Supplementary Online Material

Analysis of the symmetric game shown in Figure 1b

The matrix shown in Figure 1b is a simplified version of the payoff matrix resulting from equations (1) under the approximation $t(r+1) \approx t$ for t << 1. The full matrix is shown in (S1).

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2}t(1-s) & (t-r(t+1)) + \delta[r(t+1)-ts] \\ (1-\delta)[t-r(t+1)] & t-r(t+1) \end{pmatrix}$$
(S1)

Evolutionarily stable states in the symmetric case

A symmetric game like that of (S1) can have pure and mixed evolutionarily stable strategies (ESSs). The pure strategies are those where both insertion sequences transpose either only selfishly or only cooperativity. Biologically, one can think of a mixed strategy is a strategy where individual ISs in one transposable element are able to initiate selfish and cooperative transposition at probabilities p and (1-p).

It is straightforward to determine the evolutionarily stable states (ESSs) in a simple symmetric game like this. Relevant theorems can be found in [6, Appendix B,35, Chapter 15]. Because of the four parameters, the model can display a variety of different outcomes. Different regions in its parameter space correspond to well-characterized game theoretical models. For example, for $\delta=1$ and ts/(t+1) < r < t(s+1)/[2(t+1)], the game becomes the well-known Prisoner's dilemma game, with selfish behavior as the only ESS. In studying the model's behavior systematically as a function of its parameters, I need to distinguish several cases.

Case 1: $\delta = 1$ *(Selfishness is genetically dominant)*

In this case, it is easy to show e.g., using [35, Theorem 15.15] that there are only pure ESSs.

$$r > \frac{ts}{t+1}$$
 selfishnesss is an ESS (S2a)
 $r < \frac{ts}{t+1}$ both selfishness and cooperativity are ESSs (S2b)

Case 2: $\delta = 0$ *(Cooperativity is genetically dominant)*

This is the case most favorable to the emergence of cooperativity. There are again only pure ESSs..

$$r > \frac{t(s+1)}{2(t+1)}$$
 selfishnesss is an ESS (S3a)

$$r < \frac{t(s+1)}{2(t+1)}$$
 cooperativity is an ESS (S3b)

Case 3: 0<δ<1.

Here, one can show that the pure ESSs exist under the following conditions

$$r > \frac{t}{t+1} \left(1 - \frac{1-s}{2(1-\delta)} \right)$$
 selfishness is an ESS (S4a)
 $r < \frac{ts}{t+1}$ cooperativity is an ESS (S4b)

When both (S4a) and (S4b) hold, then both selfishness and cooperativity can be pure ESSs and the initial state of the game's evolutionary dynamics (selfishness or cooperativity) becomes important for the outcome.

The right-hand side parenthetical expression of (S4a) will be important throughout, so I introduce

$$f = f(s,\delta) \coloneqq 1 - \frac{1-s}{2(1-\delta)}$$
(S5)

The qualitative behavior of $f(s,\delta)$ is shown in Figure S1. $f(s,\delta) < 0$ whenever $\delta > (1+s)/2$. In this case (S4a) holds regardless of the value of r, and selfish transposition is a pure ESS. For $\delta < (1+s)/2$, $f(s,\delta) > 0$, and the truth-value of (S4a) depends on both t and r.

Mixed strategy ESSs are always unique in symmetric two-player games like this. Therefore, no mixed ESS is possible whenever (S4a), (S4b), or both hold. In fact, a mixed strategy ESSs can only exist if $0 < \delta < (1/2)$, i.e., whenever cooperativity shows incomplete genetic dominance. To see this, note that for $\delta > 1/2$, $f(s,\delta) < s$. Assume that tand r have values in this case such that (S4a) does not hold. It follows that $r < [t/(t+1)]f(s,\delta) < ts/(t+1)$, such that (S4b) holds. In other words, at least one of (S4a) and (S4b) holds for $\delta > 1/2$, preventing the existence of a mixed ESS.

