Phylodynamics of the HIV-1 Epidemic in Cuba

Edson Delatorre, Gonzalo Bello*
Laboratório de AIDS e Imunologia Molecular, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil

Abstract

Previous studies have shown that the HIV-1 epidemic in Cuba displayed a complex molecular epidemiologic profile with circulation of several subtypes and circulating recombinant forms (CRFs); but the evolutionary and population history of those viral variants remains unknown. HIV-1 pol sequences of the most prevalent Cuban lineages (subtypes B, C and G, CRF18_cpx, CRF19_cpx, and CRFs20/23/24_BG) isolated between 1999 and 2011 were analyzed. Maximum-likelihood analyses revealed multiple introductions of subtype B (n=66), subtype C (n=10), subtype G (n=8) and CRF18_cpx (n=2) viruses in Cuba. The bulk of HIV-1 infections in this country, however, was caused by dissemination of a few founder strains probably introduced from North America/Europe (clades BCCu1 and BCCu2), east Africa (clade CCu1) and central Africa (clades GCu1, CRF18cu1 and CRF19cu1), or locally generated (clades CRFs20/23/24_BG). Bayesian-coalescent analyses show that the major HIV-1 founder strains were introduced into Cuba during 1985–1995; whereas the CRFs_BG strains emerged in the second half of the 1990s. Most HIV-1 Cuban clades appear to have experienced an initial period of fast exponential spread during the 1990s and early 2000s, followed by a more recent decline in growth rate. The median initial growth rate of HIV-1 Cuban clades ranged from 0.4 year⁻¹ to 1.6 year⁻¹. Thus, the HIV-1 epidemic in Cuba has been a result of the successful introduction of a few viral strains that began to circulate at a rather late time of the AIDS pandemic, but then were rapidly disseminated through local transmission networks.

Introduction

The global dissemination of the Human immunodeficiency virus type 1 (HIV-1) group M clade during the second half of the twentieth century has resulted in the generation of a diverse collection of genetic variants classified into subtypes, sub-subtypes, circulating recombinant forms (CRFs) and unique recombinants forms (URFs). The HIV-1 epidemic in the Americas is typically characterized by the co-circulation of several subtypes and circulating recombinant forms (CRFs); but the evolutionary and population history of most prevalent HIV-1 clades circulating in Cuba remains to be elucidated. In this study, we used a Bayesian coalescent-based method and a comprehensive data set of HIV-1 subtype B (n=322), and non-B subtypes (n=420) pol sequences of Cuban origin isolated between 1999 and 2011, to date the origin and reconstruct the demographic history of major HIV-1 variants circulating in Cuba.

Materials and Methods

HIV-1 Cuban sequence datasets

We downloaded all HIV-1 Cuban sequences covering the entire protease and partial reverse transcriptase (PR/RT) regions of the pol gene (nt 2253–3272 relative to HXB2 clone) classified as subtypes B (n=322), C (n=49), G (n=35), CRF18_cpx (n=71), CRF19_cpx (n=167), and CRFs20/23/24_BG (n=118) that were available at the Los Alamos HIV Sequence Database (www.hiv.lanl.gov) by March 2013. HIV-1 pol sequences were retrieved from both antiretroviral therapy naive and HAART treated patients from different Cuban regions between 1999 and 2011, as described in previous studies [2,3,4,5,6]. Sequences were aligned using the CLUSTAL X program [21]. To avoid any bias on the
phylogenetic reconstructions, all sites with major antiretroviral drug resistance mutations in PR (30, 32, 46, 47, 48, 50, 54, 76, 82, 84, 88 and 90) or RT (41, 65, 67, 69, 70, 74, 100, 101, 103, 106, 115, 138, 151, 181, 184, 188, 190, 210, 215, 219 and 230) detected in at least two sequences were excluded from each alignment. All alignments are available from the authors upon request.

HIV-1 reference datasets

HIV-1 Cuban sequences were combined with reference sequences of diverse origin that matched the selected genomic region and were available at the Los Alamos HIV Sequence Database. Subtype B Cuban sequences were aligned with reference sequences representative of the viral diversity in US (n=525), France (n=348), and the Caribbean (n=417) (Table S1). Subtype C Cuban sequences were aligned with representative sequences from central (n=53), eastern (n=330) and southern (n=545) African regions (Table S2). The HIV-1 subtype G Cuban sequences were combined with all available subtype G sequences of African origin (n=457) (Table S3). The CRF19_cpx Cuban sequences were aligned with all available CRF19_cpx sequences from other countries (n=3) and subtype D sequences of African origin (n=1,112) (Table S4). Finally, the HIV-1 CRF18_cpx and CRFs_BG Cuban sequences were combined with all available

Figure 1. ML phylogenetic tree of HIV-1 subtype B pol (−1000 pb) sequences circulating in Cuba (n=322), US (n=525), France (n=348), and other Caribbean countries (n=418). The branches are colored according to the origin of each sequence, as indicated at the legend (bottom left). The circular brackets highlight the position of the pandemic (B_PANDEMIC, green line) and non-pandemic (B_CARIBBEAN, blue line) HIV-1 subtype B clades. Shaded boxes highlight the position of the two major HIV-1 subtype B Cuban clades (B_CUB, and B_CUB). The number of Cuban sequences distributed accordingly to the clade size is shown (top left). Key nodes with LRTsupport values >0.80 (*) and ≥0.90 (**) are indicated. The tree was rooted using HIV-1 subtype D reference sequences (gray branches). The branch lengths are drawn to scale with the bar at the bottom indicating nucleotide substitutions per site.

doi:10.1371/journal.pone.0072448.g001
CRF18_cpx \( (n = 15) \) and CRFs20/23/24_BG \( (n = 7) \) sequences from other countries (Tables S4 and S5).

