



Published in final edited form as:

*Comput Stat Data Anal.* 2009 January 15; 53(3): 766–775.

## Two-way Bayesian hierarchical phylogenetic models: An application to the co-evolution of gp120 and gp41 during and after enfuvirtide treatment

Christina M.R. Kitchen<sup>a,\*</sup>, Jing Kroll<sup>b</sup>, Daniel R. Kuritzkes<sup>b</sup>, Erik Bloomquist<sup>a</sup>, Steven G. Deeks<sup>c</sup>, and Marc A. Suchard<sup>a,d,e</sup>

<sup>a</sup> Department of Biostatistics, University of California, Los Angeles, CA, United States

<sup>b</sup> Section of Retroviral Therapeutics, Brigham and Women's Hospital and Division of AIDS, Harvard Medical School, Boston, MA, United States

<sup>c</sup> University of California, San Francisco General Hospital, San Francisco, CA, United States

<sup>d</sup> Department of Human Genetics, University of California, Los Angeles, CA, United States

<sup>e</sup> Department of Biomathematics, University of California, Los Angeles, CA, United States

### Abstract

Enfuvirtide (ENF) is a fusion inhibitor that prevents the entry of HIV virions into target cells. Studying the characteristics of viral evolution during treatment and after a treatment interruption can lend insight into the mechanisms of viral evolution and fitness. Although interruption of anti-HIV therapy often results in rapid emergence of an archived “wild-type” virus population, previous work from our group indicates that when only ENF is interrupted, viral gp41 continues to evolve forward and resistance mutations are lost due to back-mutation and remodeling of the envelope protein. To examine the co-evolution of gp120 and gp41 during ENF interruption we extend the Bayesian Hierarchical Phylogenetic model (HPM). Current HPMs enforce conditional independence across all outcomes while biologically all gene regions within a patient should return the same tree unless recombination confers an evolutionary selective advantage. A two-way-interaction HPM is proposed that provides middle ground between these two extremes and allows us to test for differences in evolutionary pressures across gene regions in multiple patients simultaneously. When the model is applied to a well-characterized cohort of HIV-infected patients interrupting ENF we find that across patients, the virus continued to evolve forward in both gene regions. Overall, the hypothesis of independence over dependence between the gene regions is supported. Models that allow for the examination of co-evolution over time will be increasingly important as more therapeutic classes are developed, each of which may impact other through novel and complex mechanisms.

### 1. Introduction

Studying the characteristics of viral evolution during and after exposure to antiretroviral drugs can provide insight into the relationship between drug resistance, viral fitness and disease progression. Enfuvirtide (ENF) is a fusion inhibitor that prevents the entry of HIV virions into human cells and is often used in the management of drug-resistant HIV. Previously our group conducted a study in which heavily treated patients who had exhibited virologic rebound on an enfuvirtide-based regimen underwent a “partial treatment interruption” (PTI) in which

\*Corresponding address: Department of Biostatistics, UCLA School of Public Health, 650 Charles E. Young Drive, Los Angeles, CA 90095-1772, United States. Tel.: +1 310 825 7332; fax: +1 310 276 2113., cr@ucla.edu (C.M.R. Kitchen).

enfuvirtide was interrupted while the other drugs were maintained (Deeks et al., 2006). All patients had documented ENF resistance at the time of interruption. Genotypic resistance waned 8 weeks after the interruption and by 16 weeks was undetectable in a clonal analysis in the majority of patients. An unexpected finding from this study was that during the interruption period, the loss of ENF-resistance mutations in gp41 was not due to the re-emergence of archived strains circulating before ENF initiation, but rather due back mutation of the resistance mutations and continued forward evolution in gp41 (Kitchen et al., 2006). It remains unknown how ENF therapy and its interruption affects the co-evolution of other gene regions involved in viral entry, such as gp120 which encodes for coreceptor usage. This study allows us a unique opportunity to examine the co-evolution of multiple gene regions during a short-term episode of selective pressure directed at one of gene regions.

HIV gains entry into the cell in a multi-step process that is mediated via the viral envelope proteins (gp41 and gp120) (Wyatt and Sodroski, 1998). HIV first binds to the CD4 receptor via a non-covalent binding reaction with gp120. This interaction initiates a conformational change within gp120, exposing the chemokine receptor binding sites. Once HIV binds to these receptors (either CCR5 or CXCR4), the viral gp41 undergoes a conformational change that results eventually in the fusion of the virus with the cell membrane. ENF binds to HR-1 region of gp41, thus preventing the latter step. This complex mechanism for cell entry involves a complex and poorly understood interaction between gp120 and gp41. It also remains unclear if coreceptor usage has an effect on ENF resistance (Melby et al., 2006; Morse and Maldarelli, 2007) or vice versa.

To aid in understanding the co-evolution of both gene regions in the presence and absence of directed selective pressure we extended our Bayesian hierarchical phylogenetic model (HPM) (Suchard et al., 2003; Kitchen et al., 2004) to allow for the simultaneous estimation of multiple gene regions across multiple patients and time points. This method was specifically developed to test the hypothesis of evolutionary interactions between gene regions. Biologically, all gene regions within a single patient should return the same phylogenetic tree unless recombination or differing selective pressures confers an evolutionary selective advantage. Previous HPM construction enforced conditional independence across all tree outcomes and were therefore not designed to test for co-evolution. To address this limitation, we propose a two-way interaction HPM that provides a middle ground between these two extremes. We provide novel distributions over tree topologies to allow these topologies to differ across genes within a patient and test for an evolutionary interaction. We accomplish this via model selection techniques involving a log linear prior over the tree topology outcomes. The hierarchical framework also allows for the pooling of statistical information across patients and gene regions leading to more efficient use of data to improve estimation of individual parameters while simultaneously estimating parameters across all patients. Pooling is especially useful if one of the gene regions is poorly estimated due to small number of sites sampled (sequence length) or lack of phylogenetic information (variation). Here we describe this novel statistical method and apply it to our cohort of patients. Our data suggest that there is continued evolution in both gene regions after ENF interruption and that this forward evolution occurs independently. Although prior observations had shown that ENF resistance wanes rapidly in absence of ENF (Deeks et al., 2006; Kitchen et al., 2006), this is most likely due to back-mutation of specific resistance mutations and continued forward evolution in rest of the gene region gp41. Further, the dominant quasi-species arising during ENF therapy persisting even after ENF interruption suggests that this virus is relatively more fit than the presumed archived ENF-susceptible virus. Taken together, these observations suggest that there exists a selective pressure greater than ENF that is driving evolution forward and that these pressures act on gp41 independent of its effect on gp120.