A mixed ESS exists whenever $\delta < 1/2$ and neither (S4a) nor (S4b) hold. This is the case for $s < r(t+1)/r < f(s,\delta)$. What is this ESS? To find it, one can take advantage of the fact that any mixed Nash equilibrium $\mathbf{p} = (p,1-p)$ of the game has to fulfill the condition $(\mathbf{Pp})_1 = (\mathbf{Pp})_2$, where $(\mathbf{Pp})_i$ is the *i*-th element of the product of the payoff matrix \mathbf{P} with \mathbf{p} . This condition prescribes two linear equations for p, the frequency of selfish transposition in the ESS. The solution of this equation is

$$p = \frac{\delta[r(t+1) - ts)]}{(1 - 2\delta)[\frac{t}{2}(1+s) - r(t+1)]}$$
(S6)

It is easy to verify that $0 \le p \le 1$ exactly for the conditions under which a mixed ESS exists, i.e., for $s \le r(t+1)/t \le f(s,\delta)$. It is also easy to check that $\mathbf{p} = (p,1-p)$ is not just a Nash equilibrium, but indeed an ESS [e.g., 35, Theorem 15.17].

These considerations completely characterize the game defined in (S1). Figure S2 shows a pictorial characterization of the different ESSs in different regions of parameter space. It displays the existing ESSs as a function of δ (vertical dimension) and the ratio r(t+1)/t. The ESSs shown all follow from (S2-S4), and from the qualitative behavior of $f(s,\delta)$ shown in Figure S1. Because all parameters will usually be much smaller than one, $(t+1)/t \approx 1/t$, and one can approximate ratios $r(t+1)/t \approx r/t$. I note in closing that the results above can also be obtained from an analysis of a replicator equation [e.g., 35, Theorem 15.17] with (S1) as the matrix of coefficients.

Evolutionarily stable states in the asymmetric case.

I will now examine the ESSs for the asymmetric version of the game whose payoff matrix is shown in Figure 1b. This asymmetric version can arise if two insertion sequences are different, as explained in the main text. Such an asymmetric game is characterized by two payoff matrices, P_L for IS_L and P_R for IS_R, whose structure is analogous to that of **P**. Specifically,

$$\mathbf{P}_{L} = \begin{pmatrix} \frac{1}{2}t_{L}(1-s) & t_{L} - r_{L}(t_{L}+1) + \delta[r_{L}(t_{L}+1) - t_{L}s] \\ (1-\delta)[t_{L} - r_{L}(t_{L}+1)] & t_{L} - r_{L}(t_{L}+1) \end{pmatrix}$$
(S7a)

$$\mathbf{P}_{R} = \begin{pmatrix} \frac{1}{2}t_{R}(1-s) & t_{R} - r_{R}(t_{R}+1) + \delta[r_{R}(t_{R}+1) - t_{R}s] \\ (1-\delta)[t_{R} - r_{R}(t_{R}+1)] & t_{R} - r_{R}(t_{R}+1) \end{pmatrix}$$
(S7b)

I assume that only *t* and r can differ between IS_L and IS_R , because δ results from the mechanics of the transposition process itself, and because *s* is a parameter influenced by the environment and not by the insertion sequence's structure.

I will next explore the ESSs for this asymmetric game. Both IS_L and IS_R have a continuous set of possible strategies characterized by the likelihoods $0 \le p_L \le 1$ and $0 \le p_L \le 1$, that IS_L and IS_R, respectively, undergo selfish transposition. The space of possible strategies is thus the unit square $(0,1) \times (0,1)$, whose corners correspond to the pure strategies of selfish $(p_X=1)$ and cooperative transposition $(p_X=0)$ of IS_X (X=L,R). A composite transposon's strategy is represented as a point (p_L, p_R) in the unit square.

Asymmetric games like this can never have mixed ESSs (although mixed Nash equilibria can exist). The only four possible ESSs are pure ESSs that correspond to the corners of the square, where either transposon can only transpose selfishly or cooperatively. The task is to identify which of these ESSs occur for different parameter values. There may be parts of parameter space where two pure ESSs can coexist in the unit square. In this case, the initial condition of the evolutionary dynamics of (p_L, p_R) may determine which of these ESSs is attained.

