Ultradeep sequencing reveals HIV-1 diversity and resistance compartmentalization

Eleni Giatsou, Basma Abdi, Isabelle Plu, Nathalie Désiré, Romain Palich, Vincent Calvez, Danielle Seilhean, Anne-Geneviève Marcelin, Aude Jary

Title: Ultradeep sequencing reveals HIV-1 diversity and resistance compartmentalization during HIV-encephalopathy

Short title: HIV-compartmentalization in the CNS

Eleni GIATSOU[1], Basma ABDI[1], Isabelle PLU[2], Nathalie DESIRE[1], Romain PALICH[3], Vincent CALVEZ[1], Danielle SEILHEAN[2], Anne-Geneviève MARCELIN[1], Aude JARY[1]

[1]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLESP), AP-HP, Hôpital Pitié Salpêtrière, Laboratoire de Virologie, F-75013 Paris, France

[2]Sorbonne Université, APHP, Hôpital Pitié Salpêtrière, Département de Neuropathologie Raymond Escourolle, F-75013, Paris, France

[3]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLESP), AP-HP, Hôpital Pitié Salpêtrière, Service de Maladies Infectieuses et Tropicales, F-75013 Paris, France

Corresponding author: Dr Aude Jary, Virology Laboratory, CERVI, Pitié-Salpêtrière Hospital, 83 Bd de l'hôpital, 75013, Paris, France. Email Address: aude.jary@aphp.fr

ABSTRACT

Objectives: To examine viral diversity and resistance mutations in different brain areas in cases of HIV-encephalopathy.

Design: Twelve post-mortem brain areas from 3 cases of possible or certain HIV-encephalopathy were analyzed.

Methods: After amplification of the reverse transcriptase and the V3 loop region of the gp120 protein, ultradeep sequencing was performed with Illumina® technology. Phylogenetic analysis was performed with Fastree v2.1 using the generalized time-reversible (GTR) model. Identification of resistant viral variants was performed on Geneious software, according to HIV-1 genotypic drug resistance interpretation's algorithms, 2018 administered by the French Agency for Research on AIDS and Viral Hepatitis.

Results: Phylogenetic analysis revealed significant inter-regional and intra-regional diversity reflecting persistent HIV-1 viral replication in the different brain areas. Although some cerebral regions shared HIV-variants, most of them harbored a specific HIV-subpopulation reflecting HIV compartmentalization in the central nervous system. Furthermore, proportion and distribution of resistance mutations to Nucleoside and Non-Nucleoside Reverse Transcriptase Inhibitors differed among different brain areas of the same case suggesting that penetration of antiretroviral treatment may differ from one compartment to another.

Conclusions: This study, performed with a powerful sequencing technique, confirmed HIV compartmentalization in the central nervous system already shown by classical sequencing, suggesting that there are several reservoirs within the brain.

Keywords: HIV-encephalitis; compartmentalization; viral diversity; resistance mutation; ultradeep sequencing

INTRODUCTION

The Human Immunodeficiency Virus (HIV) enters the brain causing HIV-encephalopathy. The multinucleated giant cell (MGC), formed by cell-to-cell fusion of infected macrophages with microglia is the hallmark of this disease [1].

HIV-persistence in the CNS is principally due to weak penetration of antiretroviral drugs through the blood-brain-barrier [2,3]. Sanger-sequencing has shown independent evolution of drug resistance mutations to Nucleoside and Non-Nucleoside Reverse Transcriptase Inhibitors (NRTIs/NNRTIs) and protease inhibitors (PI) in different brain areas suggesting that differential drug penetration may occur among them [4]. Ultra-deep sequencing (UDS) detects minority variants that represent up to 1% of the HIV-1 population and that were incriminated for systemic therapeutic failure in treatment naïve patients [5–8]. Moreover, phylogenetic studies based on Sanger-sequencing determined brain-specific variants [9–11]. Analysis of the envelope gene in either Sanger or Single Molecule Real Time (SMRT) sequencing showed viral strains within the CNS evolving independently in different brain areas in patients who died from HIV-encephalopathy [12]. More specifically, uniquely divergent viral strains were identified in frontal, occipital, parietal, temporal lobes and basal ganglia [12–14].

In this study, we used UDS to describe HIV-diversity in the CNS by sequencing the reverse transcriptase (RT) gene and the hypervariable V3 loop region of the HIV-1 gp120 envelope protein, to detect minority resistant variants and to identify HIV-1 tropism in specimens derived from different brain areas in three HIV+ cases of HIV encephalopathy.

METHODS

Twelve post-mortem brain tissues from 3 HIV-positive cases of probable or certain HIV-encephalopathy were provided by the Raymond Escourolle Neuropathology laboratory of the Pitié-Salpêtrière Hospital. The twelve tissues represented temporal and frontal lobe, caudate nucleus, thalamus, cerebellum, medulla oblongata, substantia nigra and spinal cord. The first case (C1) concerned a 48-year-old woman, HIV-positive for ten years, treated by multiple antiretroviral therapy. The second case (C2) concerned a 38-year-old man, HIV-positive for nine years, never treated. Concerning the third case (C3), a 29-year-old woman, she received treatment but no data was available on duration and date of HIV-diagnosis. Information on clinical course, specific HIV treatment history and biological parameters was limited, as the majority of medical records have been destroyed (Supplementary Table 1).

After DNA/RNA extraction, HIV proviral-DNA was quantified with Generic HIV DNA Cell® kit (Biocentric®). RT (RT1 and RT2) and V3 loop regions were amplified by nested PCR (Supplementary Table 2) and sequencing was performed by Illumina® MiSeq (paired-end, 2x300bp).

