# Supporting Information

In this section, standard statistical procedures yield two estimators for , **** and . They do not appear in the main text, because they prove inferior to ****.

First, after substituting  for , Eq (12) can be maximized numerically to generate a MLE estimator . Applying the derivative operator a second time in Eq (13) and negating produces the observed Fisher information

,

so

,

where the first approximation follows from the approximation , the second approximation is a standard use of Fisher information; and the final equality derives from Eq . Note that calculation of the expected folded AFS  () in Eq only requires  for .

The simplicity of Eq (19) suggests also comparing the approximate MLEs  and  to an estimate  derived from applying the method of moments to . To begin, define

.

The moment

.

In a standard notation, let  denote the inverse of a function  under functional composition (often denoted by “”). In the method of moments, the basic reproduction number  is estimated from (the observable) . Logarithms  provide a convenient scale for , so writing  temporarily (i.e., only to reduce visual noise in Eq , next), yields

.

Under the assumption that  (i.e.,  has a small bias and coefficient of variation), the variance of Eq is

,

because  implies that .

Fig 2 for  has two analogs in this Appendix: Fig S1 for  and Fig S2 for . The text surrounding Fig 2 explains the graphs for : Fig S1 and Fig S2 have the same structure as Fig 2.

**Fig S1. Plots for the maximum likelihood estimate (delta approximation) **

Fig S1 for **** has the same format as Fig 2 for **** in the main text.

Fig S1f is absent because for **** and , there were numerical difficulties associated with generating and multiplying extremely small infinite sums and extremely large binomial coefficients in Eq (4) for. Meticulous numerical programming might eliminate some of the numerical instability, but we judged the effort unjustified given the remarkable success of the estimator **** .

**Fig S2. Plots for the estimate  from the method of moments**

Fig S2 for  has the same format as Fig 2 for **** in the main text.

Fig S2 for  resembles Fig 2 for ****. It differs, however, in that near , Fig S2e displays some systematic overestimation away from the perfect estimate ****, with Fig S2f showing an unacceptable worsening of the overestimation. The error estimate  in Fig S2d is slightly more accurate than  in Figure 2d, but the error estimate deteriorates badly in Fig S2f, presumably again from numerical difficulties caused by the infinite series.

The inferior estimators  and  displayed a pronounced upward bias (of uncertain origin) in the important regime where  is near 1.0 (see Fig S1e and Fig S2f). The two inferior estimators and their estimated errors displayed occasional numerical instabilities (for , see Fig S1e, the non-existent Fig S1f; and for , see the large variations in  indicated by the dashed lines in Fig S2f). The inferior estimators can also involve staggeringly large binomial coefficients in computations that truncate infinite series. Current deep-sequencing technologies can generate 105 reads of size comparable to gp120 [[1](#_ENREF_1),[2](#_ENREF_2)]. In applications utilizing extremely large sample sizes, therefore, the numerical extremes in  and  render them undependable for routine computation.

# References

1. Carneiro MO, Russ C, Ross MG, Gabriel SB, Nusbaum C, et al. (2012) Pacific biosciences sequencing technology for genotyping and variation discovery in human data. BMC Genomics 13: 375.

2. McElroy K, Thomas T, Luciani F (2014) Deep sequencing of evolving pathogen populations: applications, errors, and bioinformatic solutions. Microbial Informatics and Experimentation 4: 1.