The evolutionary dynamics in strategy space can be described by a system of two nonlinear differential equations [17.8, and 17.15 in 35].

$$\dot{p}_L = p_L (1 - p_L) (L_{12} - (L_{12} + L_{21}) p_R) \dot{p}_R = p_R (1 - p_R) (R_{12} - (R_{12} + R_{21}) p_L),$$
(S8)

whose coefficients are defined as follows.

$$\mathbf{L} = \begin{pmatrix} 0 & L_{12} \\ L_{21} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \delta[r_L(t_L + 1) - t_L s] \\ (1 - \delta)[t_L - r_L(t_L + 1)] - \frac{1}{2}t_L(1 - s) & 0 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 0 & R_{12} \\ R_{21} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \delta[r_R(t_R + 1) - t_R s] \\ (1 - \delta)[t_R - r_R(t_R + 1)] - \frac{1}{2}t_R(1 - s) & 0 \end{pmatrix}$$
(S9)

The matrices L and R are obtained from the payoff matrices P_L and P_R by subtracting a constant from each column that renders the diagonal elements equal to zero. As for the simpler, symmetric case, I will distinguish three cases.

Case 1: $\delta = 1$ *(Selfishness is genetically dominant)*

I will first consider the case where either $r_L > t_L s / (t_L + 1)$, or $r_R > t_R s / (t_R + 1)$, or both. In this case, either L₁₂L₂₁<0, or R₁₂R₂₁<0, or both hold, and one can show [35, p 143-145] that no fixed point (unstable or saddle point) of the evolutionary dynamics (S8) exists in the interior of the unit square. No matter what the initial condition $\mathbf{p} = (p_L, p_R)$, the evolutionary dynamics will always lead to one of the corner ESSs. Which of these corner ESSs are attained? It is easy to verify [35, 17.1 and 17.2] that $\mathbf{p} = (0,0)$, where both ISs always transpose cooperatively, and $\mathbf{p} = (1,0)$, as well as $\mathbf{p} = (0,1)$, where one of the ISs always transposes cooperatively, are never ESSs. In other words, all-selfish transposition $\mathbf{p} = (1,1)$ is the only ESS in this case.

In the opposite case, where $r_L < t_L s / (t_L + 1)$ and $r_R < t_R s / (t_R + 1)$, there is an interior saddle point of the evolutionary dynamics (S8). It is given by

$$F = \left(\frac{t_R s - r_R(t_R + 1)}{t_R s - r_R(t_R + 1) + \frac{1}{2}t_R(1 - s)}, \frac{t_L s - r_L(t_L + 1)}{t_L s - r_L(t_L + 1) + \frac{1}{2}t_L(1 - s)}\right)$$
(S10)

In this case, there are two ESSs, $\mathbf{p} = (0,0)$, and $\mathbf{p} = (1,1)$. Which of these is attained will depend on the initial condition of (S8).

Case 2: $\delta = 0$ *(Selfishness is genetically recessive)*

In this case, no interior fixed point of the evolutionary dynamics exists, regardless of the parameter values. Because both L_{12} and R_{12} , as defined in (S9), are equal to zero, equation (S8) reduces to

$$\dot{p}_L = p_L (1 - p_L) (-L_{21} p_R) \dot{p}_R = p_R (1 - p_R) (-R_{21} p_L),$$

The equilibria of this system are $\mathbf{p} = (1,1)$, as well as the two edges of the unit square defined by $\mathbf{p} = (0, x)$ and $\mathbf{p} = (x,0)$, with $0 \le x \le 1$. It follows from the definition of an ESS for asymmetric games [35, 17.1 and 17.2] that no point on a continuum of such (degenerate) equilibria can be an ESS. The only candidate ESS is thus $\mathbf{p} = (1,1)$, which is an ESS for

$$r_{L} > \frac{t_{L}(1+s)}{2(t_{L}+1)}, r_{R} > \frac{t_{R}(1+s)}{2(t_{R}+1)}$$
(S11)

Case 3: 0<δ<1.

