Modelling the effects of phylogeny and body size on within-host pathogen replication and immune response: Supplementary Information

Soumya Banerjee[1*], Alan S. Perelson[2], Melanie Moses[3]
1 University of Oxford, Oxford, Oxfordshire, United Kingdom
2 Los Alamos National Laboratory, Los Alamos, New Mexico, USA
3 University of New Mexico, Albuquerque, New Mexico, USA
∗ **E-mail: Corresponding soumya.banerjee@maths.ox.ac.uk**

Supplementary Information

Overview of Methods

We use a Bayesian inference approach to estimate model parameters. Assume that a model (in our case a differential equation model describing how virus concentration changes over time in serum) contains parameters Θ. The Bayesian approach allows us to include prior knowledge about model parameters in a systematic fashion. If we have information about Θ (e.g., from experimental evidence) which needs to be incorporated in our analysis, this is represented as a prior probability distribution $P(\Theta)$. Bayes Rule allows us to incorporate the prior knowledge about parameters, $P(\Theta)$, and experimental data, D, to derive a posterior distribution of parameters:

$$P(\Theta|D) = \frac{P(D|\Theta) \cdot P(\Theta)}{P(D)} \tag{1}$$

The multi-level hierarchical Bayesian model mimics the hierarchical nature of host phylogeny. There are three levels in the hierarchical tree representation: individual, species and genus (Fig. 1). Each level of the tree has an equation describing the distribution of model parameters (differential equation models, Eqs. 2 - 5 and Eqs. 2 - 4, 6 - 7). The differential equation and hierarchical Bayesian models are explained in greater detail in the next section.

Our goal is to compare two different hierarchical Bayesian models: a multi-level model that has three levels of hierarchy (individual, species and genus) (Fig. 1) and an aggregated model with two levels of hierarchy (all individuals of all species pooled together under a single genus) (Fig. 2). We compare the accuracy of parameter estimates with respect to ground truth at the individual, species and genus level. We also experiment with different degrees of variation of parameters between species and within species.

Differential Equation Models of Viral Dynamics and Immune Response

We use two different viral dynamic models to account for the observed plasma viremia. The first model assumes that infection is target-cell limited in birds, i.e. the concentration of virus reaches a peak and then declines when few susceptible target cells remain. Models of target cell limited acute infection have been developed for HIV [1], influenza A virus [2], hepatitis C virus [3], simian immunodeficiency virus [4], dengue virus [5] and Zika virus [6, 7]. Here we use a target cell limited model with an eclipse phase, given by the following differential equations:

$$\frac{dT}{dt} = -\beta TV \tag{2}$$

$$\frac{dI_1}{dt} = \beta TV - kI_1 \tag{3}$$

$$\frac{dI_2}{dt} = kI_1 - \delta I_2 \tag{4}$$

$$\frac{dV}{dt} = pI_2 - \gamma V - \beta TV \tag{5}$$

where T is the density of uninfected target cells, and V is the viral titer in serum. Target cells become infected by virus at rate βTV, where β is the rate constant characterizing infection. The initial viral titer and the initial number of target cells are denoted V_0 and T_0, respectively. The initial density of infected cells is assumed to be zero. The separation of infected cells into two classes, I_1 cells that are infected but not yet producing virus and I_2 cells that produce virus, is similar to that in a model proposed earlier for influenza infection [2]. This separation increases the realism of the model, since delays in the production of virus after the time of initial infection are part of the viral life cycle (the eclipse phase). The parameter $1/k$ is the average transition time from I_1 to I_2. Productively infected cells (I_2) release virus at an average rate p per cell and die at rate δ per cell, where $1/\delta$ is the average life span of a productively infected cell. Free infectious virus is cleared at rate γ per infectious unit per day, for example by phagocytosis or loss of infectivity and is lost by entering cells during the infection process at rate βTV.

The Bayesian model infers (V_0,β,p,δ) from the viral titer data.