Viral diversity. Geneious Prime software (Biomatters Ltd, Auckland, NZ) was used to keep reads with a Q-score > 30 and longer than 200bp and to pair forward and reverse reads to form complete RT1, RT2 and V3 regions. Sequences in 100% agreement were grouped to form consensus sequences (CS). Second round was performed with sequences in 99% agreement then 98% and finally 97% as previously described [15]. Then, multiple alignments of all CS and HXB2 reference genome was performed using Mafft Software v7 [16]. Phylogenetic analysis was performed using approximately-maximum-likehood method with FastTree v2.1 using generalized time-reversible (GTR) model on both all CS (HIV_RT1_CS

and HIV_V3_CS) and CS after cleaning viral CS found less than 100 times in each brain area (HIV_RT1_CS100 and HIV_V3_CS100).

To compare, Sanger sequencing was also performed according to the ANRS (French Agency for HIV research and Hepatitis) technique (http://www.hivfrenchresistance.org/). Multiple alignment of nucleotide sequences was performed with Mafft [16] and phylogenetic analysis with PhyML using GTR model and 1000 bootstrap resampling.

Finally, HIV-1 tropism was determined with geno2pheno (https://coreceptor.geno2pheno.org/) according to the recommendations of the European Consensus Group on clinical management of HIV-1 tropism testing (10% FPR).

Single-nucleotide polymorphisms (SNPs). Cleaned reads of RT1/RT2 issued from UDS (see previous paragraph) were mapped against HXB2 that carried annotations for the RT to identify SNPs (synonymous and non-synonymous SNPs, the coverage and the number of reads carrying polymorphism). The minimum variant frequency was set at 1%. Finally, HIV-1 genotypic drug resistance interpretation's algorithms, 2018 (http://www.hivfrenchresistance.org/table.html) administered by the French Agency for Research on AIDS and Viral Hepatitis were used to identify resistance mutations.

RESULTS

HIV proviral-DNA load was only detected in C2 temporal lobe and medulla oblongata specimen as well as in all C3 specimens mentioned in increasing order: cerebellum (23 copies/ 10^6 cells), medulla oblongata (31 copies/ 10^6 cells), temporal lobe (91 copies/ 10^6 cells), substantia nigra (92 copies/ 10^6 cells), caudate nucleus (130 copies/ 10^6 cells), and frontal lobe (544 copies/ 10^6 cells) **(Table 1)**.

Phylogenetic trees were generated using both HIV_RT1_CS (Supplementary Figure 1) and HIV_RT1_CS100 **(Figure 1)** and viral diversity is depicted in Table 1.

Viral diversity of RT1 for C1 was very high varying from 1086 to 3453 different viral CS in each brain area. Viral variants isolated from temporal lobe, caudate nucleus and spinal cord clustered independently **(Figure 1A)**. However, a small part of viral strains derived from caudate nucleus and spinal cord was intermingled with 15 common CS between the two areas (0.6% and 0.4% of their CS, respectively)**.** Viral diversity of C2 was also very high in temporal lobe region and medulla oblongata (3259 and 3189 respectively) with a clear separation of viral population between the two compartments and only 3 common CS (0.09%). Considering HIV_RT1_CS100, no viral population was shared between C2's compartments **(Figure 1C)**. Finally, for C3, viral variants isolated from substantia nigra clustered independently (both HIV_RT1_CS and HIV_RT_CS100) from caudate nucleus, cerebellum and frontal lobe variants. However, among the last 3 brain areas, 527 viral CS were shared (20% of frontal CS, 21.5% of caudate nucleus CS and 15% of cerebellum CS) **(Figure 1D)**.

By Sanger sequencing similar results were found, specifically sequences from C1 and C2 clustered independently. However, in C3, sequences from cerebellum, caudate nucleus and frontal lobe clustered together and these results may explain the more important proportion of common CS obtained by UDS between these 3 brain areas (Supplementary Figure 2).

The analysis of the V3 loop region for C1 also showed about thousand different viral CS per brain area with a limited number of them common between temporal lobe and spinal cord **(Figure 1B)**. HIV-1 tropism was analyzed with HIV_V3_CS100: 94% (72/77) of spinal cord CS100 and 96% (126/131) of temporal lobe CS100 were predicted to use the CCR5 co-

receptor. The remaining CS100 of the two brain areas were undetermined and none was predicted to use CXCR4 co-receptor (Supplementary Table 3).

SNPs not conferring resistance to NRTIs/ NNRTIs were found for all samples amplified for RT1. They were carried by either majority or minority variants depending on brain area and they reflected the viral diversity previously found (Supplementary Table 4).
SNPs conferring resistance to ZDV, ABC, TDF/ FTC (NRTIs) and ETR (NNRTIs) if associated to other mutations of the RT gene were found: specifically, M41L conferring resistance to NRTIs and V90L and V106I to the NNRTI **(Table 1)**. In C1, the majority of caudate nucleus's and spinal cord's variants harbored M41L (98, 4% each) and V90L (96.7% and 97.6% respectively) not found in temporal lobe. In C2, no resistance mutations were identified in neither temporal lobe nor medulla oblongata. In C3, minority variants in caudate nucleus and substantia nigra carried V90I (16.2% and 1.6% respectively). However, V106I was carried only by 1% of variants in caudate nucleus and M41L only by 2.3% of variants in substantia nigra. Finally, no resistance mutations were identified in neither frontal lobe nor cerebellum.

