

Modeling and Estimation of Replication Fitness of HIV-1 *in Vitro* Experiments Using a Growth Competition Assay

hww@bst.rochester.edu

Tel: 585-275 - 6767

Fax: 585-273 - 1031

13th CROI

Poster No. 623

Hulin Wu 1, Yangxin Huang 2, Carrie Dykes 1, Dacheng Liu 3, Jingming Ma 1, Alan S. Perelson 4, and Lisa Demeter 1

1 University of Rochester School of Medicine and Dentistry, Rochester, NY, USA
3 Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT, USA

2 College of Public Health, University of South Florida, Tampa, FL, USA
4 Los Alamos National Laboratory, Los Alamos, NM, USA

Abstract

We develop mathematical models to describe viral/cellular dynamic interactions in the assay experiment from which competitive fitness indices or parameters are defined. These include the log fitness ratio (LFR), the log relative fitness (LRF), and the production rate ratio (PRR). From the population genetics perspective, we clarify the confusion and correct the inconsistency in the definition of relative fitness in the literature of HIV-1 viral fitness. Calculation and estimation methods based on two data points and multiple data points were proposed and were carefully studied. The developed methodologies are generally applicable to similar growth competitive assays. A user-friendly computational tool is also developed and publicly available on the web at: <http://www.urmc.rochester.edu/bstools/vfitness/virusfitness.htm>

Introduction

There is a growing interest in determining the replicative capacity of HIV-1 due to the potential impact of viral fitness on virus population size (viral load), drug resistance and AIDS progression. HIV-1 replication fitness has been proposed to be an important predictor of clinical outcome in HIV-1 infection. A critical question is whether assays that measure HIV-1 replication fitness *in vitro* correlate with fitness *in vivo*, and whether such assays can be used to predict prognosis or response to therapy. However, there is no clear consensus on how best to quantify HIV replication fitness, as measured *in vitro*. In previous studies, a measure of fitness is typically derived by plotting the ratio of the two competing variants on a logarithmic scale against time and estimating the linear slope of this graph which is used as the measure of fitness. However, the definition of relative fitness used is inconsistent with the conventional definition of relative fitness in population genetics, and does not take into account the death rate of the viral strains. We will introduce several parameters, based on HIV dynamic models, to quantify the relative fitness for the growth competition assay. The relationship between these parameters and the relative fitness in population genetics will be discussed. The confusion in the definitions of relative fitness in recent HIV literature will be clarified. More efficient statistical methods will be proposed to estimate the viral fitness parameters.

Materials and Methods

• **Growth competition assay:** The design and evaluation of a growth competition assay in which viral variants are detected using flow cytometry is described in detail in another publication by Dykes *et al.* We designed two vectors, pAT1 and pAT2, that are identical to the lab strain pNL4-3, except that they have the mouse *Thy1.1* or *Thy1.2* genes cloned in place of *nef*, respectively. The *Thy1.1* and *Thy1.2* proteins differ by only one amino acid and are expressed on the surface of infected cells. On days 3, 4, 5, 6, 7, and half the culture was removed and replaced with fresh media. The number of viable cells per mL of culture was determined before cells were removed, by staining with trypan blue and counting in a hemocytometer. Flow cytometry data were analyzed by Cell Quest (Becton Dickinson). Cells were analyzed using a density plot of forward scatter versus side scatter and the cell population was gated in order to eliminate debris. The number of viable wild-type or mutant infected cells was calculated by multiplying the total number of viable cells in the original culture by the percent of wild type or mutant as determined by flow cytometry.

• **Mathematical models:** A complete model with five compartments

$$\begin{aligned} dT/dt &= \lambda T - k_M^*MT - k_W^*WT - d_T T \\ dT_M/dt &= k_M^*MT - \delta_M T_M \\ dT_W/dt &= k_W^*WT - \delta_W T_W \\ dM/dt &= N_M \delta_M T_M - c_M M \\ dW/dt &= N_W \delta_W T_W - c_W W \end{aligned}$$

Where T , T_M , T_W , M and W are numbers of uninfected target cells, cells infected by infectious mutant virus, cells infected by infectious wild-type virus, mutant virus, and wild-type virus, respectively. Others are model parameters.

• **Mathematical models:** Assume the quasi-steady state and constant target cells. The model can be simplified to

$$\begin{aligned} dT_M/dt &= (k_M T - \delta_M) T_M \\ dT_W/dt &= (k_W T - \delta_W) T_W \end{aligned}$$

Where k_M and k_W are the respective infection rates at which cells become infected by M and W ; δ_M and δ_W are the respective death rates of T_M and T_W .

Materials and Methods (cont.)