Subtype assignment

The subtype assignment and recombinant structure of all sequences here included was confirmed by: REGA HIV subtyping tool v.2 [22]; Bootscanning with Simplot software v3.5.1 [23] and Maximum Likelihood (ML) phylogenetic analysis. In bootscan analyses, supporting branching with reference sequences from all HIV-1 group M subtypes were determined in Neighbor-Joining trees based on 100 re-samplings, within a 250 bp window moving in steps of 10 bases. ML phylogenetic trees were inferred under the best nucleotide substitution model (Table S6) and an uncorrelated Lognormal relaxed clock model using the PhyML program [25] and the PhyML web server [26]. Heuristic tree search was performed using the SPR branch-swapping algorithm and the reliability of the obtained topology was estimated with the approximate likelihood-ratio test \( \text{aLRT} \) [27] based on the Shimodaira-Hasegawa-like procedure. The ML trees were visualized using the FigTree v1.4.0 program [28]. All HIV-1 sequences displaying incorrect clade assignment and/or frameshift mutations were excluded from the study, with the exception of four CRF23_BG sequences that were reclassified as CRF20_BG (GenBank accession numbers FJ481689 and FJ585687) and CRF24_BG (GenBank accession numbers JQ585465 and FJ481688).

Reconstruction of evolutionary and demographic history

The evolutionary rate \( \mu \) (nucleotide substitutions per site per year, subst./site/year), the age of the most recent common ancestor \( T_{\text{mrca}}, \text{years} \), and the mode and rate \( r \) (years \(^{-1} \)) of population growth of different Cuban HIV-1 clades were jointly estimated using the Bayesian Markov Chain Monte Carlo (MCMC) approach as implemented in BEAST v1.7.5 [29,30]. Analyses were performed using the best nucleotide substitution model (Table S6) and an uncorrelated Lognormal relaxed clock model.
A Bayesian Skyline coalescent tree prior [32] was first used to estimate $\mu$, the $T_{\text{mrca}}$, and the change in effective population size through time. Estimates of the population growth rate were subsequently obtained using a logistic growth coalescent tree prior that was the model pointed out by the Bayesian Skyline plot and that also provided the best fit to the demographic signal contained in most datasets. Comparison between demographic models was performed using the log marginal likelihood (ML) estimation based on path sampling (PS) and stepping-stone sampling (SS) methods [33]. MCMC chains were run for 10–50$\times 10^6$ generations. Adequate chain mixing and uncertainty in parameter estimates were assessed by calculating the effective sample size (ESS) and the 95% Highest Probability Density (HPD) values respectively, after excluding an initial 10% using the TRACER v1.5 program [34].

**Results**

**Identification of major HIV-1 Cuban clades**

The ML analysis of HIV-1 subtypes B, C and G sequences from Cuba and other countries from the Americas, Europe and Africa revealed that most Cuban strains branched in well-supported country-specific sub-clades. Of the 322 HIV-1 subtype B Cuban sequences analyzed, 180 (56%) formed a large country-specific monophyletic sub-clade (BCU-I, $aLRT = 0.81$), 44 (14%) branched in two clusters of medium size ($15 < n < 30$), 52 (16%) formed clusters of small size ($n \leq 10$), and the remaining 46 (14%) represented non-clustered sequences (Fig. 1). Of note, all subtype B Cuban sequences branched in a large BPANDEMIC monophyletic cluster ($aLRT = 0.80$) together with most subtype B sequences from US (92%) and all sequences from France (100%); whereas most non-Cuban Caribbean sequences (60%) occupy the deepest branches within B clade (Fig. 1).
Of the 49 HIV-1 subtype C Cuban sequences analyzed, 34 (69%) branched in a single monophyletic sub-cluster (C CU-I, \(aLRT = 0.94\)), six (12%) branched in a second well supported minor clade (C CU-II, \(aLRT = 0.98\)), and the remaining nine (18%) represented non-clustered sporadic lineages (Fig. 2). The major clade CCU-I was nested within Ethiopian sequences that belongs to the previously called C EA clade [35], a viral lineage characteristic of the east African region (Fig. 2). The minor clade C CU-II, by contrast, was nested within subtype C sequences from southern Africa (Fig. 2). Non-clustered Cuban sequences were scattered among strains from Ethiopia and southern African countries.

Of the 35 HIV-1 subtype G Cuba sequences analyzed, 26 (74%) branched in a single monophyletic sub-cluster (G CU, \(aLRT = 0.87\)) and the remaining nine (26%) represented sporadic lineages of one or two sequences. Although most subtype G African strains included in our analysis were from the western region \((n = 366, 84\%)\), the clade G CU and most sporadic subtype G Cuban lineages were nested among basal sequences from the central African region (Angola, DRC and Cameroon) (Fig. 3). There was only one
Cuban sequence that branched within a major African subtype G sub-clade mainly composed by sequences from Nigeria.