We also use a more sophisticated model that assumes viral decline is due to an adaptive antibody (induced IgM) response. Humoral immunity is an essential component of the immune response to WNV, as neutralizing antibodies limit dissemination of infection [8, 9]. Diamond et al. [9] infected wild type mice subcutaneously with WNV and measured titers of neutralizing antibody (analyzed in [10]). The data can be described by the following piecewise linear function:

$$A(t) = \begin{cases} 0 & , \quad t < t_i \\ \eta(t - t_i) & , \quad t \geq t_i \end{cases} \tag{6}$$

The level of neutralizing antibody at time t, $A(t)$, measured by the plaque reduction neutralization test (PRNT) is 0 before time t_i and increases linearly with time after that with rate η. We assume that neutralizing antibody, A, binds virus, V, and neutralizes it with rate constant ρ, so that infectious virus is lost at rate $\rho A(t)V$. The model including neutralizing antibody consists of Eqs. 2-4 and Eq. 6 with Eq. 5 replaced by:

$$\frac{dV}{dt} = pI_2 - \gamma V - \beta TV - \rho A(t)V \tag{7}$$

The Bayesian model infers (V_0,β,p,δ,ρ,t_i) from the viral titer data.

The ordinary differential equations describing our viral kinetic models were solved numerically in Matlab [11]. The Runge-Kutta 4 method of integration was employed. All model parameters and virus concentration are logged (base 10) in order to stabilize variance and ensure positive estimates from the Bayesian inference. A sample plot of the target limited model prediction of virus concentration over time (compared to observed virus concentration data) using one representative set of parameters inferred from the Bayesian model is shown in Fig. 3.

Bayesian Aggregated Model

The aggregated model has individuals of different species. The individuals are aggregated to form a tree of height 2. The model is shown graphically in Fig. 2 (Left panel).

We denote the number of individuals by n and the number of experimental measurements of virus concentration on the ith individual by m_i. $y_{ij}(t_j)$ represents the experimental measurements of logarithmic virus concentration in serum for the ith individual at times $t_j (j = 1, 2, 3, ..., m_i)$. For notational convenience, let $\mu = (log_{10}V_0, log_{10}\beta, log_{10}p, log_{10}\delta)^T$, $\theta_i = (log_{10}V_{0i}, log_{10}\beta_i, log_{10}p_i, log_{10}\delta_i)^T$, $\Theta = \{\theta_i, i = 1..n\}$, $Y = \{y_{ij}, i = 1..n, j = 1..m_i\}$ and $f_{ij}(\theta_i, t_j) = log_{10}V_{ij}(\theta_i, t_j)$ where $V_{ij}(\theta_i, t_j)$ denotes the numerical solution for $V(t)$ in Eq. 5 for the ith individual at time t_j.

The Bayesian non-linear mixed effects aggregated model can be written as the following three stages [12–14]:

1. Within-individual variation

$$[y_i|\theta_i, \sigma^2] \sim Normal(f_i(\theta_i), \sigma^2) \tag{8}$$

2. Between-individual variation

$$[\theta_i|\mu, \Sigma] \sim Normal(\mu, \Sigma) \tag{9}$$

3. Prior distributions

$$\sigma^{-2} \sim Gamma(a, b), \Sigma^{-1} \sim Wishart(\Omega, \nu), \mu \sim Normal(\eta, \Lambda) \tag{10}$$

The Normal, Wishart and Gamma distributions are chosen to simplify calculations [12, 15].

The full conditional distributions for θ_i, σ^{-2}, μ and Σ^{-1} can be written as [12, 15]:

$$\left[\theta_i|\sigma^{-2}, \mu, \Sigma, \eta, \Theta, Y\right] \propto exp\left(-\frac{\sigma^{-2}}{2}\sum_{j=1}^{m_i}\left[y_{ij} - f_{ij}(\theta_i, t_j)\right]^2 - \frac{1}{2}(\theta_i - \mu)^T\Sigma^{-1}(\theta_i - \mu)\right) \tag{11}$$

$$\left[\sigma^{-2}|\mu, \Sigma, \Theta, Y\right] \sim Gamma\left(a + \frac{\sum_{i=1}^{n} m_i}{2}, A^{-1}\right) \tag{12}$$