## 2. Materials and methods

### Study population

Details of the study population are described elsewhere (Deeks et al., 2006; Kitchen et al., 2006). Briefly, subjects were enrolled in a prospective study of HIV-1 infected patients who had exhibited incomplete viral suppression on an enfuvirtide-based regimen and who had developed genotypic evidence of enfuvirtide resistance. Subsequently, patients interrupted only the ENF portion of their regimen while remaining on an optimized background antiretroviral drugs. As previously reported (Deeks et al., 2006), genotypic resistance to ENF waned in the absence of drug pressure. The dominant virus population was enfuvirtide susceptible by 16 weeks off therapy. For the present analysis, we identified subjects for whom we had clonal sequence data on gp41 and gp120 available prior to the initiation of enfuvirtide (“pre-ENF”), during ENF failure (“ENF”) and approximately 16 weeks post ENF interruption (“post-ENF”). There were seven subjects who had all three time points available for both gp41 and gp120. All subjects provided informed consent for participation in these studies.

### PCR amplification, cloning, and sequencing

Clonal analysis details are presented elsewhere (Lu et al., 2006). Briefly, viral RNA was extracted from plasma using the QIAamp viral RNA kit (Qiagen, Valencia, CA). A 650bp fragment of gp41-that encompassed the HR-1 and HR-2 regions and the C2-V3-C3 region of gp120 was amplified by nested RT-PCR PCR using QIAgen One step RT-PCR kit with forward primer gp41OS, 5' GAG GGA CAA TTG GAG AAG TGA ATT 3' and reverse primer gp41OA 5' GTG AGT ATC CCT GCC TAA CTC TAT 3' in the RT-PCR and forward primer gp41IS 5' GGA GAA GTG AAT TAT ATA AAT ATA AAG 3' and reverse primer GP41A1 5' TTA AAC CTA CCA AGC CTC C 3' in the second-round PCR. Reaction conditions were as reported (Lu et al., 2006). To minimize sampling bias, first round RT-PCR products were performed in quadruplicate, except for samples with plasma HIV-1 RNA less than 1000 copies/mL. Five microliters of each quadruplicate first-round PCR product were pooled and 3  $\mu$ L of the mixture was then carried over to quadruplicate second-round PCR. To obtain individual clones of the HR-1 and HR-2 coding region for sequence analysis, 1  $\mu$ L of pooled second-round PCR product was cloned into plasmid vector pCR-2.1 (Stratagene, La Jolla, CA) using the TOPO TA cloning system (Invitrogen, CA). A minimum of 8 clones were picked and subjected to bidirectional DNA sequencing of the purified PCR products using Applied Biosystems Taq DyeDeoxy Terminator cycle sequencing kits (Applied Biosystems, Foster City, CA). The DNA sequences were edited and translated using BioEdit. Note that the different gene regions were amplified from independent PCR replicates.

### Recombination detection

All gene sequences were tested for evidence of intragenic recombination using a Bayesian dual multiple change-point model (Minin et al., 2005). No evidence of recombination was found within each gene region. We used the two-way Bayesian HPM described below to test for possible evidence of recombination between gene regions.

### Phylogenetic analysis

The gp41 and gp120 sequences for each patient were aligned separately using Clustal w and manually checked and edited. A Blast search was conducted to ensure that the sequences were not contaminants. Clones were shown to be unique by sequence analysis. Overall, the median sequence genetic distance was 2.2% (range: 0.50%–4.30%). We extend our previous Bayesian hierarchical phylogenetic model (Suchard et al., 2003; Kitchen et al., 2004) to allow information from all patients and multiple gene regions to be analyzed concurrently and to test for an evolutionary interaction between gene regions.

To reconstruct the evolutionary trees we modeled the probability of nucleotide mutations over time via a continuous time Markov chain model using an infinitesimal rate matrix for nucleotide substitution similar to the Tamura-Nei matrix (Tamura and Nei, 1993) and site-to-site rate variation modeled with a discrete gamma distribution. For each patient  $i = 1, \dots, N$  and gene region  $j = 1, \dots, J$ , we estimate the unknown tree topology,  $\tau_{ij}$ , relating the gene variants from within a patient at the pre-ENF, ENF and post-ENF time points, the tree branch lengths,  $t_{ij}$ , the transition/transversion rate ratio between the purines,  $k_{1ij}$ , the transition/transversion rate between the pyrimidines,  $k_{2ij}$ , the prior expected divergence,  $\mu_{ij}$ , the stationary nucleotide frequencies,  $\pi_{ij}$ , and the rate variation shape parameter  $\alpha_{ij}$ . We assume that patient-specific gene-specific parameters draw from common hierarchical distributions. These distributions are characterized by unknown across-patient and across-gene parameters upon which we make “population-level” inference. We fit the population level and patient-level models simultaneously by using hierarchical priors so that the population-level feeds back information to the individual-level parameters to improve estimation. For the continuous patient/gene-level parameters  $(\alpha_{ij}, \kappa_{1ij}, \kappa_{2ij}, \mu_{ij}, \pi_{ij})$  we follow current HPM practice in defining the population level distributions. Specifically we employ the following hierarchical priors:

$$(\alpha_{ij}, \kappa_{1ij}, \kappa_{2ij}, \mu_{ij}) \sim \text{logNormal}(M, \Sigma), \quad \text{and} \quad \pi_{ij} \sim \text{Dirichlet}(N_{\Pi} \times \Pi)$$

where  $\Sigma = \text{diag}(\sigma_{\alpha}^2, \sigma_{\kappa_1}^2, \sigma_{\kappa_2}^2, \sigma_{\mu}^2), \sigma_{\alpha}^2, \sigma_{\kappa_1}^2, \sigma_{\kappa_2}^2, \sigma_{\mu}^2$  are distributed half-Cauchy,  $M$  is a diffuse Normal and  $\Pi$  is a flat Dirichlet. To interpret these parameters  $M$  and  $\Pi$  are the population-level means and  $\Sigma$  and  $N_{\Pi}$  measure the variability across patients and genes. We require and provide novel distributions over  $\tau_{ij}$  to allow topologies to differ across genes within a patient and test for an evolutionary interaction.

Alignments for each gene region consisted of four sequences: one from each of the three time points (pre-ENF, ENF, and post-ENF) and HXB2 as the outgroup. By organizing the data structure this way we can have 3 distinct and testable evolutionary histories, as with 4 taxa there are 3 possible trees. If the viral variant that was isolated approximately 16 weeks post ENF interruption (post-ENF strain) is closer to the strain isolated during ENF failure (ENF strain) then the hypothesis of continued evolution is supported (Tree E), suggesting that the virus continued to evolve in that gene region. If the post-ENF strain is more closely related to the strain isolated before ENF therapy was initiated (pre-ENF), then the re-emergence hypothesis is supported (Tree R). The third possibility is the post-ENF strain is most closely related to the outgroup (Tree O) and is unlikely a priori. For the gp41 gene region, support for the continued evolution hypothesis suggests that the loss of ENF resistance associated mutations is most likely due to back-mutation, while support of the re-emergence hypothesis suggests that the loss of resistance is most likely due to the re-emergence of the presumed archived ENF-susceptible strains that were circulating before ENF initiation. Theoretically, each of these three outcomes is possible for each gene region within each patient. The original HPM construction then enforces conditional independence across all outcomes. Biologically, all gene regions within a patient should return the same tree if recombination between the regions does not confer an evolutionary selective advantage. We propose a two-way interaction model that finds a middle ground between these extremes and enables us to test for a difference between evolutionary pressures across gene regions in multiple patients simultaneously. In this way, we take as our null hypothesis the evolutionary model in which gene regions share, through a high degree of correlation, the same tree outcomes and attempt to show differing processes acting during ENF interruption. We accomplish this via model selection techniques involving a log-linear prior (Albert, 1996) on tree outcomes.

