File S2. Sequence and frequency of the most common 25 bp long sequence, the gene name of the flanking RAYT and the number of RAYTs, in all of the bacteria analyzed in this study.

File S3. Modeling REP populations in various *Pseudomonas* and *E. coli* strains. Column 1 corresponds to the sequence network in Figure 2 Columns 2 to 4 correspond to Figure 3A, B and C.

File S4. Proportion of symmetric REPINs in all identified sequences from all studied strains. Symmetric REPINs consist of two REP sequences (from the network of highly abundant 25bp long sequences) less than 130 nucleotides apart and in opposite orientation on the genome sequence (see Methods). In *Pseudomonas* REP singlets are likely immobile and decaying elements (Bertels and Rainey, 2011). This is probably not the case in *E. coli* where this symmetry does not exist.

File S5. Error threshold for the different fitness landscapes inferred for the different organisms. First column shows the duplication rates for the five mutation classes and the second column the corresponding equilibrium frequencies.

File S6. Mathematica code for calculating equilibrium frequencies, fitness values and error thresholds for all 10 REP populations.

File S7. Mathematica code as pdf.

File S8. Shows how adjusting mutation rates and fitness values affects the equilibrium frequencies of the quasispecies model.

File S9. Sequence frequencies of the different mutation classes for all 10 REP populations.

File S10. A quasispecies model with 5 mutation classes is optimal for the available REPIN data.