$$\left[\mu|\sigma^{-2}, \Sigma, \Theta, Y\right] \sim Normal\left(B^{-1}C, B^{-1}\right) \tag{13}$$

$$\left[\Sigma^{-1}|\sigma^{-2}, \mu, \Theta, Y\right] \sim Wishart\left(D^{-1}, n + \nu\right) \tag{14}$$

where $A = b^{-1} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\left[y_{ij} - f_{ij}(\theta_i, t_j)\right]^2$, $B = n\Sigma^{-1} + \Lambda^{-1}$, $C = \Sigma^{-1}\sum_{i=1}^{n}\theta_i + \Lambda^{-1}\eta$ and $D = \Omega^{-1} + \sum_{i=1}^{n}(\theta_i - \mu)\cdot(\theta_i - \mu)^T$.

The parameters σ^{-2}, μ and Σ^{-1} are sampled using a Gibbs sampler and a Metropolis-Hastings algorithm is used to sample θ_i (both algorithms are described later).

Bayesian Multi-level Model

The multi-level hierarchical model has groups (species) with individuals in each group. The groups are arranged hierarchically to form a tree of height 3. The model is shown graphically in Fig. 1 (Left panel). Let there be m distinct species indexed by k and the number of individuals in the kth group is represented by n_k. Let us represent the number of experimental measurements on the ith individual of the kth species by m_{ik}. $y_{ijk}(t_j)$ represents the experimental measurements of logarithmic virus concentration in serum for the ith individual belonging to the kth species at times $t_j (j = 1, 2, 3, ..., m_{ik})$.

For notational convenience, we define the individual level distribution (ith individual belonging to the kth species) by $\theta_{ik} = (log_{10}V_{0ik}, log_{10}\beta_{ik}, log_{10}p_{ik}, log_{10}\delta_{ik})^T$. We define the kth species level distribution by $\mu_k = (log_{10}V_{0,k}, log_{10}\beta_k, log_{10}p_k, log_{10}\delta_k)^T$ and the genus level distribution by $\eta = (log_{10}V_0, log_{10}\beta, log_{10}p, log_{10}\delta)^T$.

Let $\Theta = \{\theta_{ik}, i = 1..n_k, k = 1..m\}$, $Y = \{y_{ijk}, i = 1..n_k, j = 1..m_{ik}, k = 1..m\}$ and $f_{ijk}(\theta_{ik}, t_j) = log_{10}V_{ijk}(\theta_{ik}, t_j)$ where $V_{ijk}(\theta_{ik}, t_j)$ denotes the numerical solution for $V(t)$ in Eq. 5 for the ith individual (belonging to the kth species) at time t_j.

We write this Bayesian non-linear mixed effects multi-level model as the following four stages:

1. Within-individual variation

$$[y_{ik}|\theta_{ik}, \sigma_k^2] \sim Normal(f_{ik}(\theta_{ik}), \sigma_k^2) \tag{15}$$

for the ith individual belonging to the kth species.

2. Between-individual variation

$$[\theta_{ik}|\mu_k, \Sigma_k] \sim Normal(\mu_k, \Sigma_k) \tag{16}$$

3. Between species variation

$$\sigma_k^{-2} \sim Gamma(a, b), \Sigma_k^{-1} \sim Wishart(\Omega, \nu), \mu_k \sim Normal(\eta, \Lambda) \tag{17}$$

for the kth species. The Normal, Wishart and Gamma distributions are chosen to simplify calculations.

4. Genus level prior distributions

$$\eta \sim Normal(x, y), \Lambda^{-1} \sim Wishart(p, q) \tag{18}$$

Fig. 1 (right panel) represents these equations graphically in a plate diagram. In a graphical model like Fig. 1 (right panel), for any node u, we can represent the remaining nodes by U_{-u} and the full conditional distribution $P(u, U_{-u})$ is $\propto P(u|parents[u]) \cdot \prod_{w \in children[u]} P(w|parents[w])$ [16]. The full conditional distribution for u contains a prior component (from the parents of u) and a likelihood component (from each child of u).