Traditional log linear models explore main effects and interactions in count data that form contingency tables. In a sense, testing for interactions generalizes the  $\chi^2$  test of independence that ask if the strata that define one dimension in the table help predict the strata in another dimension in which each datum falls. To test for gene region interactions in intra-patient evolution, we build a contingency table with  $R = 3$  rows where each row indicates one of three possible evolutionary histories and  $C = 2$  columns, one for each gene region. Suppose that we know what the true reconstructed tree for each region is for every patient; then the cells in this table  $y_{rc}$  for  $r = 1, 2, 3$  and  $c = 1, 2$  count the number of trees that fall into category  $(r, c)$  summed over all patients. In the log-linear framework one formulates the data likelihood through assuming  $y_{rc}$  are known and distributed as conditionally independent Poisson random variables with expected counts  $m_{rc}$ . In the evolutionary setting,  $y_{rc}$  are not known; rather they are estimated through the sequence data via phylogenetic reconstruction with a twist. Instead of assuming that the trees  $\tau_{ij}$  are all independent across patients and regions, we place a log-linear motivated prior (Albert, 1996) on them through their summary statistics  $y_{rc} = \sum_i 1\{\tau_{ij} = r\}$ .

The model with the most dependence structure is called the saturated model in which  $\log m_{rc} = u_0 + u_{1(r)} + u_{2(c)} + u_{12(rc)}$ ,  $r = 1, \dots, R - 1$ ,  $c = 1, \dots, C - 1$ , where  $u_0$  is a constant term,  $u_{1(r)}$  is the row effect,  $u_{2(c)}$  is the column effect and  $u_{12(rc)}$  are the interaction parameters. A model that reflects independence is  $u_{12(rc)} = 0$  for all  $r$  and  $c$ . Between the saturated and independence models exist a continuum of nested models. To test for independence, we need to compare the relative likelihoods of these models which we accomplish using Bayesian model selection (Albert, 1996). For construction of the prior, let  $u$  be a  $px1$  vector of log-linear parameters and then partition  $u$  into two parts so that  $u = (\beta_0, \beta_1, \dots, \beta_s)$  where  $\beta_0$  includes the parameters that are assumed to be 0 in a given model and the elements of  $\beta_1, \dots, \beta_s$  may be non-zero. Let  $p_0$  equal the dimension of  $\beta_0$  and let  $p_l$  be the dimension of  $\beta_l$  ( $l = 1, \dots, s$ ). In terms of hypothesis testing, a test of independence ( $H_I$ ) is equivalent to testing that all elements of  $\beta_l$ , for  $l > 0$ , are equal to zero for all  $l$  versus the alternative dependent hypothesis ( $H_D$ ) that at least one of the terms is not equal to zero. In our prior specification, we impose the corner point constraints on the marginal parameter vectors  $u_1$  and  $u_2$  by which the first component of the vector is set equal to zero. The prior on  $u$  is a multivariate Normal with mean 0 and variance-covariance matrix  $P^{-1}$ . The precision matrix  $P$  has the structure:

$$P = \begin{bmatrix} 0I_{p_0} & 0 & \dots & 0 \\ 0 & P_1I_{p_1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_sI_{p_s} \end{bmatrix},$$

where  $I_p$  is the identity matrix with dimension  $p$ . The  $\beta_l$ 's are distributed Normal with mean 0 and precision  $P_lI_{p_l}$  which represents the prior uncertainty that  $\beta_l$  is close to zero. For example, if the hypothesis of independence is true then the prior mean of  $\beta_l$  should be equal to zero and under the dependence hypothesis, one believes that there are elements of  $\beta_l$  that are not equal to zero and this is reflected in the precision prior for  $\beta_l$ . Because the size of the Bayes Factor can be sensitive to the value of  $P_l$ , we place a hyperprior distribution over  $P_l$ . The components of  $\beta_0$  have precision equal to 0 and  $P_l$  for  $l > 0$  is distributed as Gamma ( $v_l/2 + \beta_l/2$ ,

$b_l^2 v_l/2 + \sum_{t=1}^{p_l} \beta_{lt}^2/2$ ) where  $\beta_l = (\beta_{l1}, \dots, \beta_{lp_l})$  are the regression parameters that are not assumed to be equal to 0. We set  $v_1 = \dots = v_s = 1$  and  $b_1 = \dots = b_s = 2$ , the number of non-zero finite precision parameters.

Sub-models in this framework are statements about the significance of all  $s$  subsets of  $\beta_l$ . For each of the  $s$  hypotheses, we denote them as models  $M_1, \dots, M_K$ . Since each model is either

true or false there are  $K = 2^s$  models. Model selection was achieved by estimating the posterior probabilities  $\Pr(M_k | y)$  using Gibbs sampling. Let  $Z = (Z_1, Z_2)$  be the model indicator. Let  $Z_l = 0$  if the  $l$ th subset is equal to zero (the independence hypothesis,  $H_I$ ) and let  $Z_l = 1$  if the  $l$ th subset is not equal to zero (the dependence model  $H_D$ ). Further, let  $P^* = (P_1^*, \dots, P_s^*)$  be the vector of finite precision parameters  $P_l$  when hypothesis  $H_D$  is true (when hypothesis  $H_I$  is true  $P_l$  equals infinity). To understand this prior, recall that the mean vector is 0 and thus the value of  $P_l$  reflects the prior belief that  $\beta_l$  is close to zero. We used the following Gibbs Sampling method:

1. Start with an initial value of  $Z$  and precision vector  $P$  and simulate values of  $\{\beta_l\}$ ,  $Z$  and  $P^*$
2. If  $Z =$  then
  - a. Simulate  $\beta_l$  from  $N(\tilde{\beta}_l, (X_0'WX_0 + \sum_0^{-1})^{-1})$  where  $\tilde{\beta}_l$  denotes the mode
  - b. Simulate  $P_j^*$  from  $\text{Gamma}(v_l/2 + \beta_l/2, b_l^2 v_l/2 + \sum_{i=1}^{P_l} \beta_{li}^2/2)$ .
3. Simulate with probability proportional to  $\Pr(M_D | y, P)$  and  $\Pr(M_I | y, P)$  where  $M_D$  and  $M_I$  have precision values of  $P_l = P_l^*$  and  $P_l = \infty$ , respectively.