Because of the generality and complexity of this case, I will first state the mathematical conditions for the existence of the various ESSs, and then illustrate the qualitative behavior of the game in Figure S3. An interior saddle point exists for $\delta < 1/2$ if $t_X s / (t_X + 1) < r_X < t_X f(s, \delta) / (t_X + 1)$ for both X=L and X=R. This saddle exists for $\delta > 1/2$ if $t_X s / (t_X + 1) > r_X > t_X f(s, \delta) / (t_X + 1)$ for both X=L and X=R. Under the conditions where this saddle exists, its coordinates are,

 $F = (F_L, F_R)$

where

$$F_{L} = \left(\frac{\delta[r_{R}(t_{R}+1) - t_{R}s]}{\delta[r_{R}(t_{R}+1) - t_{R}s] + (1 - \delta)[t_{R} - r_{R}(t_{R}+1)] - \frac{1}{2}t_{R}(1 - s)}\right)$$
(S12)

$$F_{R} = \left(\frac{\delta[r_{L}(t_{L}+1) - t_{L}s]}{\delta[r_{L}(t_{L}+1) - t_{L}s] + (1 - \delta)[t_{L} - r_{L}(t_{L}+1)] - \frac{1}{2}t_{L}(1 - s)}\right)$$

One can show that the corners are ESSs under the following conditions

$$\mathbf{p} = (0,0) \quad \text{if} \quad r_L < t_L s / (t_L + 1), r_R < t_R s / (t_R + 1)
\mathbf{p} = (1,0) \quad \text{if} \quad r_L > t_L s / (t_L + 1), r_R < t_R f (s, \delta) / (t_R + 1)
\mathbf{p} = (0,1) \quad \text{if} \quad r_L < t_L f (s, \delta) / (t_L + 1), r_R > t_R s / (t_R + 1)
\mathbf{p} = (1,1) \quad \text{if} \quad r_L > t_L f (s, \delta) / (t_L + 1), r_R > t_R f (s, \delta) / (t_R + 1)$$
(S13)

Figure S3 shows a systematic qualitative analysis of the different regions of parameter space and the ESSs that exist in them. These equilibria follow straightforwardly (if tediously) from (S13), (S12), and (S5).

Summary of results from the asymmetric model.

Selfish transposition of *both* IS_L and IS_R is the only ESS as long as $r_L > (1+s)t_L / 2, r_R > (1+s)t_R / 2$. Cooperative transposition of *both* IS_L and IS_R is an ESS only if $r_L < st_L, r_R < st_R$. These conditions are exactly analogous to equations (2) and hold for all values of δ .

As in the symmetric model, there is a narrow gray zone in which the dominance of selfishness matters, characterized by $st_X < r_X < t_X(1+s)/2$. For the biologically most realistic case of selfish dominance (δ =1), selfish transposition is an ESS for both ISs already if $r_X > st_X$ for both X=L and X=R. For $r_X < st_X$ (both X=L and X=R), there are two coexisting ESSs, one where both ISs transpose only selfishly, the other one where both ISs transpose only cooperatively. An interior saddle point of the evolutionary dynamics separates their basins of attraction (Figure S3). As in the symmetric case, cooperative transposition faces an uphill fight, because most composite transposons may initially be composed of two identical ISs capable of selfish transposition. Again, cooperativity may have a chance only if $r_X << st_X$, in which case the basin of attraction of selfish transposition becomes very small. A simple dynamical model of transposable element spreading. A complementary perspective on the results obtained thus far can be obtained from a population biological model of transposable element spreading. The sketch below is stripped of all the complexities of the previous sections and contains only the bare essentials. Its main purpose is didactical: it provides a different way of obtaining intuition about the evolutionary dynamics of transposable elements.

Consider an (infinitely) large population of bacteria that harbors, at low frequencies, two kinds of transposable elements, selfishly transposing elements and cooperatively transposing (composite) elements that allow transfer of selectable genes to new genomes. Denote their frequencies at some point in time at *S* and *C*, respectively. If *t* is the rate at which transposition and horizontal transfer occur for both elements, if *r* is the rate at which composites get lost due to faulty transposition, and if *s* is the instantaneous probability that the selectable genes are required for survival, then one can write the following differential equation for the time derivatives of S and C, \dot{S} and \dot{C}

$$\dot{S} = tS(1-s)$$
$$\dot{C} = tC - rC$$

Here, the spreading of the selfish element is promoted by its horizontal transfer (tS) and delayed by a factor (1-s) that reflects the fact that selfish spreading does not transfer the selectable genes. Notice that if s=1 (the selectable genes are always required), the selfish element cannot spread at all. The spreading of the composite is promoted by its horizontal transfer (tC) and inhibited by its loss through faulty transposition (-rC). The key question is under what conditions the relative frequency C/(S+C) increases over time. We observe that

$$\left(\frac{C}{S+C}\right)^{\cdot} = \frac{\dot{C}S - C\dot{S}}{\left(S+C\right)^{2}} = \frac{CS}{\left(S+C\right)^{2}} [(t-r) - t(1-s)]$$

Thus, *C* increases in frequency relative to *S* at all times iff (t-r) > t(1-s), or if r < st, i.e., under the same (biologically unrealistic) conditions as in the game theoretic model.

