Characterising the Complex Mutational Pathways Associated With Maraviroc Resistant HIV-1

Xiaowei Jiang[1, *], Felix Feyertag[1], Conor Meehan[2], Simon Travers[3], Charles Craig[4, 5], Mike Westby[5], Marilyn Lewis[4, 5], David L.Robertson[1]

[1]Computational and Evolutionary Biology, Faculty of Life Sciences, University of Manchester, UK; [2]National University of Ireland Galway, Galway, Ireland; [3]South African National Bioinformatics Institute, SA MRC Bioinformatics Unit, University of the Western Cape, South Africa; [4]The Research Network, Sandwich, UK; [5]Pfizer R&D, Sandwich, UK. ***Current address:** Department of Genetics, University of Cambridge, UK.

Contact: x.jiang@cam.ac.uk, jiangx@tcd.ie

Background

The drug maraviroc differs from other HIV-1 drugs as it targets the host cell coreceptor CCR5 as opposed to a virus protein [1]. Lack of sensitivity to such entry-inhibitor drugs occurs primarily through emergence of pre-existing viruses that use the CXCR4 coreceptor, while genuine CCR5-antagonist resistance occurs through virus use of the drug-bound CCR5 coreceptor [2-4]. The aim of this project is to build a comprehensive understanding of the R5 resistance pathway based on susceptibility testing and clonal envelope (gp160) sequence data.

Methodology

We use population genetic, coevolutionary, recombination, positive selection and structural analysis to analyse the envelope protein sequences from 20 patients from the clinical trials MOTIVATE 1 and 2. All these patients have exclusively CCR5-using (R5) virus. All patients have two groups of sequences that have been sequenced before and after treatment with maraviroc/placebo-arm. Three patients under placebo-arm treatment are used as control.

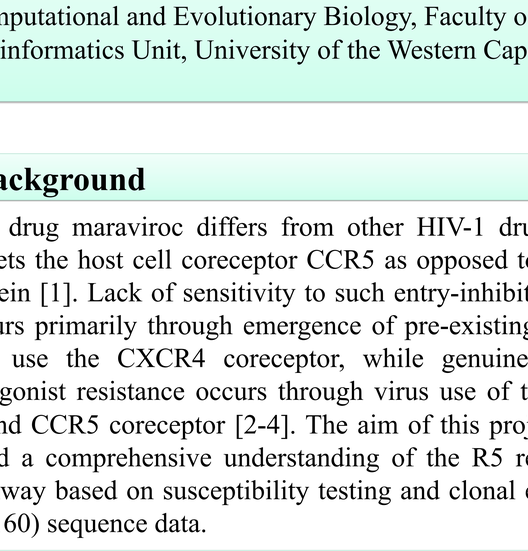


Fig.1 Evolutionary analysis. Patient 4 is presented here as a representative example. The Bayesian skyline plot (A) shows a sharp drop (blue line) in genetic diversity as therapy starts, with a rebound occurring as drug resistance develops, which is consistent with the viral load plot of the patient (red line). (B) Phylogenetic tree for patient 4 was created by BEAST v1.7.2 using a Bayesian skyline coalescent. Tree tips are calibrated by days into therapy, with the R5-sensitive strains sampled at day 1 in green, and the R5-resistant strains sampled at day 57 in red. Tip nodes indicate patient and sequence number, as well as sampling day, pooled MPI, and sequence MPI are shown in brackets. Internal nodes show the ancestral time-points in days with bars indicating the 95% highest posterior density (HPD) of the time-points.

Results

The resistant viral populations were phylogenetically distinct from those susceptible to maraviroc and associated with a clear genetic bottleneck in each patient, consistent with *de novo* emergence of resistance (Fig.1). Unique and overlap sites or point mutations were involved in the resistance. Particularly: (1) There were complex coevolutionary networks (indicating epistatic interactions in the context of protein structure and/or function) in both the sensitive and resistance populations (Fig.2). (2) A subset of coevolving sites are also under positive selection during active treatment indicating their importance for drug resistance (Fig.2). (3) CD4 binding sites formed a unique coevolutionary network that is independent of the V3 loop (Fig.3). (4) The coevolutionary network formed between the V3 loop and other regions of gp120 and gp41 intersects with sites involved in the N-linked glycosylation motif (Fig.4). (5) Key residues in the signal peptide involved in Env function and expression level coevolve with the potential N-linked glycosylation sequence motifs (NXS/TX) (Fig. 5).