Note that the same non-informative prior is assigned for  $\beta_0$  under both the independence and dependence hypotheses such that the normalizing constant cancels out during the computation of the Bayes Factor for the test of independence ( $H_I$ ) versus dependence ( $H_D$ ). Bayes Factors can be thought of as the level of support for one model versus another given the data. In the case where there is a single hypothesis of interest, independence, ( $M_I$ ), versus dependence, ( $M_D$ ), one can calculate the Bayes Factor as the posterior odds ratio:  $\Pr(M_I | y)/\Pr(M_D | y)$  as we assume each model has equal prior probability.

Models were fitted using Markov chain Monte Carlo (MCMC) with chain lengths of 350,000 iterations. We discarded the first 50,000 observations as burn-in and subsampled every 30th observation to yield 10,000 posterior samples with decreased autocorrelation. The models ran in approximately 5 hrs on a 1-GHz Pentium M PC. Examination of model log-likelihood time series plots showed that the burn-in times and the total MCMC chain lengths were significantly longer than what appeared to be required (Cowles and Carlin, 1996). Convergence was also assessed via the scale reduction statistic of Gelman and Rubin (1992).

### Nucleotide sequence accession number

The sequences of the cloned *gp41* have been deposited in Genbank under the accession numbers DQ852666-DQ852692.

## 3. Results

### 3.1. Subject characteristics

Seven subjects met our inclusion criteria. At study entry the median CD4+ T-cell count was 10 cells/mm<sup>3</sup> (range: 8–140) and median HIV-1 RNA plasma viral burden was 5.03 log<sub>10</sub> copies/mL (range: 4.45–5.44). Approximately 16 weeks post ENF interruption the median CD4+ T-cell count was 34 cells/mm<sup>3</sup> (range: 7–159) and the median HIV-1 plasma burden was 5.49 log<sub>10</sub> copies/mL (range: 4.58–5.71). Before ENF interruption all patients had genotypic evidence of ENF resistance. Approximately 16 weeks post ENF interruption, resistance-associated mutations in HR1 had waned. For each patient timepoint, coreceptor usage prediction was assessed using the PSSM algorithm (Jensen et al., 2003), the 11/25 algorithm (DeJong et al., 1992; Fouchier et al., 1995) and from the net charge of the V3 loop

(Fouchier et al., 1995; Briggs et al., 2000). Coreceptor usage for each patient was taken as the majority consensus of these 3 algorithms. In this study, 3/7 patients had predicted CXCR4 usage prior to enfuvirtide treatment that persisted throughout the study and 2/7 patients had predicted CCR5 usage during the entire study. Two of seven patients had predicted CCR5 usage at baseline but developed predicted CXCR4 usage during and after ENF treatment (Table 1).

### 3.2. Evolution of gp41 and gp120

An average of 8 clones (range: 6 to 14) for gp41 and an average of 7 clones (range 4–9) clones of the C2-V3-C3 region of gp120 were available for each time point for each patient. It is important to note that the gp41 region and the gp120 region were cloned and sequenced separately. Sequences within each gene region were tested for recombination using a Bayesian dual change point model (Minin et al., 2005). We did not find evidence of within-gene region recombination. Because of potential bias created by convergent evolution, we deleted the resistance encoding sites in gp41 and constructed maximum likelihood trees for each patient and each gene region. Within each patient gene region, the sequences clustered by time (preENF, ENF and post-ENF). Fig. 1 illustrates this time clustering for a representative patient in (a) gp41 and (b) gp120. We randomly selected clones from each patient-time-point and gene region and aligned them using HXB2 as the outgroup for analysis via the two-way HPM. Because of possibility of randomly selecting a non-representative clone, we repeated this procedure.

Using the two-way HPM, we found overwhelming support for the evolution hypothesis across patients and gene regions confirming our previous findings in gp41 (Kitchen et al., 2006). In the individual estimates of gp41, 5/7 patients supported the Evolution hypothesis (Tree E) with at least 0.95 probability. In gp120, 5/7 patients supported the Evolution hypothesis (Tree E) with at least 0.99 probability (Table 2). In both gene regions there was a difference in the transition/transversion rate ratio between the purines and the pyrimidines with the ratio between the purines higher than the ratio between the pyrimidines ( $\kappa_1$  and  $\kappa_2$ , respectively) indicating the necessity of that parameterization. In comparison to gp41, gp120 had a smaller  $\alpha$  parameter suggesting greater site-to-site variation and a higher mutation rate  $\mu$  ( $p = 0.013$  for both using a two-sided Wilcoxon Rank Sum test). Table 3 shows the population level estimates. At the population level we found overwhelming support for the Evolution hypothesis over the Re-emergence hypothesis (Tree E/Tree R) with a Bayes Factor of 1428, see Table 3. This suggests that Evolution hypothesis is over 1000 times more likely than the Re-emergence hypothesis, which in terms of  $p$ -values suggests a  $p < 0.001$ . The population estimates also show support for our parameterization with a significant difference between  $\kappa_1$  and  $\kappa_2$  as well as an  $\alpha$  parameter less than 1.

Analysis of the gp41 region alone yielded a Bayes Factor of 30.2 of the Evolution hypothesis over Re-emergence hypothesis (Tree E/Tree R) suggesting that pooling data across genes and patients allowed us to strengthen our conclusions in terms of a more decisive Bayes Factor and more precise estimation of the evolutionary parameters. Estimates of the transition/transversion rate ratios between the purines and pyrimidines ( $\kappa_1$  and  $\kappa_2$ , respectively) as well as the  $\alpha$  rate variation parameter have smaller standard errors. The  $\alpha$  parameter being less than 1 implies that there exists rate variation across gene regions and thus the estimation of  $\alpha$  gives one confidence that the inferred tree is not biased by ignoring the rate variation among sites in the differing gene regions.

Using our novel two-way HPM, we found that the hypothesis of independence ( $H_I$ ) was supported over the dependence hypothesis ( $H_D$ ) with a Bayes Factor equal to 3. In this framework we test the null hypothesis that the two gene regions do not differ with respect to the relative frequency of tree support, that is the null hypothesis is that there is no evolutionary

interaction. Thus, strong disagreement in trees is evidence for the dependence hypothesis while strong agreement for the same tree supports the independence hypothesis. If there exists differing selective pressures acting upon the gene regions, we would expect there to be support for different trees, that is, we would expect an evolutionary interaction ( $H_D$ ). Support for differing trees could also come about via recombination between gp41 and gp120, thus support for  $H_D$  is also consistent with a recombination event. The data yielded evidence in favor of the independence hypothesis suggesting that the loss of drug-resistance associated mutations in gp41 was consistent with back-mutation in gp41. Further, the lack of support for an evolutionary interaction during ENF interruption suggests that there exists similar selective pressures on gp41 and gp120 that is forcing the evolution forward in both of these regions. This unmeasured pressure is greater than the selective pressure put forth by enfuvirtide as evidenced by the fact that the dominant quasi-species circulating during ENF therapy persists even after ENF interruption. This also suggests that this virus is relatively more fit than the presumed archived ENF-susceptible virus. These findings may reflect continued immune pressure in *env* even in advanced HIV disease.