Test the monophyletic origin of the HIV-1 CRF_{18_cpx} Cuban sequences was very much complicated because the scarcity of CRF_{18_cpx} (n = 12) and the absence of CRF_{19_cpx} pol sequences of African origin available in public databases. Because CRF_{19_cpx} is subtype D in the pol fragment analyzed, we decided to include all available subtype D pol sequences of African origin in our analysis. ML analysis revealed that all (except one) CRF_{18_cpx} and all CRF_{19_cpx} sequences from Cuba branched in highly supported (aLRT≥0.90) monophyletic sub-clusters (CRF_{18CU} and CRF_{19CU}) that were nested within CRF_{18_cpx} and subtype D pol sequences of central African origin, respectively (Fig. 4). The few CRF_{18_cpx} isolated in Europe (n = 2) and South America (n = 1) were intermixed among basal African strains; whereas all CRF_{19_cpx} detected in Europe (n = 3) branched within the clade CRF_{19CU} (Fig. 4).

As expected, the CRF\textsubscript{20/23/24\_BG} Cuban sequences formed three well-supported (aLRT≥0.90) monophyletic lineages (Fig. 5). The few CRF\textsubscript{20\_BG} (n = 4) and CRF\textsubscript{24\_BG} (n = 3) sequences isolated outside Cuba (Spain and Greece) were intermixed among Cuban strains (Fig. 5); thus supporting a Cuban origin for all those European sequences.

**Time scale of major HIV-1 Cuban clades**

Bayesian MCMC analyses under a relaxed molecular clock model were used to estimate the substitution rate and T\textsubscript{MRCA} of all HIV-1 Cuban clades with a minimum size of 25 sequences. A few subtype B (n = 4) and CRF_{19_cpx} (n = 2) sequences with anomalously long branches in the phylogenetic tree, were excluded. The final number of HIV-1 Cuban sequences included in the evolutionary analyses is shown in Table 1. The median estimated evolutionary rates for the pol region of the different HIV-1 clades were roughly similar, ranging from 2.0\times10^{-3} subst./site/year (CRF_{19CU} clade) to 3.4\times10^{-3} subst./site/year (CRF_{19CU} clade), with a considerable overlap of the 95% HPD intervals (Table 1). The coefficient of rate variation was higher than zero for all HIV-1 datasets analyzed (Table 1), thus supporting the use of a relaxed molecular clock model to reconstruct the time-scale of major HIV-1 Cuban lineages.

The median T\textsubscript{MRCA} of those HIV-1 clades imported into Cuba range between 1987 (CRF_{19CU}) and 1994 (GCU clade); whereas the median T\textsubscript{MRCA} of those CRF\textsubscript{BG} variants locally generated varied between 1996 and 1998 (Table 1). The T\textsubscript{MRCA} of CRF\textsubscript{20\_BG} and CRF\textsubscript{24\_BG} clades estimated from the single CRF datasets were almost identical to those estimated from the combined CRF\textsubscript{20/23/24\_BG} data set (Table 1), indicating that all Cuban CRFs\textsubscript{BG} evolved at quite similar rates. A previous study [20], proposed that Cuban CRF\textsubscript{BG} viruses derive from a common recombinant ancestor generated by recombination between clade GCU and the second most prevalent subtype B clade (B\textsubscript{SUBCH}) (Fig. 1). The analysis of the combined CRF\textsubscript{20/23/24\_BG} data set allows us to estimate the median T\textsubscript{MRCA} of that putative BG recombinant ancestor at 1991, roughly coinciding