Following the principle above and simplifying, we derived the full conditional distributions for θ_{ik}, σ_k^{-2}, μ_k, Σ_k^{-1}, η and Λ^{-1}:

$$\begin{aligned}\left[\theta_{ik}|\sigma_k^{-2}, \mu_k, \Sigma_k, \eta, \Lambda^{-1}, \Theta, Y\right] \propto exp\Bigg(&-\frac{\sigma_k^{-2}}{2}\sum_{j=1}^{m_{ik}}\left[y_{ijk} - f_{ijk}(\theta_{ik}, t_j)\right]^2 \\ &-\frac{1}{2}(\theta_{ik} - \mu_k)^T\Sigma_k^{-1}(\theta_{ik} - \mu_k)\end{aligned} \tag{19}$$

$$\left[\sigma_k^{-2}|\mu_k, \Sigma_k, \eta, \Lambda^{-1}, \Theta, Y\right] \sim Gamma\left(a + \frac{\sum_{i=1}^{n} m_{ik}}{2}, A_k^{-1}\right) \tag{20}$$

$$\left[\mu_k|\sigma_k^{-2}, \Sigma_k, \eta, \Lambda^{-1}, \Theta, Y\right] \sim Normal\left(B_k^{-1}C_k, B_k^{-1}\right) \tag{21}$$

$$\left[\Sigma_k^{-1}|\sigma_k^{-2}, \mu_k, \eta, \Lambda^{-1}, \Theta, Y\right] \sim Wishart\left(D_k^{-1}, n_k + \nu\right) \tag{22}$$

$$\left[\eta|\sigma_k^{-2}, \mu_k, \Sigma_k^{-1}, \Lambda^{-1}, \Theta, Y\right] \sim Normal\left(z, Z\right) \tag{23}$$

$$\left[\Lambda^{-1}|\sigma_k^{-2}, \mu_k, \Sigma_k^{-1}, \eta, \Theta, Y\right] \sim Wishart\left(E, q + m\right) \tag{24}$$

where $A_k = b^{-1} + \frac{1}{2}\sum_{i=1}^{n_k}\sum_{j=1}^{m_{ik}}\left[y_{ijk} - f_{ijk}(\theta_{ik}, t_j)\right]^2$, $B_k = n_k\Sigma_k^{-1} + \Lambda^{-1}$, $C_k = \Sigma_k^{-1}\sum_{i=1}^{n_k}\theta_{ik} + \Lambda^{-1}\eta$, $D_k = \Omega^{-1} + \sum_{i=1}^{n_k}(\theta_{ik} - \mu_k) \cdot (\theta_{ik} - \mu_k)^T$, $Z = \left(U^{-1} + y^{-1}\right)^{-1}$, $z = Z\left(U^{-1}u + y^{-1}x\right)$ and $E^{-1} = p^{-1} + \sum_{k=1}^{m}(\mu_k - \eta) \cdot (\mu_k - \eta)^T$.

The parameters σ_k^{-2}, μ_k, Σ_k^{-1}, η and Λ^{-1} are sampled using a Gibbs sampler and a Metropolis-Hastings algorithm is used to sample θ_{ik} (both algorithms are described later).

Metropolis-Hastings Sampler

We use two popular MCMC algorithms known as the Metropolis-Hastings sampler and the Gibbs sampler. We use these algorithms to sample from the posterior distribution $P(\Theta|D)$. Let $\Theta^{(t)}$ represent the state of the Markov chain at iteration t of the algorithm. The Metropolis-Hastings sampler uses a proposal distribution $q(\Theta|\Theta^{(t-1)})$ to generate a candidate point, Θ^*, that is conditioned on the previous state of the sampler, $\Theta^{(t-1)}$. In the next step, the algorithm either accepts or rejects the candidate point based on an acceptance probability:

$$\alpha = \min(1, \frac{P(\Theta^*)}{P(\Theta^{(t-1)})} \cdot \frac{q(\Theta^{(t-1)}|\Theta^*)}{q(\Theta^*|\Theta^{(t-1)})}) \tag{25}$$

The Metropolis-Hastings algorithm is as follows:

1. Set $t = 1$
2. Generate an initial value u, and set $\Theta^{(t)} = u$
3. Repeat

   $t = t + 1$

   Generate a candidate Θ^* from the proposal distribution $q(\Theta|\Theta^{(t-1)})$

   Evaluate the probability of acceptance, $\alpha = \min(1, \frac{P(\Theta^*)}{P(\Theta^{(t-1)})} \cdot \frac{q(\Theta^{(t-1)}|\Theta^*)}{q(\Theta^*|\Theta^{(t-1)})})$

   Generate a x from a Uniform(0,1) distribution

   If $x \leq \alpha$, accept the candidate solution and set $\Theta^t = \Theta^*$, else set $\Theta^t = \Theta^{(t-1)}$
4. Until $t =$ required number of iterations

The Metropolis-Hastings sampler has the advantage of not requiring knowledge of the normalizing constant, $P(D)$. In our case the parameter Θ is multi-variate and has different components corresponding to the different ODE model parameters. We accept or reject the candidate solution involving all components as a block simultaneously. This is known as a block-wise updating scheme. A component-wise updating scheme updates each component in turn and independently of each other.

We use the Metropolis-Hastings sampler to sample the ordinary differential equation model parameters θ_i (Bayesian aggregated model) and θ_{ik} (Bayesian multi-level model).

Gibbs Sampler

The Gibbs sampler is an algorithm in which all samples are accepted and hence is computationally more efficient than the Metropolis-Hastings sampler. However, the Gibbs sampler can only be applied in the case when we know the full conditional distribution of each component of the parameter Θ conditioned on all other components.

As an illustrative example, we show the workings of the Gibbs sampler for the joint two-component distribution $g(\theta_1, \theta_2)$. The Gibbs sampler can be used if we can derive the distribution of each component conditioned on the other, i.e. in this case the conditional distributions $g(\theta_1|\theta_2 = \theta_2^{(t)})$ and $g(\theta_2|\theta_1 = \theta_1^{(t)})$. The Gibbs sampler for the case of two-components is as follows:

1. Set $t = 1$
2. Generate an initial value $u = (u_1, u_2)$, and set $\Theta^{(t)} = u$

3. Repeat

   $t = t + 1$

   Sample $\theta_1^{(t)}$ from the conditional distribution $g(\theta_1|\theta_2 = \theta_2^{(t-1)})$

   Sample $\theta_2^{(t)}$ from the conditional distribution $g(\theta_2|\theta_1 = \theta_1^{(t)})$

4. Until $t =$ required number of iterations

We update each component sequentially in turn (systematic-scan Gibbs sampler). In the Bayesian aggregated model, the parameters σ^{-2}, μ and Σ^{-1} are sampled using a Gibbs sampler. In the Bayesian multi-level model, the Gibbs sampler is used to sample the parameters σ_k^{-2}, μ_k, Σ_k^{-1}, η and Λ^{-1}.

Our approach is to use the Metropolis-Hastings sampler within the Gibbs sampler; this involves using the computationally efficient Gibbs update steps whenever the full posterior conditional distributions are known and using Metropolis-Hastings acceptance steps when the full conditional distribution is not available.

References

1. Stafford M, Cao Y, Ho D, Corey L (2000) Modeling plasma virus concentration and CD4+ T cell kinetics during primary HIV infection. J Theor Biol 203: 285–301.

2. Baccam P, Beauchemin C, Macken C, Hayden F, Perelson A (2006) Kinetics of influenza A virus infection in humans. J Virol 80: 7590–7599.

3. Rong L, Dahari H, Ribeiro RM, Perelson AS (2010) Rapid emergence of protease inhibitor resistance in hepatitis C virus. Sci Transl Med 2: 30ra32–30ra32.

4. Vaidya NK, Ribeiro RM, Miller CJ, Perelson AS (2010) Viral dynamics during primary simian immunodeficiency virus infection: effect of time-dependent virus infectivity. J Virol 84: 4302–4310.

5. Clapham HE, Quyen TH, Kien DTH, Dorigatti I, Simmons CP, et al. (2016) Modelling virus and antibody dynamics during dengue virus infection suggests a role for antibody in virus clearance. PLoS Comput Biol 12: e1004951.