DISCUSSION

The CNS is an important viral reservoir of HIV and can be particularly difficult to target as a consequence of limited drug penetration [17,18]. This study is the first to use Illumina® technology to describe viral diversity and to analyze resistance mutations to NRTIs/NNRTIs in diverse areas of the CNS in three cases diagnosed with probable or certain HIV-encephalopathy. These cases concern a woman and a man with a long HIV disease

course, either treated or not respectively, as well as a treated woman for whom disease duration is unknown.

Firstly, we found that HIV proviral-DNA load varied both among different cases and among different brain areas of a single case as previously reported for HIV-RNA load in different brain regions of HIV+ cases [19]. However, the highest rates were not necessarily found in the same areas among studies and such discordance may be expected because DNA load reflects the size of the viral reservoir and not cell-free replicating virus [20]. In C1, although HIV proviral-DNA load was undetectable, sequencing and viral diversity analysis were effective in 3 of the 4 samples. This discrepancy may be explained by a higher sensibility of nested PCR compared to real-time PCR or the use of different primers between the two techniques.

Although RT1 sequencing for the C1's and C2's cerebellum and C2's thalamus specimens, RT2 for all cases and V3 for C2 and C3 failed, our results of viral diversity and tropism were consistent with those previously obtained by SMRT on an HIV+/cART+ case diagnosed with HIV encephalopathy. Indeed, the authors showed by sequencing full-length envelope gene that frontal lobe sequences clustered independently of occipital and parietal lobes and all of them were predicted to use CCR5 co-receptor while most non-brain sequences were predicted to use CXCR4 co-receptor. In our study, the majority of brain areas harbored a distinct HIV-subpopulation and those with effective V3 sequencing showed that strains used CCR5 co-receptor. While some variants isolated from caudate nucleus were intermixed to various degrees with sequences from spinal cord, frontal lobe or cerebellum region, brainstem (substantia nigra and medulla oblongata) harbored a specific subpopulation in C2 and C3. Overall, our results confirm previous evidence by Sanger-sequencing that several HIV-reservoirs exist within the CNS [13,14] and prove a high intra-regional and inter-regional viral diversity just like a study based on SMRT [12], reflecting persistent viral replication in

the CNS. Compartmentalization is evident in all of our three cases regardless of treatment status. However, in C2, who received no treatment, there is a clear separation of viral population between the two compartments examined, while in C1 and C3 who received treatment, we note some common viral strains between two regions.

HIV-1 resistance mutations to antiretroviral drugs were reported to be regionally distributed in diverse areas of the brain by classical sequencing (15). Our study, detecting minority variants up to 1% by UDS, found similar results with different distribution of resistance mutations among brain areas of the same case. These results suggest that selection pressure may vary across brain compartments and that antiretroviral treatment does not penetrate equally all of them. Finally, resistance mutations were expected in C1 and C3 who received treatment unlike C2 for whom no mutation was found.

In conclusion, this study shows significant inter-regional and intra-regional viral diversity and confirms HIV-compartmentalization in different brain areas already shown by studies based on classical sequencing suggesting that there are several reservoirs within the CNS.

ACKNOWLEDGEMENTS

Funding. This work was funded by the *Agence Nationale de Recherche sur le SIDA et les Hépatites Virale*s (ANRS) AC43 "Next-generation sequencing and resistance" working group.

Authors contributions. Conception and design of the study: IP, VC, DS, AGM; acquisition and analysis of the data: EG, BA, ND, RP, AJ; drafting of significant portion of the manuscript or figures: EG, AJ. All the authors read, corrected and approved the final manuscript.

Conflicts of interest. No conflicts of interest to disclose.

Meeting presentation. This work was presented as an e-poster at the Conference on Retroviruses and Opportunistic Infections (CROI), in March 2020 (poster no 425).

REFERENCES

1 González-Scarano F, Martín-García J. **The neuropathogenesis of AIDS**. *Nat Rev Immunol* 2005; **5**:69–81.

2 Bhaskaran K, Mussini C, Antinori A, Walker AS, Dorrucci M, Sabin C, *et al.* **Changes in the incidence and predictors of human immunodeficiency virus-associated dementia in the era of highly active antiretroviral therapy**. *Ann Neurol* 2008; **63**:213–221.

3 Robertson KR, Smurzynski M, Parsons TD, Wu K, Bosch RJ, Wu J, *et al.* **The prevalence and incidence of neurocognitive impairment in the HAART era**. *AIDS* 2007; **21**:1915–1921.

4 Smit TK, Brew BJ, Tourtellotte W, Morgello S, Gelman BB, Saksena NK. **Independent evolution of human immunodeficiency virus (HIV) drug resistance mutations in diverse areas of the brain in HIV-infected patients, with and without dementia, on antiretroviral treatment**. *J Virol* 2004; **78**:10133–10148.

5 Cozzi-Lepri A, Noguera-Julian M, Di Giallonardo F, Schuurman R, Däumer M, Aitken S, *et al.* **Low-frequency drug-resistant HIV-1 and risk of virological failure to first-line NNRTI-based ART: a multicohort European case-control study using centralized ultrasensitive 454 pyrosequencing**. *J Antimicrob Chemother* 2015; **70**:930–940.

6 Li JZ, Paredes R, Ribaudo HJ, Svarovskaia ES, Metzner KJ, Kozal MJ, *et al.* **Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis**. *JAMA* 2011; **305**:1327–1335.

7 Stella-Ascariz N, Arribas JR, Paredes R, Li JZ. **The Role of HIV-1 Drug-Resistant Minority Variants in Treatment Failure**. *J Infect Dis* 2017; **216**:S847–S850.

8 Hodkinson BP, Grice EA. **Next-Generation Sequencing: A Review of Technologies and Tools for Wound Microbiome Research**. *Adv Wound Care (New Rochelle)* 2015; **4**:50–58.