• **Competitive Fitness Parameters**

$$\begin{aligned} \bullet \text{Log-relative fitness (LRF): } & d = \ln(1+s) = gm - gw = (km - kw)T / (\delta_M - \delta_W) \\ \bullet \text{Log-fitness ratio (LFR): } & r = gm/gw \\ \bullet \text{Production rate ratio (PRR): } & p = km/kw \\ \bullet \text{Relative fitness (RF): } & 1+s = \exp(gm - gw) \end{aligned}$$

where $gm = km T - \delta_M$ and $gw = kw T - \delta_W$ are the net growth rates of mutant and wild-type infected cells

• **Calculations of Fitness Parameters from Two Data Points**

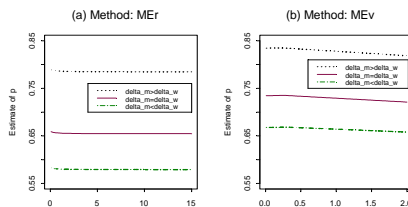
$$\begin{aligned} 1. \text{Log-relative fitness (LRF): } & d = g_M - g_W = \frac{1}{\Delta t} \ln \left[\frac{T_M(t_2) T_W(t_1)}{T_W(t_2) T_M(t_1)} \right] \\ 2. \text{Log-fitness ratio (LFR): } & r = g_M / g_W = \frac{\ln[T_M(t_2) / T_M(t_1)]}{\ln[T_W(t_2) / T_W(t_1)]} \\ 3. \text{Production rate ratio (PRR): } & p = \frac{k_M}{k_W} = \frac{\ln[T_M(t_2) / T_M(t_1)] + \delta_M \Delta t}{\ln[T_W(t_2) / T_W(t_1)] + \delta_W \Delta t} \\ 4. \text{Relative fitness (RF): } & 1+s = \exp(d) = \exp \left\{ \frac{1}{\Delta t} \ln \left[\frac{T_M(t_2) T_W(t_1)}{T_W(t_2) T_M(t_1)} \right] \right\} \end{aligned}$$

• **Estimation Methods of PRR from Multiple Data Points**

1. **AM method:** Calculate the average production rate ratio among all pairs of successive time-points according to Marder *et al.*
2. **LS method:** Standard linear regression model without intercept.
3. **MEr method:** Linear regression with measurement error in covariate when ratio (p) of measurement variances is known.
4. **MEv method:** Linear regression with measurement error in covariate when measurement variance (σ_e^2) in covariate is known.

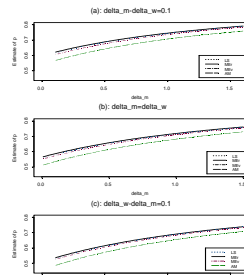
Results

• **Figure 1: Effects of Measurement Errors on PRR Estimates**

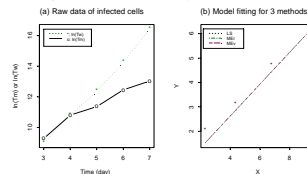


Results (cont.)

• **Figure 2: Effects of Death Rate on PRR Estimates**



• **Figure 3: Experiment Data Analysis**



• **Table 1. Comparison of the methods of fitness parameter estimation with a data set from our growth competitive assay.**

Parameter	Method	Estimate	SD
p	LS	0.654	0.034
	MEr	0.656	0.042
	MEv	0.655	0.042
r	AM	0.614	NA
	LS	0.558	0.043
	MEr	0.560	0.052
d	MEv	0.559	0.053
	AM	0.510	NA
	LS	-0.921	0.271
1+s	AM	-0.921	NA
	LS	0.398	0.108
AM	AM	0.443	NA

Conclusion

• Developed mathematical models to describe viral and cellular dynamics for *in vitro* experiments using a growth competition assay.

• Among three fitness indices, the log relative fitness (LRF) was shown to be the easiest to measure and most convenient to use in practice.

• Both production rate ratio (PRR) and log fitness ratio (LFR) require calculations of the ratio for mutant and wild-type infected cells at two different time points.

• Proposed methods to calculate the relative fitness indices from two data points, and statistical methods to estimate these indices from multiple data points.

• Simulation studies show that we may choose one of the LS, MEr and MEv methods to estimate the relative fitness parameters based on the information availability of measurement error variances.

• The effect of death rates of infected cells on the estimate of production rate ratio (PRR) is significant.

• Two newly defined fitness indices, the log fitness ratio (LFR) and the log relative fitness (LRF), are more attractive in practical use since they do not depend on the death rates of infected cells and are easier to estimate from the experimental data.

• Computational tools are available at <http://www.urmc.rochester.edu/bstools/vfitness/virusfitness.htm>.

• **Table 2. A general guideline for selecting methods to estimate fitness parameters.**

known	p known	Method recommended
No	No	LS
No	Yes ($p > 1$)	LS
No	Yes ($p \leq 1$)	MEr
Yes	No	MEv
Yes	Yes ($p > 1$)	MEv
Yes	Yes ($p \leq 1$)	MEr