Supplementary Figures and Captions

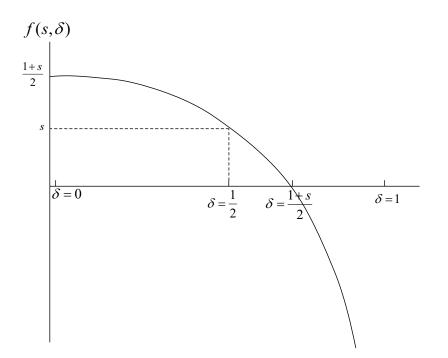


Figure S1

Figure S1: Qualitative behavior of the function $f(s, \delta) = 1 - (1 - s)/[2(1 - \delta)]$ from (S5). The function attains a maximum of (1+s)/2 at $\delta=0$ (dominant cooperativity), a value of *s* at $\delta=(1/2)$, a value of 0 at $\delta=(1+s)/2$, and approaches negative infinity as δ approaches one.

Figure S2 (next page): Evolutionarily stable strategies if both ISs are so similar that they receive symmetric payoffs. The vertical dimension shows five different values or intervals of the parameter δ for which the symmetric game shows qualitatively different behavior. The genetic dominance of selfish behavior over cooperative behavior increases with δ . For each of these five categories, the position along the horizontal arrow indicates the value of the ratio r(t+1)/t, where r is the probability of a DNA rearrangement eliminating an insertion sequence, and t is the joint probability of transposition/lateral gene transfer. If t<<1, then this ratio is well-approximated by r/t. Each of the five horizontal arrows is subdivided by tick-marks labeled by values at which the behavior of the model changes from one ESS to another. The possible ESSs are labeled as "C" (pure, cooperative ESS). "S" (pure, selfish ESS), and "M" (mixed ESS, as given by (S6)). Two symbols in a line-segment indicate that two pure ESSs coexist. The function $f(s, \delta)$ is defined by equation (S5).

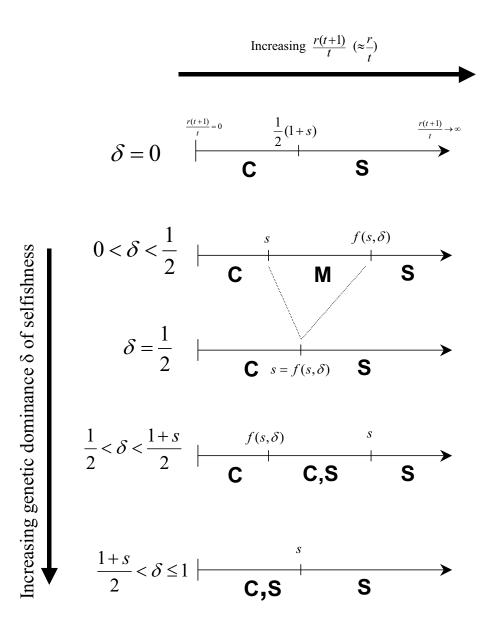
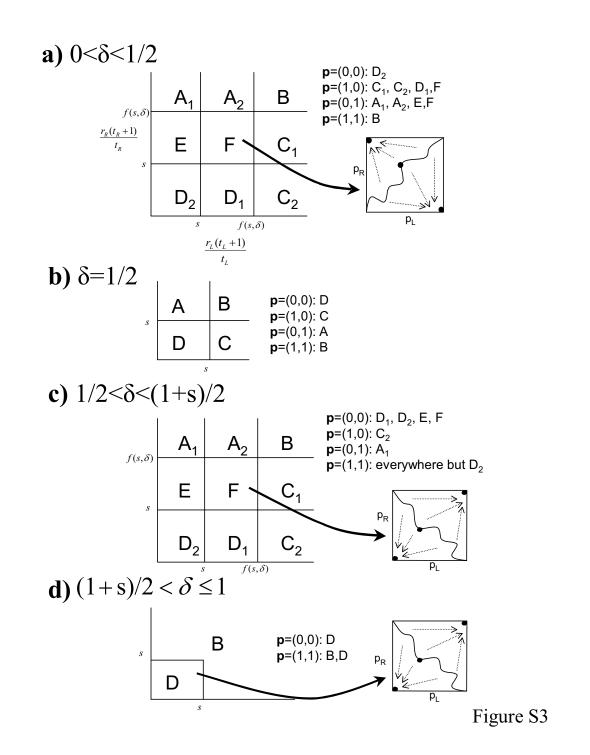