9 Epstein LG, Kuiken C, Blumberg BM, Hartman S, Sharer LR, Clement M, *et al.* **HIV-1 V3 domain variation in brain and spleen of children with AIDS: tissue-specific evolution within host-determined quasispecies**. *Virology* 1991; **180**:583–590.

10 Caragounis E-C, Gisslén M, Lindh M, Nordborg C, Westergren S, Hagberg L, *et al.* **Comparison of HIV-1 pol and env sequences of blood, CSF, brain and spleen isolates collected ante-mortem and post-mortem**. *Acta Neurol Scand* 2008; **117**:108–116.

11 Gonzalez-Perez MP, O'Connell O, Lin R, Sullivan WM, Bell J, Simmonds P, *et al.* **Independent evolution of macrophage-tropism and increased charge between HIV-1 R5 envelopes present in brain and immune tissue**. *Retrovirology* 2012; **9**:20.

12 Brese RL, Gonzalez-Perez MP, Koch M, O'Connell O, Luzuriaga K, Somasundaran M, *et al.* **Ultradeep single-molecule real-time sequencing of HIV envelope reveals complete compartmentalization of highly macrophage-tropic R5 proviral variants in brain and CXCR4-using variants in immune and peripheral tissues**. *J Neurovirol* 2018; **24**:439–453.

13 Shapshak P, Segal DM, Crandall KA, Fujimura RK, Zhang BT, Xin KQ, *et al.* **Independent evolution of HIV type 1 in different brain regions**. *AIDS Res Hum Retroviruses* 1999; **15**:811–820.

14 Chang J, Jozwiak R, Wang B, Ng T, Ge YC, Bolton W, *et al.* **Unique HIV type 1 V3 region sequences derived from six different regions of brain: region-specific evolution within host-determined quasispecies**. *AIDS Res Hum Retroviruses* 1998; **14**:25–30.

15 Nguyen T, Delaugerre C, Valantin M-A, Amiel C, Netzer E, L'yavanc T, *et al.* **Shared HCV Transmission Networks among HIV-1 Positive and Negative Men Having Sex with Men by Ultra-Deep Sequencing**. *J Acquir Immune Defic Syndr* Published Online First: 21 May 2019. doi:10.1097/QAI.0000000000002099

16 Katoh K, Standley DM. **MAFFT multiple sequence alignment software version 7: improvements in performance and usability**. *Mol Biol Evol* 2013; **30**:772–780.

17 Gray LR, Roche M, Flynn JK, Wesselingh SL, Gorry PR, Churchill MJ. **Is the central nervous system a reservoir of HIV-1?** *Curr Opin HIV AIDS* 2014; **9**:552–558.

18 Spudich S, Robertson KR, Bosch RJ, Gandhi RT, Cyktor JC, Mar H, *et al.* Persistent HIV-infected cells in cerebrospinal fluid are associated with poorer neurocognitive performance. 2019. doi:10.1172/JCI127413

19 Kumar AM, Borodowsky I, Fernandez B, Gonzalez L, Kumar M. **Human immunodeficiency virus type 1 RNA Levels in different regions of human brain: quantification using real-time reverse transcriptase-polymerase chain reaction**. *J Neurovirol* 2007; **13**:210–224.

20 Chun T-W, Murray D, Justement JS, Hallahan CW, Moir S, Kovacs C, *et al.* **Relationship between residual plasma viremia and the size of HIV proviral DNA reservoirs in infected individuals receiving effective antiretroviral therapy**. *J Infect Dis* 2011; **204**:135–138.

Table 1: HIV-1 reservoir quantification, viral diversity and resistance mutations among the different brain areas of cases 1, 2 and 3.

	CASE 1		CASE 2	CASE 3	
	Temporal lobe			**Frontal lobe**	
HIV proviral DNA (copies/10^6 cells)	<40		91	544	
HIV_RT1_CS No No shared (%)	1086 1 (0.09) with CN		3259 3 (0.09) with MO	2624 527 (20.1) with CN and Cer 55 (2.1) only with CN 83 (3.2) only with Cer	
HIV_RT1_CS100 No No shared (%)	2 0 (0)		146 0 (0)	30 7 (23.3) with CN and Cer 3 (10) only with Cer	
Resistance mutation Nucleotide substitution (position) coverage % reads carrying mutation No of reads carrying mutation	0 - - - -		0 - - - -	0 - - - -	
	Caudate nucleus		**Thalamus**	**Caudate nucleus**	
HIV proviral DNA (copies/10^6 cells)	<40		<40	130	
HIV_RT1_CS No No shared (%)	2590 15 (0.6) with SC 1 (0.04) with TL		Not amplified	2451 527 (21.5) with FL and SN 234 (9.5) only with Cer 55 (2.2) only with FL	
HIV_RT1_CS100 No No shared (%)	141 2 (1.4) with SC		Not amplified	60 7 (11.7) with FL and Cer 9 (15) only with Cer	
Resistance mutation Nucleotide substitution (position) coverage % reads carrying mutation No of reads carrying mutation	**M41 L** A-C (586) 289 139 98,4 284 513	**V90 I** G-A (733) 552 684 96,7 534 445	Not amplified	**V90 I** G-A (733) 328 900 16,2 53 282	**V106 I** G-A (781) 328 913 1 3 289
	Cerebellum				
HIV proviral DNA (copies/10^6 cells)	<40		<40	23	
HIV_RT1_CS No No shared (%)	Not amplified		Not amplified	3526 527 (14.9) with FL and CN 234 (6.6) only with CN 83 (2.4) only with FL	
HIV_RT1_CS100 No No shared (%)	Not amplified		Not amplified	73 7 (9.6) with FL and CN 9 (12.3) only with CN 3 (4.1) only with FL	
Resistance mutation Nucleotide substitution (position) coverage % reads carrying mutation No of reads carrying mutation	Not amplified		Not amplified	0 - - - -	
	Spinal cord		**Medulla oblongata**	**Substantia nigra**	
HIV proviral DNA (copies/10^6 cells)	<40		31	92	
HIV_RT1_CS No No shared (%)	3453 15 (0.4) with CN		3189 3 (0.09) with TL	1650 0 (0)	
HIV_RT1_CS100 No No shared (%)	158 2 (1.3) with CN		72 0 (0)	15 0 (0)	
Resistance mutation Nucleotide substitution (position) coverage % reads carrying mutation No of reads carrying mutation	**M41 L** A-C (586) 384 658 98,4 378 503	**V90 I** G-A (733) 618 797 97,6 603 946	0 - - - -	**M41 L** A-C (586) 62 637 2,3 1 441	**V90 I** G-A (733) 100 327 1,6 1 605