---

**Table 1.** Evolutionary rate and time-scale of major HIV-1 Cuban clades.

| HIV-1 clade | N  | Sampling interval | $\mu$ (subst./site/year) | Coefficient of variation | T\textsubscript{MRCA} |
|------------|----|-------------------|--------------------------|--------------------------|----------------------|
| $B_{CU}$   | 176| 2003–2011         | 3.0\times10^{-3}         | 0.30                     | 1992                 |
|            |    |                   | (2.4\times10^{-3}–3.6\times10^{-3}) | (0.21–0.39) | (1988–1994) |
| $B_{CUI}$  | 27 | 1999–2011         | 2.4\times10^{-3}         | 0.25                     | 1991                 |
|            |    |                   | (1.6\times10^{-3}–3.2\times10^{-3}) | (0.01–0.45) | (1986–1994) |
| $C_{CU}$   | 34 | 2003–2011         | 2.8\times10^{-3}         | 0.41                     | 1994                 |
|            |    |                   | (2.0\times10^{-3}–3.8\times10^{-3}) | (0.19–0.65) | (1990–1998) |
| $G_{CU}$   | 26 | 1999–2011         | 2.0\times10^{-3}         | 0.56                     | 1988                 |
|            |    |                   | (1.0\times10^{-3}–3.3\times10^{-3}) | (0.36–0.81) | (1976–1995) |
| CRF_{18\_CU} | 61 | 1999–2011         | 2.6\times10^{-3}         | 0.40                     | 1992                 |
|            |    |                   | (1.9\times10^{-3}–3.5\times10^{-3}) | (0.25–0.59) | (1987–1996) |
| CRF_{19\_CU} | 158| 1999–2011        | 3.4\times10^{-3}         | 0.38                     | 1987                 |
|            |    |                   | (2.9\times10^{-3}–4.0\times10^{-3}) | (0.30–0.47) | (1983–1991) |
| CRF\textsubscript{20/23/24\_BG} | 117| 1999–2011       | 2.6\times10^{-3}         | 0.35                     | 1991                 |
|            |    |                   | (2.1\times10^{-3}–3.1\times10^{-3}) | (0.25–0.45) | (1986–1994)* |
| CRF\textsubscript{20\_BG}  | 56 | 1999–2011        | 2.6\times10^{-3}         | 0.35                     | 1996                 |
|            |    |                   | (2.1\times10^{-3}–3.1\times10^{-3}) | (0.25–0.45) | (1994–1998)* |
|            |    |                   | 2.4\times10^{-3}         | 0.27                     | 1996                 |
|            |    |                   | (1.8\times10^{-3}–3.0\times10^{-3}) | (0.10–0.44) | (1994–1998) |
| CRF\textsubscript{23\_BG} | 11 | 2003–2011        | 2.6\times10^{-3}         | 0.35                     | 1998                 |
|            |    |                   | (2.1\times10^{-3}–3.1\times10^{-3}) | (0.25–0.45) | (1996–2000)* |
| CRF\textsubscript{24\_BG} | 50 | 2003–2011        | 2.6\times10^{-3}         | 0.35                     | 1997                 |
|            |    |                   | (2.1\times10^{-3}–3.1\times10^{-3}) | (0.25–0.45) | (1996–1999)* |
|            |    |                   | 2.2\times10^{-3}         | 0.36                     | 1998                 |
|            |    |                   | (1.6\times10^{-3}–2.8\times10^{-3}) | (0.19–0.54) | (1996–2000) |

*Estimates obtained from the combined CRF\textsubscript{20/23/24\_BG} data set.

doi:10.1371/journal.pone.0072448.t001
with the estimated TMRCA of the parental clades GCU and BCU-II (Table 1).

Demographic history of major HIV-1 Cuban clades

The Bayesian skyline plot (BSP) analyses suggest that all HIV-1 Cuban clades experienced an initial phase of fast exponential growth followed by a more recent decline in growth rate (Fig. 6). The growth rate of most HIV-1 Cuban clades seems to start to decrease around the early 2000s, except for clades BCU-II and CRF24_BG that seem to stabilize at earlier (before 2000) and later (after 2005) time points, respectively. The BSP analyses also suggest that the coalescent model of logistic population growth fits the demographic information contained in all HIV-1 Cuban data sets better than the exponential one.

To test this, log ML for the logistic and exponential growth models were calculated using both PS and SS methods. The model of logistic population growth was strongly supported over the exponential one for most HIV-1 Cuban clades (log BF > 3), with exception of GCU and CRF24_BG for which only a weak support was obtained (log BF = 0.9–1.0) (Table 2). Such a low BF support to the logistic growth model could be explained by the low number of sequences in clade GCU (n = 26) and the very recent stabilization of clade CRF24_BG (after 2005). Moreover, the overall time-scale and demographic pattern obtained from both BSP and logistic growth coalescent tree priors were very similar for all HIV-1 Cuban clades (Fig. 6).

According to the logistic model, the median initial growth rates of HIV-1 Cuban clades range between 0.40 year$^{-1}$ (CRF19CU) to 1.57 year$^{-1}$ (BCU-II) with some overlap of the 95% HPD intervals for most lineages, except between CRF19CU and clades BCU-I, BCU-II, CCI-I, CRF20_BG and CRF24_BG (Fig. 7).

**Discussion**

The Cuban HIV epidemic is unique in the Americas because of the exceptionally low HIV prevalence, estimated at 0.20% in adults in 2011 [36], and the unusually high HIV-1 genetic diversity with circulation of subtype B and several non-B subtypes [2,3,4,5,6,7]. Our study indicates that most HIV-1 infections in Cuba derived from the dissemination of a few founder viruses that were either introduced from the Americas/Europe (subtype B) and Africa (subtype C, subtype G, CRF18_cpx and CRF19_cpx) or were locally generated (CRF20/23/24_BG).