### 3.3. Simulations

To assess convergence we used the scale reduction statistic suggested by Gelman and Rubin (1992). We conducted 10 independent simulations of our model to estimate the scale reduction statistic for our values of  $\beta$ . The scale reduction statistics ranged from 1–1.1 suggesting that we had achieved convergence (data not shown). Fig. 2 shows the distributions of the 2 interaction variables  $\beta_1$  and  $\beta_2$ . Note that both interaction terms are centered at zero indicating support for the independence hypothesis. To assess the properties of our Bayes Factor in testing for evolutionary interactions we conducted several simulations of our model. We simulated sequences using Seq-Gen (Rambaut and Grasly, 1997) to create simulated  $3 \times 2$  tables using 4 taxa with a sample size of 7 subjects and 2 genes. We created matrices that had independent gene regions and we created matrices that had dependence between the gene regions. We ran 500 independent simulations of size 10,000. For the simulations we defined the Bayes Factor as the probability of the Dependence model given the data over the probability of the Independence model given the data ( $\Pr(M_D | y) / \Pr(M_I | y)$ ). Each simulation generated random matrices of size  $3 \times 2$ , for each simulation the Bayes Factor and the Chi Square statistic was computed. In 500 simulations of independent sequence matrices, our model supported the independence hypothesis 473 out of 500 times. In 500 simulations from dependence sequence matrices, our model supported the dependence hypothesis 489 out of 500 times. Fig. 3 shows the level of agreement between the log of the Bayes Factor versus the Chi square statistic. Note that the way the Bayes Factor is defined for the simulation, values greater than 1 suggest support for the Dependence hypothesis and values less than 1 show support for the Independence hypothesis. Of course, the Chi square statistic ignores the uncertainty in the underlying cell counts. These simulations show that the model performs well and can detect departures from independence even in this relatively small sample size.

### 3.4. Sensitivity analysis

To assess the sensitivity of our Bayes factors for the test of Independence over Dependence to our priors we conducted a sensitivity analysis. As stated previously,  $\beta_1$ 's are distributed Normal with mean 0 and precision  $P_I I_{pI}$  and we place a hyperprior distribution over  $P_I$ . We conduct sensitivity analyses on the hyperpriors  $v$  and  $b$ . The Bayes factor was robust for values of  $v$  and  $b$  under which the precision ranges between 0.22 and 6. In this range, the Bayes factor varied between 2.1 and 3.46. Hyperparameter values outside of this range resulted in much larger (or smaller) Bayes factors. The range found in the sensitivity analysis is in agreement with results obtained in Raftery (1993), who suggested that each interaction term should be modeled by a Normal with mean 0 and precision,  $P$ , where  $RC/25 < P < RC$ ; where  $R$  = the number of rows and  $C$  = the number of columns so that the functional form will have minimal impact on the

Bayes factor. This result is also in agreement with the findings in Albert (1996). The values of  $v$  and  $b$  chosen for the analysis result in precisions that are within the suggested functional form and result in robust Bayes factors.

#### 4. Discussion

We describe an extension to our Bayesian hierarchical phylogenetic model that incorporates novel distributions over tree topologies to allow for the testing of evolutionary interactions between gene regions. This model can also be useful for outlier detection. In our current analysis we used this novel model to study the co-evolution of two gene regions during and after exposure to a potent selective pressure which targets a focused region of the HIV envelope (i.e. the HR-1 region of gp41), and found that the viral population continued to evolve forward after the removal of this selective pressure. This is in contrast to other interruption studies in which the re-emergence of an archived drug-susceptible virus emerged rapidly in absence of the selective pressure (Miller et al., 2000; Kijak et al., 2002; Deeks et al., 2003; Ruiz et al., 2003; Boucher et al., 2005). In our previous work (Kitchen et al., 2006), we focused specifically at changes in gp41 and did not examine changes in other regions. By extending our model we were able to address the co-evolution of gp120 and gp41, regions intimately involved in receptor binding. Using this model we found decisive support for continuing evolution in both gene regions with a Bayes Factor of 1428 in favor of the evolution hypothesis over the re-emergence hypothesis. More importantly, our data indicate that as gp41 remodeled itself after the removal of drug pressures, the closely related gp120 molecule continued to evolve as well without strong evidence of an evolutionary interaction between these two regions.

Analysis of the gp41 region alone yielded a Bayes Factor of 30.2 of evolution over re-emergence suggesting that pooling data across genes and patients allowed us to strengthen our conclusions in terms of a more decisive Bayes Factor and more precise estimation of the evolutionary parameters. Estimates of the transition/transversion rate ratios between the purines and pyrimidines ( $\kappa_1$  and  $\kappa_2$ , respectively) as well as the  $\alpha$  rate variation parameter have smaller standard errors. We found significant differences in the  $\alpha$  and  $\mu$  parameters between gene regions. Not surprisingly, these parameters showed greater levels of site-to-site variation and mutation in gp120 versus gp41. These differences highlight the necessity of our model implementation and argue against concatenation of the gene regions as the simultaneous estimation of parameters by concatenation force evolutionary pressures across gene regions to be the same, an assumption clearly not supported by the data.

Biologically within a single patient, we expect that in the absence of recombination or differing selective pressures all gene regions would support the same tree. Strong disagreement in trees indicates support for the dependence hypothesis, i.e. an evolutionary interaction while strong agreement in trees favors the independence hypothesis. In this cohort, we found that the level of support for the independence hypothesis was three times greater than the support for the dependence hypothesis ( $BF = 3$ ). We conducted simulation studies and found that even in this small sample size we were able to detect departures from independence. The totality of the evidence points to the fact that there exists a similar selective pressure against both gene regions that is forcing their continued evolution. Further, the fact that the dominating viral swarm during ENF treatment persisted even after ENF interruption suggests that the older archived strains were not sufficiently fit to emerge in the host environment and most likely reflects unmeasured immune pressure against *env*.

Labrosse et al. (2006) examined *env* replicative capacity during ENF treatment. They found that that ENF resistance did not result in a significant reduction in *env* replicative capacity for 5/6 patients studied. This result is consistent with our findings. The study also examined longitudinal sequences of gp120 and gp41 in two patients who were taking ENF and then

interrupted ENF. Although Labrosse et al. (2006) did not specifically test this hypothesis, they found that during ENF treatment both the gp120 and gp41 gene region had changes in the dominant sequences in the plasma population and that were sustained after ENF interruption. Ray et al. (2007) examined the coreceptor usage pattern in five patients who had virologic failure on ENF. Similar to Labrosse et al., Ray et al. used phylogenetic branching patterns and examination of mutations in the alignment to assess evolutionary patterns. They conclude that the branching patterns seen by constructing phylogenetic trees suggests that the R5 and X4 lineages that harbored ENF-resistance evolved independently. All of these results are supported by our data. However, neither of these papers based their arguments on a statistical model, thus they were not able to make meaningful probability statements about their conclusion essentially ignoring any uncertainty in their findings. These studies emphasize that there is an existing need for novel phylogenetic methods with which researchers can statistically address evolutionary questions.