Cer: cerebellum; CN: caudate nucleus; CS: consensus sequences; DNA: desoxyribonucleic acid; FL: frontal lobe; HIV: human immunodeficiency virus; MO: medulla oblongata; No: number; SC: spinal cord; SN: substantia nigra; TL: temporal lobe; -: not applicable; No: number

HIV_RT1_CS: all cleaned consensus sequences of RT1 fragment; HIV_RT1_CS100: all cleaned consensus sequences of RT1 fragment after filtering out consensus sequences found less than 100 times

In case 1, the majority of caudate nucleus's and spinal cord's variants shared the same resistance mutations M41L and V90I. M41L: the substitution of methionine for leucine in position 41 of RT1 confers resistance to ZDV, ABC and TDF/FTC (NRTIs) on condition that this substitution is associated to two others specific mutations within the RT gene. V90I: The substitution of valine for isoleucine in position 90 of RT1 confers resistance to ETR (NNRTI) only if two others mutations are presents within the RT gene. No resistance mutation was identified in any of brain areas studied in case 2 (temporal lobe and medulla oblongata). Case 3 presented resistance mutations only in caudate nucleus and substantia nigra. V106I: The substitution of valine for isoleucine in position 106 confers resistance to ETR only if associated to two others specific mutations of the RT gene.

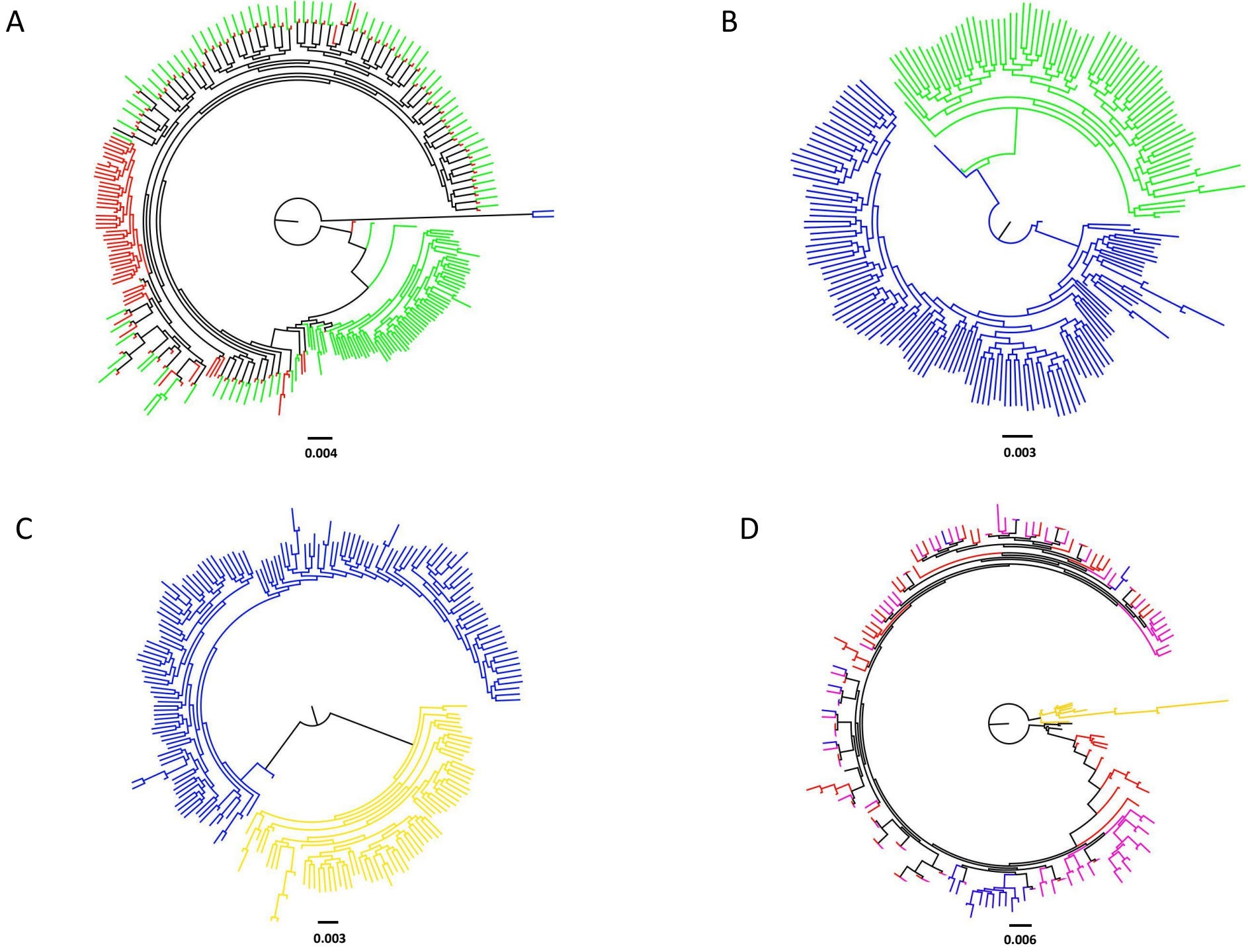
Figure 1: Approximately maximum-likelihood phylogenetic trees constructed with Fastree (2.1) of RT1 consensus viral sequences issued from the different brain areas. Phylogenetic trees were inferred with viral consensus after filtered out those found less than 100 times (HIV_RT1_CS100 or HIV_V3_CS100). **A.** RT1 of C1, **B.** V3 of C1, **C.** RT1 of C2, **D.** RT1 of C3.