The most accepted model of worldwide HIV-1 subtype B dissemination suggests that the virus moved from Haiti to other Caribbean islands and to the United States (US), and then from the US to the rest of the world establishing a “BPANDEMIC” clade [37]. The phylogenetic analysis here performed revealed multiple (n=66) introductions of HIV-1 BPANDEMIC strains in Cuba, although the bulk of the subtype B epidemic in this country resulted from the dissemination of only a few clades. The two most prevalent clades BCU-I and BCU-II comprises about 55% and 8% of all subtype B sequences from Cuba here included, respectively. We estimate that these clades most probably emerged in Cuba in the early 1990s, much later than the estimated origin of subtype B epidemics in Haiti and the US (1960–1970) [37,38,39]. The estimated TMRCA of clades BCU-I and BCU-II coincides with a crisis in the Cuban economy caused by the collapse of the Soviet Union in 1991 that precipitated important investments in the tourist industry and a sharp increase in the number of tourist mostly from...
sequences strongly suggests that the CRF19 CU clade also derives CCU-I. Thus, although Cuban personnel were stationed in several CRF18CU comprise 69%, 74% and 98% of all subtype C, subtype C (n) massive influx of subtype B PANDEMIC strains and the apparent generated by recombination between parental clades BCU-II and diognosed in 2003 [3]. It was proposed that all Cuban CRFs_BG increased from 0% in those diagnosed in 1998 to 31% in those the proportion of BG infections among MSM in Havana City recombinants were diagnosed with HIV-1 infection earlier than collected in 2003, none of the individuals harboring BG Such a recent expansion of BG recombinants in Cuba is fully of the CRFs_BG. North America and Europe [40], regions with a widespread circulation of the subtype B PANDEMIC clade. This may explain the massive influx of subtype B PANDEMIC strains and the apparent absence of “non-pandemic” subtype B Caribbean lineages in Cuba. Similarly to subtype B, there were multiple introductions of subtype C (n≥10), subtype G (n≥6) and CRF18_cpx (n≥2) viruses in Cuba, but only a few of them were able to get established and disseminate. The clades CCL, GCU and CRF18CU comprise 69%, 74% and 98% of all subtype C, subtype G and CRF18_cpx sequences from Cuba included in this study, respectively. The monophyletic clustering of CRF19_cpx-like pol Cuban sequences within subtype D radiation, the paucity of this genetic variant in Africa, and the recent TRECUBA of Cuban sequences strongly suggests that the CRF19CU clade also derives from a single founder event. HIV-1 clades GCU, CRF18CU and CRF19CU probably originate in central Africa, whereas clade CCL probably derives from east Africa. Our study suggests that clades CRF19CU and GCU began to circulate in Cuba around the late 1980s, followed shortly thereafter by clades CRF18CU and CCL. Thus, although Cuban personnel were stationed in several African countries since the 1970s, HIV-1 African strains were successfully disseminated within Cuba only from the late 1980s onwards. Our data suggest that HIV-1 CRFs_BG (20_BG, 23_BG and 24_BG) started to spread in Cuba in the second half of the 1990s. Such a recent expansion of BG recombinants in Cuba is fully consistent with epidemiological data showing that in samples collected in 2003, none of the individuals harboring BG recombinants were diagnosed with HIV-1 infection earlier than 1996, and all but three were diagnosed since 2000 [20]. Similarly, the proportion of BG infections among MSM in Havana City increased from 0% in those diagnosed in 1998 to 31% in those diagnosed in 2003 [3]. It was proposed that all Cuban CRFs_BG evolved from a common BG recombinant ancestor locally generated by recombination between parental clades BCLH and GCU [20]. According to our estimations, that common BG recombinant ancestor was generated in the early 1990s, thus around or immediately after the estimated onset date of parental clades BCLH and GCU and some years earlier than the emergence of the CRFs_BG. The reconstruction of the demographic history indicates that most HIV-1 Cuban clades followed a very similar growth pattern characterized by rapid dissemination until the early 2000s after which the epidemic growth rate of those epidemics started to slow-down. The expansion of the BCLH clade, by contrast, seems to decrease during 1990s; whereas the growth rate of the CRF24_BG clade probably only stabilized in the second half of the 2000s. The initial expansion of the major HIV-1 Cuban clades coincides with a sustained increase in the number of infected HIV-positive individuals in Cuba from 1991 to 2000 [41]. UNAIDS estimations indicate that the total number of people living with HIV in Cuba continued to growth in the last decade, rising from 3,100 (2,600– 4,300) in 2000 to 14,000 (12,000–16,000) in 2011 [36]. Our demographic analysis, however, suggests a trend toward stability in the effective number of infections of all major HIV-1 Cuban clades over time consistent with recent epidemiological data that shows a decrease of HIV incidence in Cuba, mainly among men, in the biennium 2010–2011 [42]. Our coalescent-based analyses suggest that CRF20_BG, CRF24_BG and BCLH have displayed a more explosive initial growth (1.0 year^{-1}–1.6 year^{-1}) than clades GCU, CRF18CU and CRF19CU (~0.4–0.6 year^{-1}); whereas clades BCLH and GCU displayed intermediate initial growth rates (~0.3 year^{-1}). Notably, all those HIV-1 Cuban clades with the fastest initial expansion rates (BCLH, BCUH, BCUH and CRFs_BG) were much more prevalent among MSM than among heterosexual (HET) persons [3]. Thus, some HIV-1 Cuban clades may have spread faster than others because they encountered, by chance, local transmission chains with higher rates of partner exchange. Transmission within a transmission network of small size and high rate of partner exchange may also explain the fast, but self limited, dissemination phase of clade BCLH. Additional influence of virological factors cannot be excluded. These results must also be interpreted with caution as most growth rate estimates here obtained displayed quite large overlapping 95% HPD intervals. It has been proposed that Cuba’s low rate of HIV infection is due to several factors that served to prevent sexual transmission of the virus, including wide-scale HIV screening and subsequent contact tracing of HIV-positive individuals, mandatory quarantine of the first HIV-infected individuals at sanatoria, free access to a well structured public health system, comprehensive HIV education campaigns, coordinated work of Cuban government agencies