Initiation of ENF did not appear to cause a shift in coreceptor utilization. Two patients experienced a switch from CCR5 utilizing strain to a CXCR4 utilizing strain from Pre-ENF to ENF time points. There was no change in coreceptor utilization after ENF initiation. The support for the independence hypothesis suggests that this result is most likely not due to recombination. It also suggests that ENF resistance is not the dominant selective pressure and thus ENF resistance is unlikely to affect the susceptibility to coreceptor inhibitors. Ray et al. (2007) explicitly test this hypothesis biologically and found that ENF resistance did not affect susceptibility to coreceptor inhibitors or other fusion inhibitors.

For both gene regions we found 5/7 patients supported the Evolution Tree with at least 95% support. Only 1 patient, patient 3514, did not support the Evolution hypothesis in either gene region. In gp41 patient 3514 only had 3% support for Tree E and 72% support for Tree E in gp120. This patient was an outlier. Closer examination showed that this patient missed over 50% of their ENF doses during the ENF treatment phase. Inclusion of this patient did not alter our overall conclusions and showed the utility of our model even with outliers.

Understanding viral evolution during therapy is critical to our understanding of viral pathogenesis and optimal treatment. It is especially important to understand how therapies that target specific regions of envelope may affect other regions as well. Models that examine co-evolution over time will be increasingly important as more classes of HIV therapeutics are developed, each of which may impact the other through novel and complex mechanisms.

## Acknowledgments

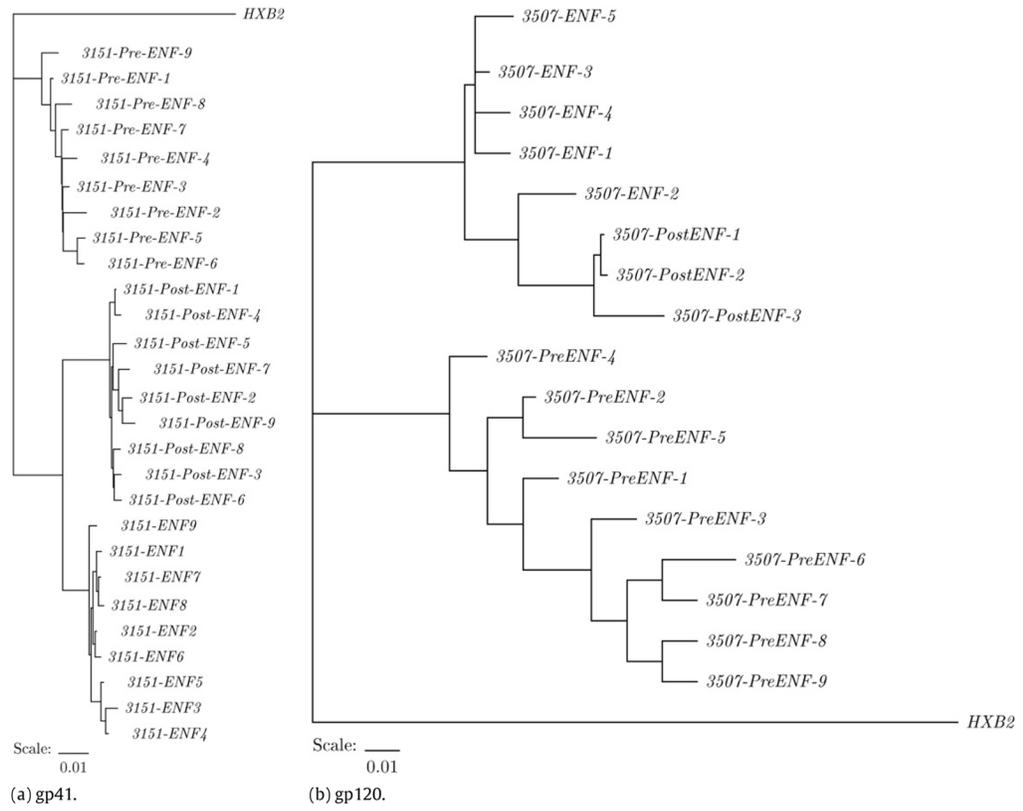
This work was partially supported in part by grants from the NIAID (AI058845-02S1, AI052745, AI055273, AI055357, NCRR RR 16482), the California AIDS Research Center (CC99-SF, ID01-SF-049), the UCSF/Gladstone Institute of Virology & Immunology Center for AIDS Research (P30 MH59037), the Harvard Medical School Center for AIDS Research (P30 AI 60354), the Center for AIDS Prevention Studies (P30 MH62246), the General Clinical Research Center at San Francisco General Hospital (5-MO1-RR0083-37), the UCLA AIDS Institute and the James P. Pendelton Foundation.

## References

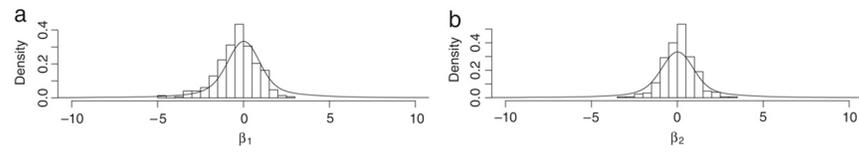
- Albert J. Bayesian selection of log-linear models. *Canadian Journal of Statistics* 1996;24:327–347.
- Boucher S, Recordon-Pinson P, Neau D, Ragnaud J, Titier K, Faure M, Fleury H, Masquelier B. Clonal analysis of HIV-1 variants in proviral DNA during treatment interruption in patients with multiple therapy failures. *Journal of Clinical Virology* 2005;34:288–294. [PubMed: 16286052]
- Briggs D, Tuttle D, Sleasman JW, Goodenow M. Envelope V3 amino acid sequence predicts HIV-1 phenotype (co-receptor usage and tropism for macrophages). *AIDS* 2000;14:2937–2939. [PubMed: 11153675]