Figure S2

Figure S3 (next page): Evolutionarily stable strategies if the ISs receive asymmetric payoffs. The game is defined by the matrices (S7). The four different panels a)-d) distinguish between regions of parameter space with different values of δ , as shown. In each panel, the squares inscribed with capital letters A-F and open to the right and upper side, correspond to regions of parameter space characterized by different values of t_L , r_L , t_R , and r_R . Specifically, the horizontal axis of each square indicates the value of $r_L(t_L+1)/t_L$. The vertical lines through the square indicate points where $r_L(t_L+1)/t_L$ has the values s, and $f(s, \delta)$, as shown. Analogously, the vertical axis of each square indicates the value of $r_R(t_R + 1)/t_R$, and horizontal lines through the square indicate points where $r_R(t_R + 1)/t_R$ has the values s and and $f(s, \delta)$, as shown. The axes are labeled with $r_L(t_L+1)/t_L$ and $r_R(t_R+1)/t_R$ only in panel a), but the same labeling is implied for the three other panels. The squares are open to the right because the expressions parametrizing the axes can become arbitrarily large if t << r. The existence of particular corner ESSs are indicated to the right of the large squares. For example, " $\mathbf{p}=(0,0)$: D_2 " means that all-cooperative transposition is an ESS only in parameter space region D_2 . a) $0 < \delta < 1/2$: Only region F harbors a fixed point (saddle) of the game dynamics. The small square to the right of the panel indicates the state space of the dynamical system. It indicates the corner ESSs and the position of the saddle as small dots. The illustration is merely a schematic to help understand the qualitative dynamic, and both the location of this saddle and the basin of attraction depend on actual parameter values. b) $\delta = 1/2$. As δ approaches the value of $\frac{1}{2}$, the size of the regions F, A₂, D₁, E, and C₁ from panel a) shrink to zero, because $f(s,\delta)$ approaches s. c) $\frac{1}{2} < \delta < \frac{1+s}{2}$: Region F, as in a) contains a saddle point. However, notice that the corner ESSs are different from a) in region F. They now are (0,0) and (1,1). d) $(1+s)/2 \le \delta \le 1/2$. Region D contains a saddle that separates the basins of attraction of the two corner ESSs (0,0) and (1,1).



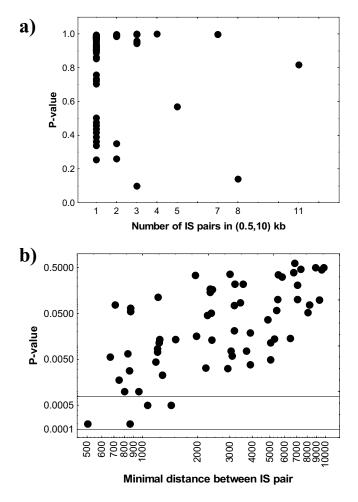


Figure S4

Figure S4: No significant enrichment of composite transposons in 376 completely sequenced bacterial genomes and plasmids. a) The horizontal axis shows the number of nearest neighbor insertion sequence pairs at a distance between 0.5 and 10kbp in the bacterial genomes examined. The P-value on the vertical axis is the likelihood that this number of insertion sequence pairs is greater than expected by chance alone, according to a randomization test described in Material and Methods. b) The horizontal axis shows the minimal distance between nearest neighbor IS pairs that are between 0.5 and 10 kb apart in a genome. The vertical axis shows the likelihood that this minimal distance is observed by chance alone. The upper and lower horizontal lines represent Bonferroni-corrected P-values of P=0.05 (0.00064=0.05/78) and P=0.01 (0.000128=0.01/78). Both panels contain n=63 data points derived from insertion sequences in 28 different families.