| Table 2. Best fit demographic model for major HIV-1 Cuban clades. |
| --- |
| Dataset | PS |  |  |
|  | Log ML LG | Log ML EG | Log BF (LG vs EG) | Log ML LG | Log ML EG | Log BF (LG vs EG) |
| BCLU-I | −1289.71 | −1293.39 | 43.68 | −1289.56 | −1293.47 | 43.91 |
| BCLU-II | −3563.92 | −3568.51 | 4.59 | −3563.99 | −3568.71 | 4.72 |
| GCU | −4114.63 | −4118.50 | 3.87 | −4114.74 | −4118.72 | 3.98 |
| GCU | −4196.64 | −4197.66 | 1.02 | −4196.87 | −4197.73 | 0.86 |
| CRF18 | −6109.29 | −6118.91 | 9.62 | −6109.54 | −6119.01 | 9.47 |
| CRF19 | −14552.57 | −14564.21 | 11.64 | −14553.91 | −14565.09 | 11.18 |
| CRF20 | −5103.74 | −5100.80 | 17.06 | −5083.91 | −5101.05 | 17.14 |
| CRF24 | −4159.40 | −4160.43 | 1.03 | −4159.81 | −4160.70 | 0.89 |

Log marginal likelihood (ML) estimates for logistic growth (LG) and exponential growth (EG) demographic models obtained using the path sampling (PS) and stepping stone sampling (SS) methods. The Log Bayes factor (BF) is the difference of the Log ML between of alternative (H1 = LG) and null (H0 = EG) models. Log BF>1 indicates that model H1 is more strongly supported by the data than model H0.

doi:10.1371/journal.pone.0072448.t002

Phylodynamics of HIV-1 in Cuba

PLOS ONE | www.plosone.org 10 September 2013 | Volume 8 | Issue 9 | e72448
HIV-1 variants already established in the Cuban population. The estimated initial growth rates of the major HIV-1 Cuban clades (~0.4–1.6 year⁻¹), however, were comparable to those obtained for different HIV-1 epidemics in the Americas (~0.3–1.3 year⁻¹) [38,47,48,49,50,51,52], Europe (~0.4–1.3 year⁻¹) [52,53,54,55], Africa (~0.2–0.8 year⁻¹) [47,52,56,57,58] and Asia (~0.8 year⁻¹) [59]. This suggests that several factors may have contributed to delay the introduction and/or dissemination of HIV-1 in Cuba for many years; but once some HIV-1 strains got established in vulnerable HET and MSM transmission groups they spread quickly.

In summary, this study indicates that only a few subtype B and non-B subtype founder viral strains were successfully disseminated in Cuba. Some of those HIV-1 viral strains were probably introduced from North America/Europe, central Africa and east Africa between the middle 1980s and the middle 1990s; whereas other were locally generated around the late 1990s. Changes in the social and economic landscapes of Cuba occurring at the beginning of the 1990s may have fueled the introduction and/or initial dissemination of major HIV-1 Cuban clades. Although the main HIV-1 Cuban lineages began to circulate at a rather late time of the AIDS pandemic, further dissemination within vulnerable groups was rapid. These results reinforce the importance of maintaining, reviewing and updating permanently the public health measures aimed at controlling the spread of those HIV-1 variants already established in the Cuban population.