6. Osuna CE, Lim SY, Deleage C, Griffin BD, Stein D, et al. (2016) Zika viral dynamics and shedding in rhesus and cynomolgus macaques. Nature Medicine 22: 1448–1455.

7. Best K, Guedj J, Madelain V, de Lamballerie X, Lim SY, et al. (2017) Zika plasma viral dynamics in nonhuman primates provides insights into early infection and antiviral strategies. Proc Natl Acad Sci USA 114: 8847–8852.

8. Diamond MS, Sitati EM, Friend LD, Higgs S, Shrestha B, et al. (2003) A critical role for induced IgM in the protection against West Nile virus infection. J Exp Med 198: 1853–1862.

9. Diamond MS, Shrestha B, Marri A, Mahan D, Engle M (2003) B cells and antibody play critical roles in the immediate defense of disseminated infection by West Nile encephalitis virus. J Virol 77: 2578–2586.

10. Banerjee S, Guedj J, Ribeiro RM, Moses M, Perelson AS (2016) Estimating biologically relevant parameters under uncertainty for experimental within-host murine West Nile virus infection. J R Soc Interface 13: 20160130.

11. The MathWorks Inc (2011). MATLAB version 7.13.0.564.

12. Davidian M, Giltinan D (1995) Nonlinear Models for Repeated Measurement Data. Chapman & Hall/CRC.

13. Huang Y, Wu H, Acosta EP (2010) Hierarchical Bayesian inference for HIV dynamic differential equation models incorporating multiple treatment factors. Biometrical J 52: 470–486.

14. Huang Y, Liu D, Wu H (2006) Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. Biometrics 62: 413–423.

15. Gelfand A, Hills S, Racine-Poon A, Smith A (1990) Illustration of bayesian inference in normal data models using gibbs sampling. J Am Stat Assoc 85: 972–985.

16. Gilks W, Richardson S, Spiegelhalter D (1995) Markov Chain Monte Carlo in practice: interdisciplinary statistics, volume 2. Chapman & Hall/CRC.

17. Komar N, Langevin S, Hinten S, Nemeth N, Edwards E, et al. (2003) Experimental infection of North American birds with the New York 1999 strain of West Nile virus. Emerg Infect Dis 9: 311–22.

Figure Legends

Figure 1. Left Panel: Multi-level hierarchical model with two groups. Each group has three individuals. Also shown are the genus, species and individual levels. Right Panel: Plate diagram for the multi-level hierarchical model. The plate denotes iteration of parameters and the number enclosed in the plate shows the number of iterations.

Figure 2. Left Panel: Aggregated model with two groups combined. Each group has three individuals. Also shown are the genus, species and individual levels. Right Panel: Plate diagram for the aggregated model. The plate denotes iteration of parameters and the number in the plate shows the number of iterations.

Figure 3. A sample prediction of the ODE model given by Eqs. 2-5 for plasma virus concentration (in log_{10} PFU/mL) over time post infection (blue) and experimental data on virus concentration (red). Data show viremia of great-horned owls from [17].