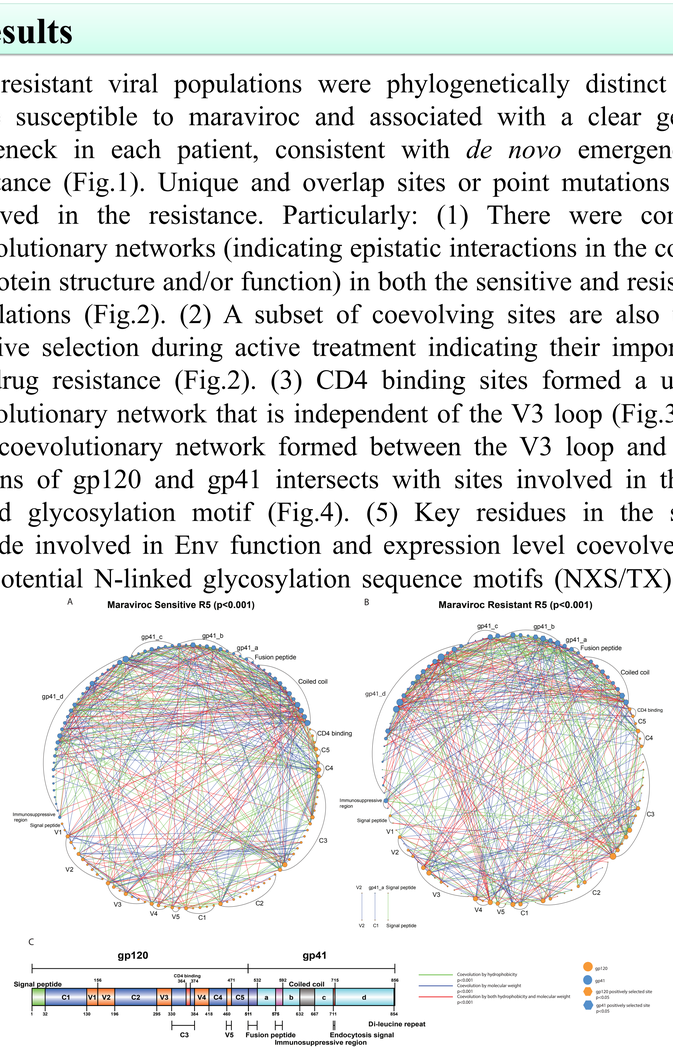


Fig.2 Coevolutionary networks of gp120 and gp41 in the sensitive and resistant R5 viruses. Coevolutionary networks of HIV-1 envelope protein explained by hydrophobicity and/or molecular weight covariation for maraviroc sensitive (A) and resistant (B) CCR5 viruses for the combined data of all patients, respectively. Coevolving amino acid sites are organised and labelled by protein domains in gp120 (orange circles) and gp41 (blue circles), accordingly. Circle size indicates the relative number of interactions in the coevolutionary networks. Coevolving sites are linked by coloured lines, which indicate the physicochemical properties of coevolution (see key and table 1 for summary).

	Coevolving (Nodes) Sites		Coevolving (Edges) Pairs	
Type	Sensitive	Resistant	Sensitive	Resistant
BLOSUM64	130	125	1413	1400
Hydrophobicity (Hydro)	103	99	295	305
Molecular Weight (MW)	108	106	319	326
Hydro & MW	81	77	131	147
Hydro/MW/Hydro & MW	120	114	483	484

Table 1. Summary of the sensitive and resistant coevolutionary networks. The numbers of coevolving sites and pairs for sensitive and resistant networks are summarised in the table. Four types of networks are compared: coevolution by BLOSUM scores covariation, coevolution by hydrophobicity and/or molecular weight covariation.

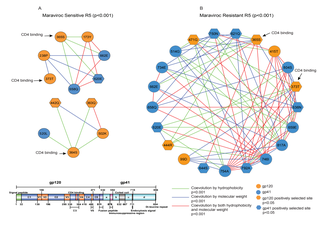


Table 2 (Sensitive)

HXB2 site	Degree	Selection	Site	Domain	Protein	N-Gly Motif
173Y	13	null	151	V2	gp120	
216P	15	null	211	C2	gp120	-3
363Q	17	0.996267	331	C3	gp120	NQS/TC
364S	4	null	332	CD4 Binding Site	gp120	NQS/TC
365S	4	null	333	CD4 Binding Site	gp120	NQS/TC
372T	1	null	341	CD4 Binding Site	gp120	
443Q	21	0.999411	392	C3	gp120	-3
502K	7	null	440	C4	gp120	
529L	4	null	464	Fusion Peptide	gp41	
626E	24	1	504	Coiled Coil	gp41	-4
669G	5	null	602	Coiled Coil	gp41	
682E	22	null	606	Coiled Coil	gp41	

Table 3 (Resistant)

HXB2 site	Degree	Selection	Site	Domain	Protein	N-Gly Motif
98C	9		70	C1	gp120	
279Y	13	0.968434	251	CD4 Binding Site	gp120	NQS/TX
365R	13	0.998585	341	CD4 Binding Site	gp120	
416T	12	null	385	C4	gp120	
417S	13	0.999503	386	C4	gp120	
471S	12	0.972233	411	C4	gp120	
514Q	10	null	454	Fusion peptide	gp41	
626E	24	0.999999	504	Coiled Coil	gp41	-4
627Q	22	0.999999	505	Coiled Coil	gp41	-3
634N	22	0.971102	580	Coiled Coil	gp41	-3
635K	19	0.999746	581	Coiled Coil	gp41	-4
638Q	10	null	602	Coiled Coil	gp41	
669E	12	null	603	Coiled Coil	gp41	
682E	19	null	606	Coiled Coil	gp41	
734S	24	null	670	gp41_ct	gp41	
748I	24	null	695	gp41_ct	gp41	
749A	24	0.978634	696	gp41_ct	gp41	
750A	21	null	697	gp41_ct	gp41	
792A	19	0.999070	736	gp41_ct	gp41	
817A	20	null	761	gp41_ct	gp41	