Branches are colored according to the tissue origin as follow: red: caudate nucleus; blue (C1 and C2: temporal lobe, C3: frontal lobe); green: spinal cord; yellow: brainstem (C2: medulla oblongata and C3: substantia nigra); pink: cerebellum

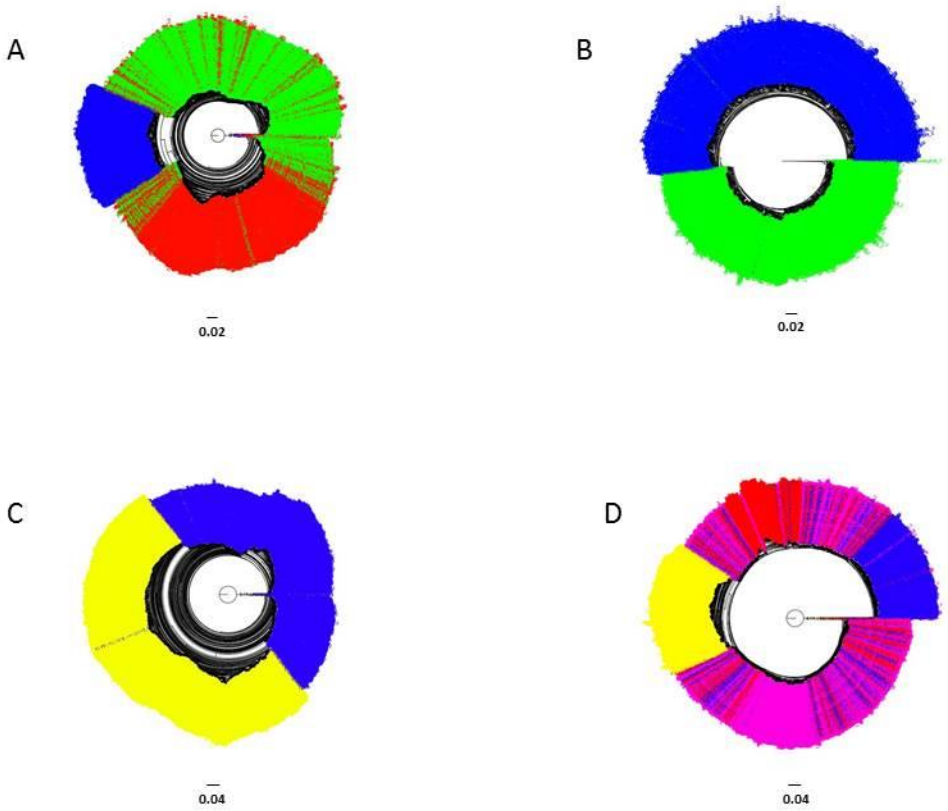
Supplementary Figure 1: Approximately maximum-likelihood phylogenetic trees constructed with Fastree (2.1) of RT1 consensus viral sequences issued from the different brain areas. Phylogenetic trees were inferred with all viral consensus found (HIV_RT1_CS or HIV_V3_CS). **A.** RT1 of C1, **B.** V3 of C1, **C.** RT1 of C2, **D.** RT1 of C3. *Branches are colored according to the tissue origin as follow: red: caudate nucleus; blue (C1 and C2: temporal lobe, C3: frontal lobe); green: spinal cord; yellow: brainstem (C2: medulla oblongata and C3: substantia nigra); pink: cerebellum*

Supplementary Figure 2: Maximum-likelihood phylogenetic tree constructed with PhyML of RT1 nucleotide sequences issued from the different brain areas. *Phylogenetic tree were inferred with nucleotide sequences generated by Sanger sequencing and rooted with HXB2 reference sequence. Branches are colored according to the tissue origin as follow: red: caudate nucleus; blue (C1 and C2: temporal lobe, C3: frontal lobe); green: spinal cord; yellow: brainstem (C2: medulla oblongata and C3: substantia nigra); pink: cerebellum*