- Cowles MK, Carlin BP. Markov chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association* 1996;91:883–904.
- Deeks SG, Grant RM, Wrin T, Paxinos EE, Leigler TJ, Hoh R, Martin JN, Petropoulos CJ. Persistence of drug-resistant HIV-1 after structured treatment interruption and its impact on treatment response. *AIDS* 2003;17:361–370. [PubMed: 12556690]
- Deeks SG, Lu J, Hoh R, Neilands TB, Beatty G, Huang W, Liegler T, Hunt P, Martin JN, Kuritzkes DR. Interruption of Enfuvirtide in HIV-1 infected adults with incomplete viral suppression on an Enfuvirtide-based regimen. *Journal of Infectious Diseases* 2006;195 (3):387–391. [PubMed: 17205477]
- DeJong J, Keulen DAW, Tersmette M, Goudsmit J. Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium inducing phenotype: Analysis by a single amino acid substitution. *Journal of Virology* 1992;66:757–765. [PubMed: 1731110]
- Fouchier R, Brouwer M, Broersen S, Schuitemaker H. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *Journal of Clinical Microbiology* 1995;33:906–911. [PubMed: 7790458]
- Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical Science* 1992;7 (4):457–472.
- Goodman SN. Toward evidence-based medical statistics. 2: The Bayes Factor. *Annals of Internal Medicine* 1999;130:1005–1013. [PubMed: 10383350]
- Jensen M, Li F, van't Wout A, Nickle D, Shriner D, He X, McLaughlin S, Shankarappa R, Margolick JB, Mullins JI. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of HIV-1 *env* V3 loop sequences. *Journal of Virology* 2003;77:13376–13388. [PubMed: 14645592]
- Kijak GH, Simon V, Balfe P, Vanderhoeven J, Pampuro SE, Zala C, Ochoa C, Markowitz M, Salomon H. Origin of Human Immunodeficiency Virus Type 1 quasispecies emerging after antiretroviral treatment interruption in patients with therapeutic failure. *Journal of Virology* 2002;76 (14):7000–7009. [PubMed: 12072500]
- Kitchen C, Lu J, Suchard M, Hoh R, Martin J, Kuritzkes D, Deeks SG. Continued evolution in gp41 after interruption of enfuvirtide in subjects with advanced HIV type 1 disease. *AIDS Research and Human Retroviruses* 2006;22 (12):1260–1266. [PubMed: 17209769]
- Kitchen C, Philpott S, Burger H, Weiser B, Anastos K, Suchard M. Evolution of human immunodeficiency virus type 1 coreceptor usage during antiretroviral therapy. *Journal of Virology* 2004;78 (20):11296–11302. [PubMed: 15452249]
- Labrosse B, Morand-Joubert L, Goubard A, Rochas S, Labernardiere J, Pacaonowski J, Meynard J, Hance A, Clavel F, Mammano F. Role of the envelope genetic context in the development of enfuvirtide resistance in human immunodeficiency virus type 1-infected patients. *Journal of Virology* 2006;80:8807–8819. [PubMed: 16912327]
- Lu J, Deeks S, Hoh R, Beatty G, Kuritzkes B, Martin JN, Kuritzkes DR. Rapid emergence of enfuvirtide resistance in HIV-1 infected patients: Results of a clonal analysis. *Journal of Acquired Immune Deficiency Syndromes* 2006;43:60–64. [PubMed: 16885776]
- Melby T, Despirito M, Demasi R, Heilek-Snyder G, Greenberg M, Graham N. HIV-1 coreceptor use in triple-class-treatment-experienced patients: Baseline prevalence, correlates and relationship to enfuvirtide response. *Journal of Infectious Diseases* 2006;194:238–246. [PubMed: 16779731]
- Miller V, Sabin C, Hertogs K, Bloor S, Martinez-Picado J, D'Aquila RT. Virological and immunological effects of treatment interruption in HIV-1 infected patients with treatment failure. *AIDS* 2000;14 (18):2857–2867. [PubMed: 11153667]
- Minin VN, Dorman KS, Fang F, Suchard M. Dual multiple change-point models leads to more accurate recombination detection. *Bioinformatics* 2005;21(13):3034–3042. [PubMed: 15914546]
- Morse C, Maldarelli F. Enfuvirtide antiviral activity despite rebound viremia and resistance: Fitness tampering or a case of persistent braking on entering. *Journal of Infectious Diseases* 2007;185:318–321. [PubMed: 17205468]
- Raftery, AE. Technical Report. Department of Statistics, University of Washington; 1993. Approximate Bayes factors and accounting for model uncertainty in generalised linear models; p. 255

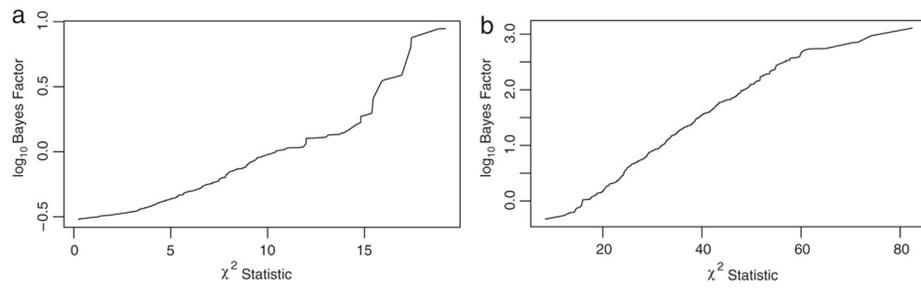
- Rambaut A, Grasly NC. Seq-Gen: An application for the Monte Carlo simulation of DNA sequences along phylogenetic trees. *Computer Applications in the Biosciences* 1997;13:235–238. [PubMed: 9183526]
- Ray N, Harrison JE, Blackburn LA, Martin JN, Deeks S, Doms RW. Clinical resistance to Enfuvirtide does not affect susceptibility of Human Immunodeficiency Type 1 to other classes of entry inhibitors. *Journal of Virology* 2007;81 (7):3240–3250. [PubMed: 17251281]
- Ruiz L, Ribera E, Bonjoch A, Romeu J, Martinez-Picado J, Paredes R. Role of structured treatment interruption before a 5-drug salvage antiretroviral regimen: The Retrogene Study. *Journal of Infectious Diseases* 2003;188 (7):977–985. [PubMed: 14513417]
- Suchard M, Kitchen C, Sinsheimer J, Weiss R. Hierarchical phylogenetic models for analyzing multipartite sequence data. *Systematic Biology* 2003;52 (5):649–664. [PubMed: 14530132]
- Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 1993;10:512–526. [PubMed: 8336541]
- Wyatt R, Sodroski J. The HIV-1 envelope glycoproteins: Fusogens, antigens, and immunogens. *Science* 1998;91:9770–9774.



**Fig. 1.** Phylogenetic tree illustrating clustering by time for (a) gp41 and (b) gp120 from a representative patient. A scale bar is drawn for each tree and represents the expected number of nucleotide substitutions per site.



**Fig. 2.** Frequency histograms showing the posterior distributions of the two interaction variables  $\beta_1$  and  $\beta_2$ . The prior distribution is plotted with the solid line. Substantial mass around 0 for both interaction variables reflects support for independence.

**Fig. 3.**

Results from 500 independent simulations under the (a) independence hypothesis and (b) dependence hypothesis. Values of the  $\log_{10}$  Bayes Factor are plotted against the  $\chi^2$  statistic. For purposes of clarity, the Bayes Factor for the simulation is modeled as the probability of the dependence hypothesis given the data over the probability of the independence hypothesis given the data. Thus values less than 1 support the independence hypothesis and values greater than 1 support the dependence hypothesis.

