organization (2014). Antimicrobial resistance: Global report on surveillance 2014. ISBN: 978 92 4 156474 8.

Wu, R., Yi, L.-X., Yu, L.-F., Wang, J., Liu, Y., Chen, X., Lv, L., Yang, J., and Liu, J.-H. (2018). Fitness advantage of mcr-1-bearing inci2 and incx4 plasmids in vitro. *Front. Microbiol.*, 9:Article 331.

Yates, C. M., Shaw, D. J., Roe, A. J., Woolhouse, M. E. J., and Amyes, S. G. B. (2006). Enhancement of bacterial competitive fitness by apramycin resistance plasmids from non-pathogenic escherichia coli. *Biology letters*, 2(3):463–465.

Zhong, X., Droesch, J., Fox, R., Top, E. M., and Krone, S. M. (2012). On the meaning and estimation of plasmid transfer rates for surface-associated and well-mixed bacterial populations. *J Theor Biol*, 294:144–152.

Zhong, X., Król, J. E., Top, E. M., and Krone, S. M. (2010). Accounting for mating pair formation in plasmid population dynamics. *Journal of Theoretical Biology*, 262(4):711–719.

Supplement

Preparation of HT broth for mating experiments

HT broth contained 1 g L^{-1} R2A broth, 11.5 g L^{-1} M9 salts, and 10 mL L^{-1} HayStock I. The latter solution was prepared by autoclaving 25 g of Cereal Grass Medium (Ward's Science, Canada) dissolved in 500 mL of water, followed by centrifugation and filtration of the supernatant (5 μm filter). Aliquotes of HayStock I were kept in the freezer until use.

Fitted growth constants

Table S.1: Fitted growth constants (defined by Eqn. 4) of the *Serratia marcescens* donor strain, the 16 *E. coli* recipients and the corresponding transconjugant strains. The reported values are medians obtained from at least 12 replicate growth curves.

	Without RP4			Harboring RP4		
Strain	μ	K	τ	μ	K	τ
Donor	–	–	–	0.68	0.32	4.3
E 1	0.99	0.71	0.0	1.08	0.65	0.9
E 2	0.90	0.69	0.0	1.14	0.68	1.4
E 3	1.20	0.61	0.3	1.29	0.60	0.8
E 4	0.95	0.64	0.0	1.02	0.63	1.3
E 5	0.98	0.68	0.0	1.13	0.64	0.1
E 6	1.06	0.67	0.0	1.20	0.65	1.9
E 7	0.89	0.72	0.7	0.97	0.68	1.3
E 8	0.81	0.68	0.0	1.04	0.61	0.7
S 1	0.90	0.53	0.9	0.97	0.57	0.4
S 2	0.90	0.56	0.4	1.12	0.56	0.1
S 3	0.87	0.56	0.2	0.93	0.46	1.2
S 4	0.83	0.75	0.0	1.10	0.71	1.0
S 5	0.95	0.67	0.1	1.04	0.63	0.7
S 6	0.89	0.77	0.0	1.06	0.65	1.9
S 7	0.84	0.70	0.0	0.98	0.67	1.8
S 8	0.89	0.73	0.0	0.97	0.70	0.2

Fitted values of γ including uncertainty estimates

Table S.2: Fitted plasmid transfer rate constants (mL cells^{-1} hour^{-1}) of the sixteen recipient strains ($\log_{10}$ transformed). The 95% CI refers to the estimate from the full sample (column 3). It is based on the approximate Hessian at the best-fit solution.

Strain	1st subsample	2nd subsample	Full sample	95% conf. interval
E 1	-12.89	-12.77	-12.83	-12.91 / -12.75
E 2	-12.23	-12.15	-12.19	-12.27 / -12.10
E 3	-12.83	-12.74	-12.79	-12.86 / -12.71
E 4	-12.69	-12.50	-12.61	-12.69 / -12.53
E 5	-14.03	-13.98	-14.00	-14.05 / -13.96
E 6	-13.45	-13.38	-13.42	-13.46 / -13.37
E 7	-13.52	-13.46	-13.49	-13.54 / -13.44
E 8	-14.12	-14.09	-14.11	-14.16 / -14.05
S 1	-12.25	-12.13	-12.19	-12.27 / -12.12
S 2	-12.53	-12.50	-12.52	-12.60 / -12.44
S 3	-11.96	-12.22	-12.09	-12.17 / -12.00
S 4	-12.87	-12.99	-12.93	-13.03 / -12.83
S 5	-13.58	-13.57	-13.58	-13.63 / -13.53
S 6	-14.07	-14.04	-14.06	-14.11 / -14.00
S 7	-14.22	-14.14	-14.18	-14.24 / -14.13
S 8	-14.37	-14.41	-14.39	-14.45 / -14.33

Exploration of the method's inherent uncertainty

Table S.3: Multi-step procedure to infer the method's inherent uncertainty from virtual experiments. Results presented in Fig. 7.

0	Pick a desired γ and assign values to the strains' growth parameters.
1a	Simulate the dynamics of transconjugant abundance (Eqn. 1).
1b	Pick values from the generated series, e.g. at hourly intervals, and translate abundances into probabilities (Eqn. 2).
1c	Transform the probabilities into binary data by comparing against random numbers drawn from a uniform distribution with range 0–1.
1d	Treat the binary data from step 1c as a virtual sample and estimate the value of γ as outlined in Sect. 2.1.
2	Repeat the above steps n times to mimic replicate experiments.
3	Repeat all above steps k times (e.g. k=10^3), to capture the variability in the estimates of γ.
4	Compute confidence intervals of γ from the output of step 3. Compare against the "true" γ picked in step 0.