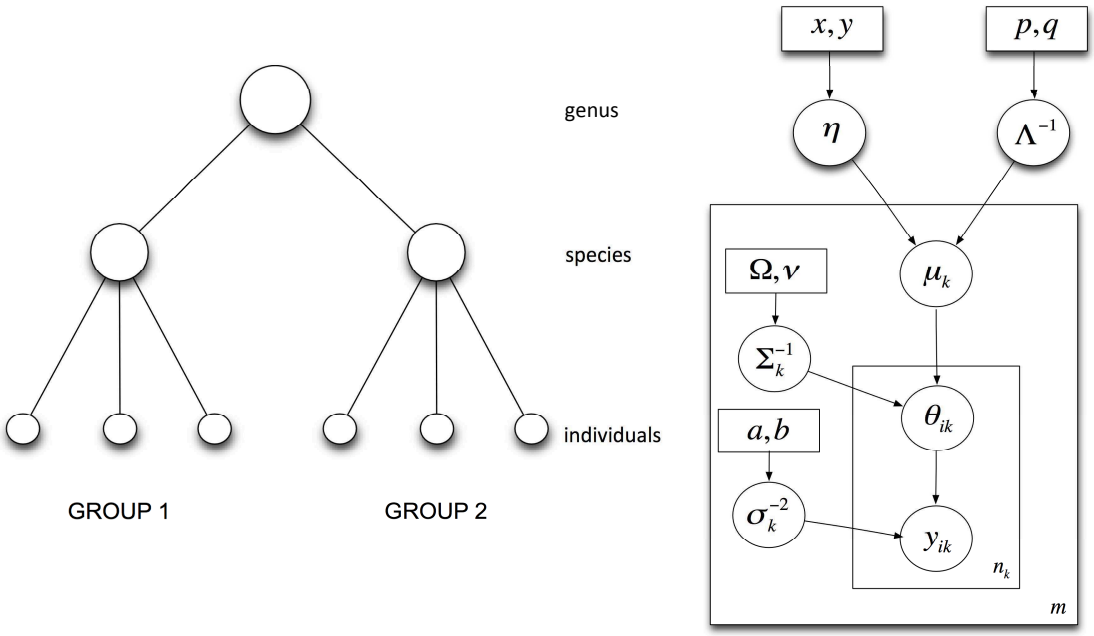


Figure 1. Left Panel: Multi-level hierarchical model with two groups. Each group has three individuals. Also shown are the genus, species and individual levels. Right Panel: Plate diagram for the multi-level hierarchical model. The plate denotes iteration of parameters and the number enclosed in the plate shows the number of iterations.