Table 1

Table 1 illustrates for each patient and time point the CD4+ T-cell count, log viral load, amino acid sequence of the V3 loop in gp120 and the predicted coreceptor utilization using three different methods: Position Specific Scoring Matrices (PSSM), the 11/25 method which looks at the charge at positions 11 and 25, and the net charge method which looks at the overall net charge of the V3 loop

Patient	Time point	CD4 count	Log viral load	V3 sequence	PSSM	11/25	Net charge
3019	1: Pre-ENF	ND	4.58	CTRPGNNTRRSITIGPGRAFYNIIIGDIRQAYC	R5	R5	R5
	2: ENF	62	4.55	----S-K-K--R----VY-T-G--K-F-	X4	R5	X4
	3: Post-ENF	49	4.12	-----K-K--R----VY-T-R-EN-K-S-	X4	R5	X4
3151	1: Pre-ENF	87	5.08	CVRPNNNTRRSIPMGPGKAFYTTGEIIGDIRQAHK	R5	R5	R5
	2: ENF	130	4.92	-T-----Q-----F-----	R5	R5	R5
	3: Post-ENF	120	5.16	-T-----Q-----F-----	R5	R5	R5
3502	1: Pre-ENF	140	4.45	CTRPNNNTRKGIQIIGHIGPGRVFYTTGKVVGDIRKAHC	X4	X4	X4
	2: ENF	193	5.29	-----	X4	X4	X4
	3: Post-ENF	159	5.49	-----	X4	X4	X4
3505	1: Pre-ENF	11	5.44	CTRPGNNTKGIHIGPGRAYATKRIIGDIRRAHC	R5	X4	X4
	2: ENF	55	5.65	-----K--	X4	X4	X4
	3: Post-ENF	14	5.58	-----K--	X4	X4	X4
3507	1: Pre-ENF	8	4.92	CVRPNNNTRRRGINIGPGRAYATGDIIGNIRQAYC	R5	R5	R5
	2: ENF	14	4.40	-T---PT---R-H-----KN-Q-S---S-	X4	X4	X4
	3: Post-ENF	7	5.42	-T-----R-H-----KT-Q-S---S-	X4	X4	X4
3508	1: Pre-ENF	9	5.03	CTRPNNNTRKSIPIGPGRAFAYATGDIIGDIRKAYC	R5	R5	R5
	2: ENF	18	5.66	-----	R5	R5	R5
	3: Post-ENF	34	5.71	----D-----	R5	R5	R5
3414	1: Pre-ENF	12	5.14	CTRPNNNTRKSVTMGPGRVWYSTGIVGDIRRAHC	X4	R5	X4
	2: ENF	57	5.24	-----	X4	R5	X4
	3: Post-ENF	23	5.56	-----T-	X4	R5	X4

There are three time points: (1) before ENF therapy (Pre-ENF), (2) during ENF therapy (ENF) and (3) after ENF interruption (Post-ENF). ND = Not Done.

**Table 2**

Individual estimates of evolutionary parameters for each patient in (a) gp41 and (b) gp120

Patient	Tree					$\mu \times 100$	
	E	R	O	$\kappa_1$	$\kappa_2$		$\alpha$
<b>(a)gp41</b>							
3019	0.62	0.29	0.09	4.56 (1.23)	2.24 (0.75)	0.73 (1.30)	3.90 (1.49)
3151	>0.99	<0.01	<0.01	4.03 (0.99)	2.50 (0.79)	0.88 (1.51)	3.93 (1.50)
3502	>0.99	<0.01	<0.01	4.08 (1.14)	3.52 (1.24)	0.78 (1.50)	3.41 (1.37)
3505	0.95	0.04	0.01	3.83 (0.99)	3.41 (0.82)	0.88 (1.63)	4.00 (1.50)
3507	0.97	0.03	<0.01	2.32 (0.62)	1.38 (0.45)	0.91 (1.70)	8.61 (3.27)
3508	>0.99	<0.01	<0.01	3.73 (1.04)	2.83 (0.99)	0.83 (1.48)	3.42 (1.36)
3514	0.03	<0.01	0.97	4.31 (1.11)	2.73 (0.88)	0.82 (1.73)	4.08 (1.52)
<b>(b)gp120</b>							
3019	>0.99	<0.01	<0.01	3.68 (0.88)	2.67 (0.82)	0.63 (2.17)	5.96 (2.17)
3151	>0.99	<0.01	<0.01	4.52 (1.18)	2.55 (0.88)	0.75 (1.31)	5.12(1.86)
3502	0.99	0.01	<0.01	3.82 (1.06)	2.96 (1.02)	0.70 (1.58)	4.84 (1.81)
3505	>0.99	<0.01	<0.01	3.51 (0.92)	3.3 (1.08)	0.79 (1.36)	5.07 (1.81)
3507	>0.99	<0.01	<0.01	3.33 (0.78)	2.87 (0.82)	0.82 (1.32)	6.02 (2.11)
3508	0.84	0.11	0.05	3.89 (1.11)	2.11 (0.75)	0.82 (1.45)	4.55 (1.73)
3514	0.72	0.1	0.18	3.34 (0.86)	1.92 (0.65)	0.69 (1.15)	6.08 (2.20)

For each patient, the estimates of the tree are for the posterior probability of each tree. Tree E is supported when the Post-ENF strain is nearest neighbors to the ENF strain and represents continued evolution. Tree R is favored when the Post-ENF strain is most closely related to the Pre-ENF strain suggesting re-emergence of an archived strain and Tree O is achieved when the Pre-ENF and ENF strains cluster together. The posterior mean and standard error are presented for  $\kappa_1$ , the transition/transversion rate ratio between the purines,  $\kappa_2$ , the transition/transversion rate ratio between the pyrimidines,  $\alpha$  the rate variation shape parameter, and  $\mu$ , the expected divergence.

**Table 3**

Hierarchical population level posterior estimates and standard errors

Parameter	Estimate
$T_E$	>0.99 (0.02)
$T_R$	<0.01 (<0.01)
$T_O$	<0.01 (<0.01)
$K_1$	3.65 (0.52)
$K_2$	2.43 (0.42)
$M$	4.57 (0.87)
$A$	0.77 (0.93)
$\sigma_{\kappa 1}$	1.45 (0.13)
$\sigma_{\kappa 2}$	1.55 (0.16)
$\sigma_{\mu}$	1.59 (0.17)
$\sigma_a$	0.49 (0.14)
$\Pi_A$	0.37 (0.02)
$\Pi_G$	0.22 (0.01)
$\Pi_C$	0.18 (0.01)
$\Pi_T$	0.23 (0.01)
Bayes factor	1428

Tree E ( $T_E$ ), Tree R ( $T_R$ ), and Tree O ( $T_O$ ) represent the population level probabilities of each tree. The Bayes Factor of Tree E over Tree R represents the level of support for the continued evolution hypothesis over the recall of archived strains. Bayes Factors greater than 100 are considered to be strong evidence of one hypothesis over another (Goodman, 1999).