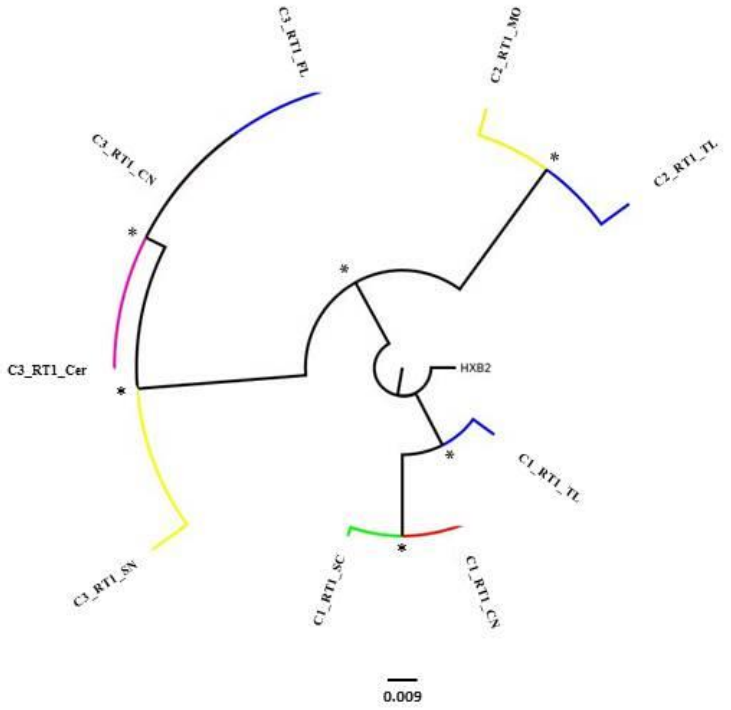
Nodes presenting a branch support > 70% (bootstrap analysis with 1000 replicates) are indicated by an asterisk.


A
0.004
B
0.003
C
0.003
D
0.006

Supplementary Figure 1: Approximately maximum-likelihood phylogenetic trees constructed with Fastree (2.1) of RT1 consensus viral sequences issued from the different brain areas. *Phylogenetic trees were inferred with all viral consensus found (HIV_RT1_CS or HIV_V3_CS). A. RT1 of C1, B. V3 of C1, C. RT1 of C2, D. RT1 of C3. Branches are colored according to the tissue origin as follow: red: caudate nucleus; blue (C1 and C2: temporal lobe, C3: frontal lobe); green: spinal cord; yellow: brainstem (C2: medulla oblongata and C3: substantia nigra); pink: cerebellum*



Supplementary Figure 2: Maximum-likelihood phylogenetic tree constructed with PhyML of RT1 nucleotide sequences issued from the different brain areas. *Phylogenetic tree were inferred with nucleotide sequences generated by Sanger sequencing and rooted with HXB2 reference sequence. Branches are colored according to the tissue origin as follow: red: caudate nucleus; blue (C1 and C2: temporal lobe, C3: frontal lobe); green: spinal cord; yellow: brainstem (C2: medulla oblongata and C3: substantia nigra); pink: cerebellum*
Nodes presenting a branch support >70% (bootstrap analysis with 1000 replicates) are indicated by an asterisk



Supplementary Table 1: Participants' characteristics and selected brain areas

Case	Age (years)	HIV diagnosis	Year of death	Death cause	cART	Selected brain areas	Anatomopathology
1	48	1996	2006	Pulmonary Embolism	Yes	Temporal lobe Caudate nucleus Cerebellum Spinal cord	Microglial activation and rare MGC positive for P24 antigen
2	38	2001	2010	Sepsis/ARDS	None	Temporal lobe Thalamus Cerebellum Medulla oblongata	Rare toxoplasma cysts without necrosis associated with numerous microglial nodules
3	29	Unknown	2007	Pulmonary Embolism	Yes	Frontal lobe Caudate nucleus Cerebellum Substantia nigra	Rare toxoplasma cysts without necrosis associated with numerous microglial nodules

ARDS: Acute respiratory distress syndrome. cART: combination antiretroviral therapy. MGC: multinucleated giant cells; HIV: human immunodeficiency virus; ART: antiretroviral therapy; MGC: multinucleated giant cells

Supplementary Table 2: Experimental conditions for the amplification by nested PCR of RT1, RT2 and V3 region before sequencing on MiSeq Illumina® system.

Amplified fragments	Primers	Direction	Sequences (5'-3')
RT1	Outer	Forward	TAG TCC TAT TGA RAC TGT ACC AGT
		Reverse	ATC CTA CAT ACA ART CAT CCA TG
	Inner	Forward	**AAG ACT CGG CAG CAT CTC** CAA TGG CCA TTG ACA GAA GAA A
		Reverse	**GCG ATC GTC ACT GTT CTC CA**T GGA ATA TTG CTG GTG ATC C
RT2	Outer	Forward	AGT CTT TTG ATG GGT CAT AAT A
		Reverse	GGG ARG TYA ATT AGG AAT ACC
	Inner	Forward	**AAG ACT CGG CAG CAT CTC** CAG ATG TGG GGA TGC ATA TTT
		Reverse	**GCG ATC GTC ACT GTT CTC CA**C TGT ATG TCA TTG ACA GTC CAG
V3	Outer	Forward	CAG TAC AAT GTA CAC ATG G
		Reverse	ATG GGA GGG GCA TAC ATT G
	Inner	Forward	**AAG ACT CGG CAG CAT CTC** CAT TAC AGT AGA AAA ATT CCC CTC
		Reverse	**GCG ATC GTC ACT GTT CTC CA**A ATG GCA GTC TAG CAG AAG

RT1 or RT2 PROTOCOL AMPLIFICATION			**V3 PROTOCOL AMPLIFICATION**		
1st PCR			**1st PCR**		
RT-PCR	50°C	30 min	RT-PCR	50°C	30 min
denaturation	94°C	7 min	denaturation	94°C	7 min
35 cycles	94°C	10 sec	35 cycles	94°C	10 sec
	55°C	30 sec		53°C	30 sec
	68°C	1 min		68°C	1 min
1 cycle	68°C	7 min	1 cycle	68°C	7 min
Nested PCR			**Nested PCR**		
denaturation	98°C	1 min	denaturation	98°C	1 min
3 cycles	98°C : 10 sec ; 66-64°C : 30sec ; 72°C : 15sec		40 cycles	98°C	10 sec
3 cycles	98°C: 10sec ; 64-62 °C: 30sec ; 72°C: 15sec			60°C	30 sec
3 cycles	98°C: 10sec ; 62-60 °C: 30sec ; 72°C: 15sec			72°C	20 sec
30 cycles	98°C: 10sec ; 60 °C: 30sec ; 72°C: 15sec		1 cycle	72°C	2 min
1 cycle	72°C	7 min		-	

Min : minute ; sec : second ; PCR : polymerase chain reaction ; RT-PCR: reverse transcriptase PCR

Universal Adapters necessary for libraries' preparation are represented in bold.