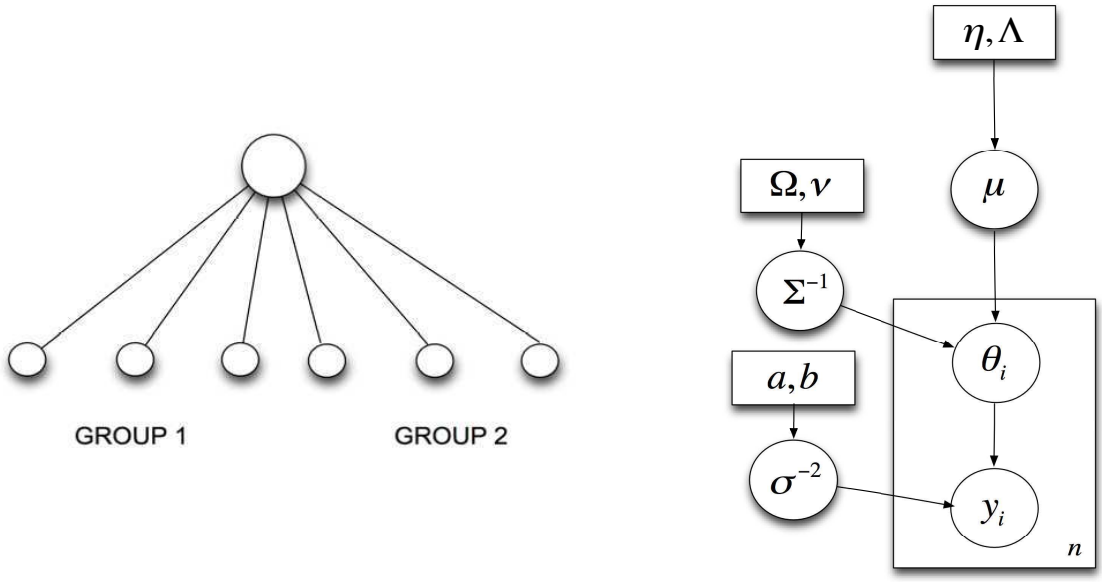
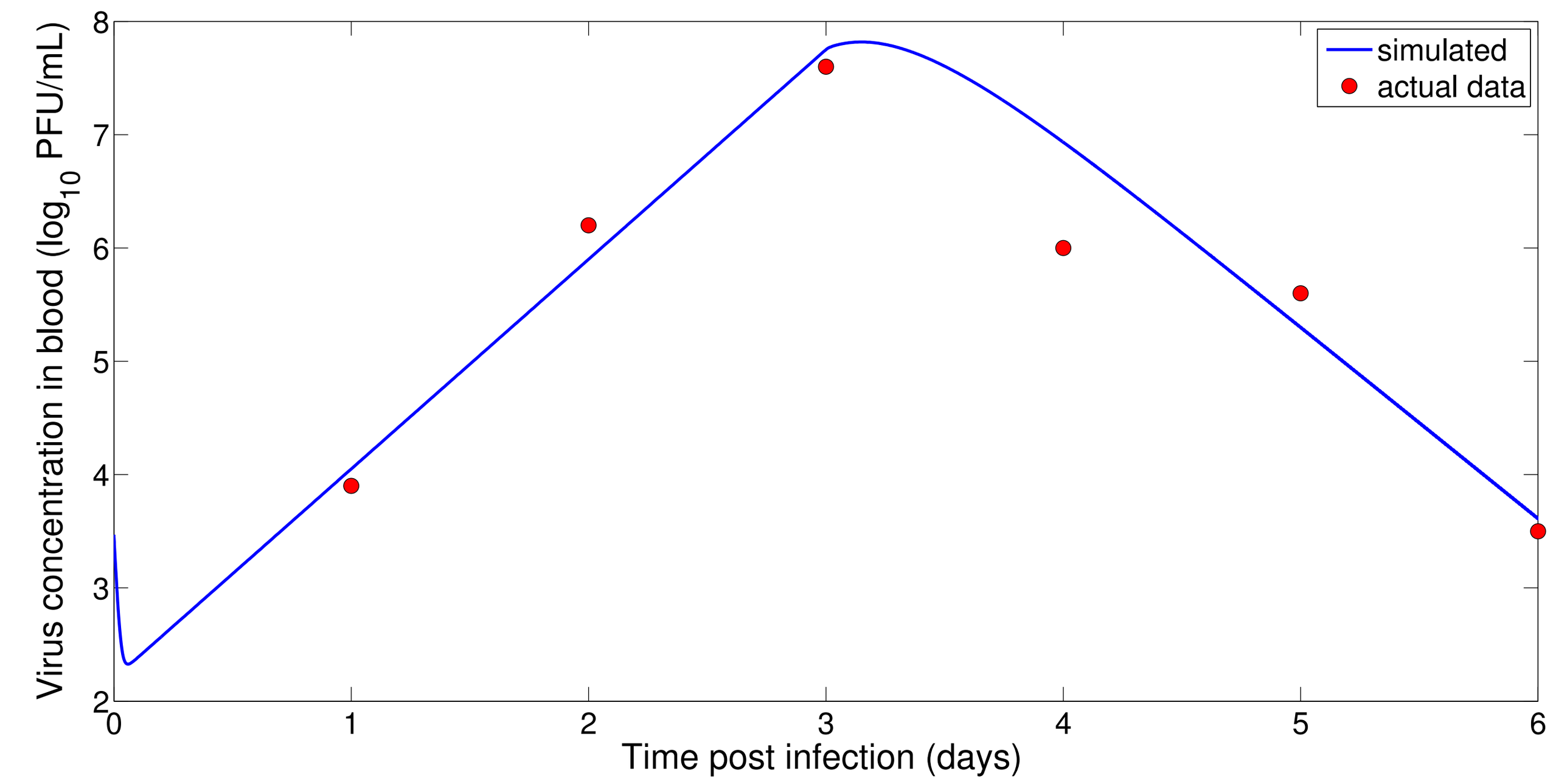


Figure 2. Left Panel: Aggregated model with two groups combined. Each group has three individuals. Also shown are the genus, species and individual levels. Right Panel: Plate diagram for the aggregated model. The plate denotes iteration of parameters and the number in the plate shows the number of iterations.


simulated
actual data
Virus concentration in blood (log10 PFU/mL)
Time post infection (days)
8
7
6
5
4
3
2
0
1
2
3
4
5
6