References

1. Thomson MM, Najera R (2005) Molecular epidemiology of HIV-1 variants in the global AIDS pandemic: an update. AIDS Rev 7: 210–224.
2. Cuevas MT, Ruibal I, Villahermosa ML, Diaz H, Delgado E, et al. (2002) High HIV-1 genetic diversity in Cuba. AIDS 16: 1643–1659.
3. Perez I, Thomson MM, Bleda MJ, Aragones C, Gmezcz Z, et al. (2006) HIV Type 1 molecular epidemiology in Cuba: high genetic diversity, frequent mosaicism, and recent expansion of B subtypes recombinant forms. AIDS Res Hum Retroviruses 22: 724–733.
4. Kouri V, Aleman Y, Perez I, Perez J, Fonseca C, et al. (2012) High frequency of antiviral drug resistance and non-B subtypes in HIV-1 patients failing antiviral therapy in Cuba. J Clin Virol 53: 348–355.
5. Machado LY, Blanco M, Dubed M, Diaz HM, Ruiz NM, et al. (2012) HIV type 1 genetic diversity in newly diagnosed Cuban patients. AIDS Res Hum Retroviruses 28: 956–960.
6. Machado LY, Dubed M, Diaz H, Ruiz N, Romay D, et al. (2013) Transmitted HIV type 1 drug resistance in newly diagnosed Cuban patients. AIDS Res Hum Retroviruses 29: 411–414.
7. Perez I, Kouri V, Aleman Y, Abrahantes Y, Correa C, et al. (2013) Antiretroviral drug resistance in HIV-1 therapy-naive patients in Cuba. Infect Genet Evol 16C: 144–150.
8. Thomson MM, Casado G, Posada D, Sierra M, Najera R (2005) Identification of a novel HIV-1 complex circulating recombinant form (CRF18_cpx) of Central African origin in Cuba. AIDS 19: 1155–1163.
9. Casado G, Thomson MM, Sierra M, Najera R (2005) Identification of a novel HIV-1 circulating ADG intersubtype recombinant form (CRF05_cpx) in Cuba. J Acquir Immune Defic Syndr 40: 532–537.
10. Bartolo I, Rocha C, Bartolomeu J, Gama A, Marcelino R, et al. (2009) Highly divergent subtypes and new recombinant forms prevail in the HIV/AIDS epidemic in Angola: New insights into the origins of the AIDS pandemic. Infect Genet Evol 9: 672–682.
11. Afonso JM, Bello G, Guimaraes ML, Soja M, Morgado MG (2012) HIV-1 genetic diversity and transmitted drug resistance mutations among patients from the North, Central and South regions of Angola. PLoS ONE 7: e39296.
12. Kafui ML, Robbins KE, Pienuaz D, Schaffer A, Nizamani L, et al. (2004) Recombinant viruses and early global HIV-1 epidemic. Emerg Infect Dis 10: 1227–1234.
13. Vidal N, Peer M, Mulanga-Kabeya C, Nizamani L, Robertson D, et al. (2000) Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. J Virol 74: 10490–10507.
14. Nzeima FR, Toure-Kane C, Vidal N, Obeunay P, Bikanoua B, et al. (2006) HIV-1 subtypes and recombinants in the Republic of Congo. Infect Genet Evol 15.
15. Pircher M, Diafoiska M, Papachon J, Recordon-Pinson P, Mahambou D, et al. (2012) Molecular characterization of HIV type 1 in Brazzaville, Republic of Congo, and first data on resistance to antiretroviral drugs. AIDS Res Hum Retroviruses 22: 1798–1802.
16. Marechal V, Javuin S, Selekon B, Leal J, Pelembi P, et al. (2006) Increasing HIV type 1 polymorphic diversity but no resistance to antiretroviral drugs in untreated patients from Central African Republic: a 2005 study. AIDS Res Hum Retroviruses 22: 1036–1044.
17. Carr JK, Torimiro JN, Wolfe ND, Ezel MN, Kim B, et al. (2001) The AG recombinant BNG and novel strains of group M HIV-1 are common in Cameroon. Virology 280: 168–181.
18. Vergne L, Bourgeois A, Mpoudi-Ngole E, Moungouni R, Mnbugo J, et al. (2003) Biological and genetic characteristics of HIV infections in Cameroon reveals dual group M and O infections and a correlation between SI-inducing phenotype of the predominant CRF02_AG variant and disease stage. Virology 310: 254–266.
19. Powell R, Barendgols D, Mayr L, Nyambi P (2010) The Evolution of HIV-1 Diversity in Rural Cameroon and its Implications in Vaccine Design and Trials. Viruses 2: 639–654.
20. Sierra M, Thomson MM, Posada D, Perez I, Aragones C, et al. (2007) Identification of 3 phylogenetically related HIV-1 BG intersubtype circulating recombinant forms in Cuba. J Acquir Immune Defic Syndr 45: 151–160.
21. Thompson JD, Gibson TJ, Pevnak F, Jeamoujou F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876–4882.
22. de Oliveira T, Deforche K, Cassel S, Salminen M, Parakkaev D, et al. (2003) An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics 21: 3797–3800.
23. Ray SC (1999) Simplot v2.5.0. Available: http://www.med.unl.edu/ SCHRoftware/simplot/. Accessed 2013 July 16.
24. Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25: 1253–1266.
25. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704.
26. Guindon S, Lethiec F, Duroux P, Durou P, Gascuel O (2002) PHYLML Online—a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res 33: W537–559.
27. Anisimova M, Gascuel O (2006) Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol 55: 539–552.
28. Ramana A (2009) FigTree v1.4.0: Tree Figure Drawing Tool. Available: http://tree.bio.ed.ac.uk/software/figtree/. Accessed 2013 July 16.
29. Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161: 1307–1320.
30. Drummond AJ, Ramah A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.

Supporting Information

Table S1 HIV-1 subtype B dataset. (PDF)
Table S2 HIV-1 subtype C dataset. (PDF)
Table S3 HIV-1 subtype G dataset. (PDF)
Table S4 HIV-1 CRF18_cpx and CRF19_cpx/subtype D datasets. (PDF)
Table S5 HIV-1 CRF20/23/24_cpx datasets. (PDF)
Table S6 Nucleotide substitution models selected using jModeltest program. (PDF)

Acknowledgments

We wish to thank Dr Vera Bongertz for critical reviewing of the manuscript.

