Reconstructing the HIV-1 Epidemics in Burkina Faso Using Early Samples

#236

Gonzalo Yebra[1], Marcia L. Kalish[2] and Andrew J. Leigh Brown[1]

[1]*Institute for Evolutionary Biology, University of Edinburgh, Edinburgh, UK*

[2]*Institute for Global Health, Vanderbilt University, Nashville TN, USA*

CONTACT: gonzalo.yebra@ed.ac.uk | http://hiv.bio.ed.ac.uk

BACKGROUND

Two HIV-1 circulating recombinant forms (CRFs) dominate the epidemic in Burkina Faso: CRF06_cpx and CRF02_AG [1].

A recent paper described early sequences of both CRFs from serum samples collected in Burkina Faso in 1986 [2]. For CRF02_AG, these are the earliest available sequences.

Here, we reconstruct the phylodynamics of both CRF epidemics in Burkina Faso and throughout Western Africa using, for the first time, early and current sequences extracted from GenBank.

METHODS

The early sequences consisted of 52 protease (PR) and 37 gp41 sequences for CRF06_cpx and 15 PR and 25 gp41 for CRF02_AG.

We analysed them together with GenBank samples (1993-2013) in 3 datasets:

- African CRF06_cpx: early Burkina Faso sequences and all GenBank sequences from Africa (163 PR / 41 gp41).
- Burkina Faso CRF02_AG: early sequences and GenBank sequences from this country (142 / 39).
- Western/Central African CRF02_AG: sequences from Western Africa (380 PR; 170 gp41) and the closest sequences from Central Africa (311 / 128).

For each dataset, we analysed sequence data for both genes jointly using the BEAST multilocus analysis [3] and conducted a phylogeographic analysis by using the discrete traits analysis implemented in BEAST to reconstruct the main migration routes between countries [4].

REFERENCES

1. Tebit et al. J Med Virol 2009; 81: 1691-701.
2. Fonjungo et al. PLoS ONE 2014; 9:e92423.
3. Heled et al. Mol Biol Evol 2010; 27: 570-80.
4. Lemey et al. PLoS Comput Biol 2009; 5:e1000520.

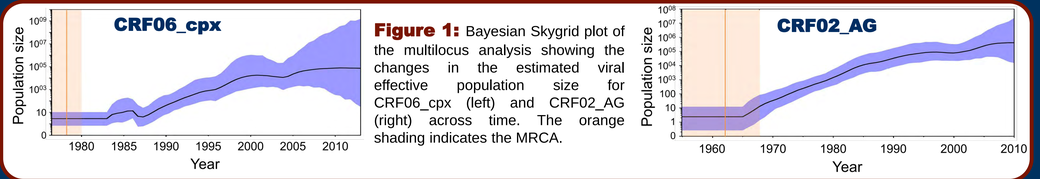


Figure 1: Bayesian Skygrid plot of the multilocus analysis showing the changes in the estimated viral effective population size for CRF06_cpx (left) and CRF02_AG (right) across time. The orange shading indicates the MRCA.

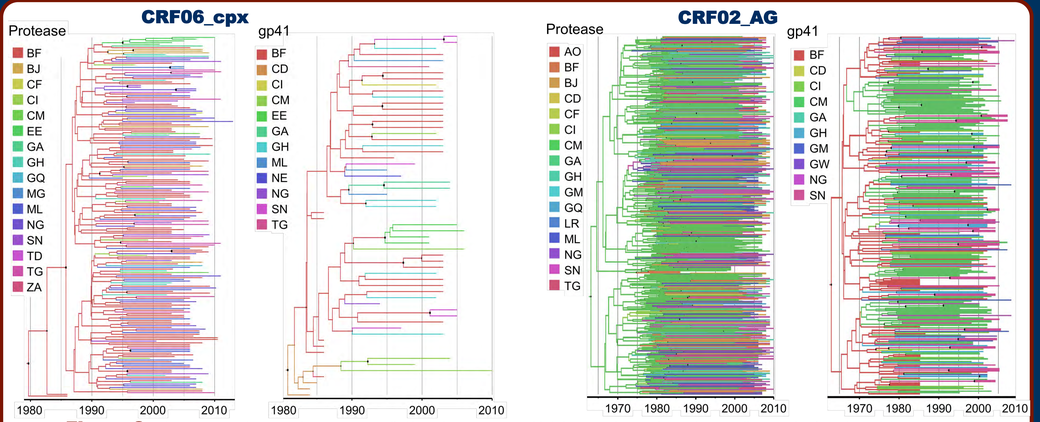


Figure 2: Time-stamped trees of CRF06_cpx (left) and CRF02_AG (right). Branches coloured according to the location of the sequences.

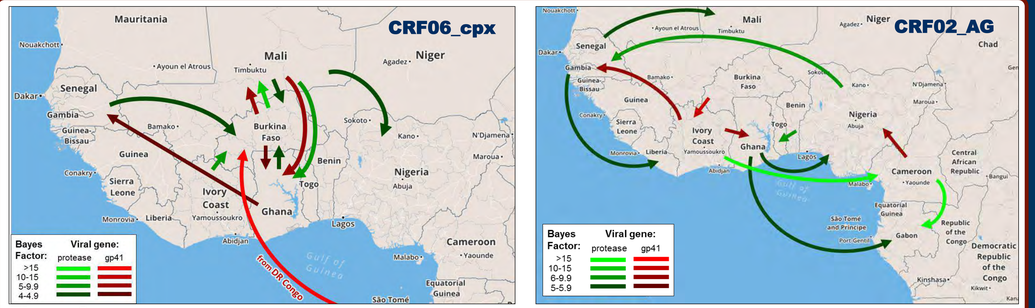


Figure 3: Migration routes of CRF06_cpx (left) and CRF02_AG (right) inferred from the phylogeographic analyses of PR (green) and gp41 (red). The colour of the arrows is proportional to the statistical support (see legend).

RESULTS

Phylodynamics of CRF06_cpx and CRF02_AG:

Dataset	MRCA		Evolutionary rate (× 10^{-3} substitutions/site/year)	
	PR	gp41	PR	gp41
African CRF06_cpx	1979 (73-83)	1981 (78-83)	5.5 (4.8-6.3)	6.1 (4.8-7.5)
Burkina Faso CRF02_AG	1979 (77-82)	1980 (77-81)	2.3 (1.8-2.9)	5.7 (4.1-7.4)
African CRF02_AG	1962 (54-68)	1963 (58-68)	1.4 (1.3-1.6)	4.5 (4.0-5.1)

The phylogeographic analysis of gp41 showed that CRF06_cpx (or at least its parental subtype G lineage) emerged in the Democratic Republic of Congo (DRC) but arrived soon after (~1981) in Burkina Faso. Both PR and gp41 showed that this recombinant radiated to the rest of Western Africa only around 1990.

In West Africa, the phylogenetic trees showed CRF02_AG sequences interspersed regardless of their sampling country, and the phylogeographic analysis showed much interconnection between countries, with Cameroon, Ghana, Ivory Coast and Senegal as main transmission hubs.

CONCLUSIONS

Burkina Faso presents a relatively young HIV epidemic, with the two main variants originating around 1980.

CRF06_cpx might have emerged in DRC but radiated throughout Western Africa from Burkina Faso.

However, CRF02_AG entered Burkina Faso though multiple introductions. The CRF02_AG epidemic showed much interchange between Western African countries but also connections with Central Africa, which suggests a great mobility in the region.

ACKNOWLEDGMENTS

GY was supported through the Pangea_HIV Consortium with support provided by the Bill & Melinda Gates Foundation and by NIH GM110749.