Fig.3 Coevolutionary networks of the CD4 binding sites. Coevolutionary networks of the CD4 binding sites explained by hydrophobicity and/or molecular weight for sensitive (A) and resistant (B) R5 viruses, respectively. Coevolving sites are mapped to the gp120 protein crystal structure. Sites under positive selection are hexagon shaped. Coevolving sites are linked by coloured lines indicating their nature of coevolution. The size of each circle indicates the relative number of interactions in the coevolutionary networks (fig. 2). Information regarding positive selection, protein domain and glycosylation can be checked in table 2 and table 3 for coevolving sites in sensitive and resistant viruses, respectively.

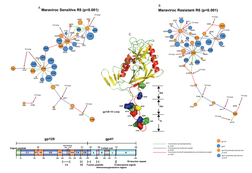


Table 4 (Sensitive)

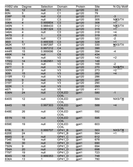


Table5(Resistant)

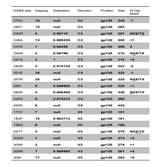


Fig.4 Coevolutionary networks of the V3 loop. Coevolutionary networks of the HIV-1 envelope protein V3 loop explained by hydrophobicity and/or molecular weight for sensitive (A) and resistant (B) R5 viruses, respectively. Information regarding positive selection, protein domain and glycosylation can be checked in table 4 and table 5 for coevolving sites in sensitive and resistant viruses, respectively.

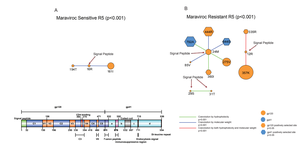


Table 6 (Sensitive)

HXB2 site	Degree	Selection	Domain	Protein	Site	N-Gly Motif
15R	2	null	SIGNAL PEPTIDE	gp120	15	
194T	7	null	V2	gp120	167	-3
181I	2	null	V2	gp120	139	

Table 7 (Resistant)

HXB2 site	Degree	Selection	Domain	Protein	Site	N-Gly Motif
30V	1	0.997736	C1	gp120	84	
275V	10	null	C2	gp120	247	
276T	10	0.998715	C2	gp120	248	NXS/TX
337A	25	null	C3	gp120	325	NXS/TX
366R	6	0.999998	C3	gp120	328	+1
444R	15	0.999052	C4	gp120	394	-4
12R	2	0.999959	SIGNAL PEPTIDE	gp120	12	
24M	6	null	SIGNAL PEPTIDE	gp120	23	
29S	1	null	SIGNAL PEPTIDE	gp120	28	
31Y	1	null	SIGNAL PEPTIDE	gp120	30	
644S	10	0.999716	COILED COIL	gp41	568	+4
792A	18	0.999976	GP41_D	gp41	736	

Fig.5 Coevolutionary networks of the signal peptide.

Conclusions

1. Coevolution between V3 loop, N-linked glycosylation motifs and the signal peptide are key to V3/gp120/gp41 folding into particular conformations that can utilize drug-bound CCR5 for cell entry. 2. Each viral population's resistance pathway is context dependent. 3. The evolutionary trajectories of the viruses are significantly constrained accounting for the low incidence of resistant R5 virus.

Acknowledgements

We would like to thank the participants and investigators in the MOTIVATE trails. The data was kindly supplied by ViiV Healthcare. XJ was supported by the MRC (G1001806/1) and WT (097820/Z/11/A). FF is supported by a BBSRC studentship to DLR. We would like to thank Paul Simpson at Pfizer R&D.

References

1. Dorr P, Westby M, Dobbs S, Griffin P, Irvine B, et al. (2005) Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. Antimicrob Agents Chemother 49: 4721-4732.
2. Westby M, Lewis M, Whitcomb J, Youle M, Pozniak AL, et al. (2006) Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1-infected patients following treatment with the CCR5 antagonist maraviroc is from a pretreatment CXCR4-using virus reservoir. J Virol 80: 4909-4920.
3. Archer J, Rambaut A, Taillon BE, Harrigan PR, Lewis M, et al. (2010) The evolutionary analysis of emerging low frequency HIV-1 CXCR4 using variants through time--an ultra-deep approach. PLoS Comput Biol 6: e1001022.
4. Westby M, Smith-Burchnell C, Mori J, Lewis M, Mosley M, et al. (2007) Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. J Virol 81: 2359-2371.