Author Contributions

Conceived and designed the experiments: GB ED. Performed the experiments: ED GB. Analyzed the data: ED GB. Wrote the paper: GB ED.
31. Drummond AJ, Ho SY, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol 4: e88.
32. Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22: 1183–1192.
33. Baed G, Lemey P, Bedford T, Rambaut A, Suchard MA, et al. (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol Biol Evol 29: 2157–2167.
34. Rambaut A, Drummond A (2007) Tracer v1.4. Available: http://tree.bio.ed.ac.uk/software/tracer/. Accessed 2013 July 16.
35. Delatorre EO, Bello G (2012) Phylodynamics of HIV-1 Subtype C Epidemic in East Africa. PLoS ONE 7: e41904.
36. UNAIDS (2013) AIDS INFO DATABASE. Available: http://www.unaids.org/en/dataanalysis/datatools/aidsinfo/. Accessed 2013 July 16.
37. Gilbert MT, Rambaut A, Wlasiuk G, Spira TJ, Pitchenik AE, et al. (2007) The emergence of HIV/AIDS in the Americas and beyond. Proc Natl Acad Sci U S A 104: 18566–18569.
38. Robbins KE, Lemey P, Pybus OG, Jaffe HW, Youngpairoj AS, et al. (2003) U.S. Human immunodeficiency virus type I epidemic: date of origin, population history, and characterization of early strains. J Virol 77: 6359–6366.
39. Junqueira DM, de Medeiros RM, Matte MC, Araujo LA, Chies JA, et al. (2011) Reviewing the history of HIV-1: spread of subtype B in the Americas. PLoS ONE 6: e27489.
40. Figueiras MA (2001) International Tourism and the Formation of Productive Clusters in the Cuban Economy. Latin American Studies Association (LASA) XXII Congress, Washington, DC Available: http://www3.dogus.edu.tr/memek/DersNotlari/TOUR%20229/TOUR%20229-2004%20fall/cuba_contribution.pdf. Accessed 2013 July 16.
41. Hsieh YH, Chen CW, Lee SM, de Arazoza H (2001) On the recent sharp increase in HIV detections in Cuba. AIDS 15: 426–428.
42. UNAIDS (2012) Country progress reports. Available: http://www.unaids.org/en/dataanalysis/knowyourresponse/countryprogressreports/2012countries/ce_CU_Narrative_Report.pdf. Accessed 2013 July 16.
43. Perez-Stable EJ (1991) Cuba’s response to the HIV epidemic. Am J Public Health 81: 363–367.
44. Hansen H, Groce NE (2001) From quarantine to condoms: shifting policies and problems of HIV control in Cuba. Med Anthropol 19: 259–292.
45. de Arazoza H, Joanes J, Lounes R, Legeai C, Clemenceau S, et al. (2007) The HIV/AIDS epidemic in Cuba: description and tentative explanation of its low HIV prevalence. BMC Infect Dis 7: 130.
46. Anderson T (2009) HIV/AIDS in Cuba: lessons and challenges. Rev Panam Salud Publica 26: 78–86.
47. Salemi M, de Oliveira T, Soares MA, Pybus O, Dumas AT, et al. (2005) Different epidemic potentials of the HIV-1B and C subtypes. J Mol Evol 60: 590–605.
48. Bello G, Eyer-Silva WA, Couto-Fernandez JC, Guimaraes ML, Chequer-Fernandez SL, et al. (2007) Demographic history of HIV-1 subtypes B and F in Brazil. Infect Genet Evol 7: 263–270.
49. Bello G, Guimaraes ML, Passaes CP, Matos Almeida SE, Veloso VG, et al. (2009) Short communication: Evidences of recent decline in the expansion rate of the HIV type 1 subtype C and CRF31_BC epidemics in southern Brazil. AIDS Res Hum Retroviruses 25: 1063–1069.
50. Bello G, Ashurino PC, Ruchansky D, Guimaraes ML, Lopez-Galindez C, et al. (2010) Phylodynamics of HIV-1 circulating recombinant forms 12_BF and 38_BF in Argentina and Uruguay. Retrovirology 7: 22.
51. Ristic N, Zukanov J, Alkmim W, Diaz RS, Janini LM, et al. (2011) Analysis of the origin and evolutionary history of HIV-1 CRF20_BF and CRF29_BF reveals a decreasing prevalence in the AIDS epidemic of Brazil. PLoS ONE 6: e17485.
52. Bello G, Afonso JM, Morgado MG (2012) Phylodynamics of HIV-1 subtype F1 in Angola, Brazil and Romania. Infect Genet Evol 12: 1079–1086.
53. Hue S, Pillay D, Clewley JP, Pybus OG (2005) Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. Proc Natl Acad Sci U S A 102: 4425–4429.
54. Saleri M, de Oliveira T, Ciccozzi M, Rezza G, Goodenow MM (2008) High-Resolution Molecular Epidemiology and Evolutionary History of HIV-1 Subtypes in Albania. PLoS ONE 3: e1390.
55. Zehender G, Ebranati E, Lai A, Santoro MM, Altiri C, et al. (2010) Population dynamics of HIV-1 subtype B in a cohort of men-having-sex-with-men in Rome, Italy. J Acquir Immune Defic Syndr 55: 156–160.
56. Walker PR, Pybus OG, Rambaut A, Holmes EC (2005) Comparative population dynamics of HIV-1 subtypes B and C: subtype-specific differences in patterns of epidemic growth. Infect Genet Evol 5: 199–208.
57. Tully DC, Wool C (2010) Chronology and evolution of the HIV-1 subtype C epidemic in Ethiopia. AIDS 24: 1577–1582.
58. Delatorre E, Bello G (2013) Spatiotemporal dynamics of the HIV-1 CRF06_cpx epidemic in western Africa. AIDS.
59. Chen JH, Wong KH, Chan KC, To SW, Chen Z, et al. (2011) Phylodynamics of HIV-1 subtype B among the men-having-sex-with-men (MSM) population in Hong Kong. PLoS ONE 6: e25286.