Supplementary Table 3: V3 viral diversity and tropism among two different brain area of case 1

Brain area	HIV_V3_CS		HIV_V3_CS100		HIV1-tropism		
	No of CS	Consensus sequences shared	No of CS	Consensus sequences shared	CCR5 (FPR>10%)	Undetermined (5%<FPR<10%)	CXCR4 (FPR<5%)
Temporal lobe	2086	3 (0.14) with TL	131	0	126 (96%)	5 (6%)	0
Caudate nucleus	Not amplified	-	Not amplified	-	-	-	-
Cerebellum	Not amplified	-	Not amplified	-	-	-	-
Spinal cord	1944	3 (0.15) with SC	77	0	72 (94%)	5 (4%)	0

CS: consensus sequences; DNA: desoxyribonucleic acid; FPR: false positive rate; HIV: human immunodeficiency virus; No: number; SC: spinal cord; TL: temporal lobe; -: not applicable

HIV_RT1_CS: all cleaned consensus sequences of RT1 fragment; HIV_RT1_CS100: all cleaned consensus sequences of RT1 fragment after filtering out consensus sequences found less than 100 times

Supplementary Table 4: Synonymous and non-synonymous polymorphisms not conferring resistance to Nucleoside and Non-Nucleoside Reverse Transcriptase Inhibitors in different brain areas of cases 1, 2 and 3.

CASE 1						
	TEMPORAL LOBE	**CAUDATE NUCLEUS**		**SPINAL CORD**		
Polymorphism	**D67 D**	**K65 K**	**D67 D**	**K65 K**	**D67 D**	**L74 L**
Nucleotide substitution (position)	C -T (666)	A-G (660)	C-T (666)	A-G (660)	C-T (666)	T-C (685)
Coverage (number of reads)	105 496	289 746	293 115	379 378	377 542	371 256
% Reads carrying mutation	98,5	98,1	97	97,6	98,7	1,2
Number of reads carrying mutation	103 914	284 241	284 322	370 273	372 634	4 455

CASE 2							
	TEMPORAL LOBE						
Polymorphism	**K65 E**	**D67 D**	**T69 T**	**L74 L**	**V90 V**	**A98 A**	**E138 G**
Nucleotide substitution (position)	A-G (658)	C-T (666)	T-G (672)	T-C (685)	T-C (735)	A-G (759)	A-G (878)
Coverage (number of reads)	374 271	379 027	745 373	745 392	745 393	745 393	371 453
% Reads carrying mutation	1,2	97,2	9,2	92,7	93,5	94	1,5
Number of reads carrying mutation	4 491	368 414	68 574	690 978	696 942	700 669	5 572
	MEDULLA OBLONGATA						
Polymorphism	**D67 D**	**L74 L**	**V90 V**	**A98 A**	**K101 K**		
Nucleotide substitution (position)	C-T (666)	T-C (685)	T-C (735)	A-G (759)	A-G (768)		
Coverage (number of reads)	319 846	312 409	488 146	504 794	505 950		
% reads carrying mutation	98,5	1,4	95,1	96,3	1,2		
Number of reads carrying mutation	315 048	4 374	464 227	486 117	6 071		

CASE 3								
	FRONTAL LOBE							
Polymorphism	**D67 D**	**L74 L**	**A98 A**	**L100 L**	**K101 K**	**K103 K**		
Nucleotide substitution (position)	C-T (666)	T-C (685)	A-G (759)	T-C (763)	A-G (768)	A-G (774)		
Coverage (number of reads)	215 992	212 686	364 358	364 746	364 304	361 107		
% Reads carrying mutation	98,5	97,4	96,9	97,5	98,1	97,2		
Number of reads carrying mutation	212 752	207 156	353 063	355 627	357 382	350 996		
	CAUDATE NUCLEUS							
Polymorphism	**M41 V**	**D67 D**	**T69 T**	**L74 L**	**A98 A**	**L100 L**	**K101 K**	**K103 K**
Nucleotide substitution (position)	A-G (586)	C-T (666)	T-G (672)	T-C (685)	A-G (759)	T-C (763)	A-G (768)	A-G (774)
Coverage (number of reads)	175 388	176 620	328 900	328 900	328 900	328 900	328 900	328 912
% Reads carrying mutation	1,5	98,1	7,4	94,1	96	96,8	97,2	96,2
Number of reads carrying mutation	2 631	173 264	24 339	309 495	315 744	318 375	319 691	316 413
	CEREBELLUM							
Polymorphism	**D67 D**	**L74 L**	**A98 A**	**L100 L**	**K101 K (768)**	**K103 K (774)**		
Nucleotide substitution (position)	C-T (666)	T-C (685)	A-G (759)	T-C (763)	A-G (768)	A-G (774)		
Coverage (number of reads)	286 874	282 159	486 300	486 737	486 197	482 086		
% Reads carrying mutation	98,8	97,9	97,3	98	98,4	97,5		
Number of reads carrying mutation	283 431	276 234	473 170	477 002	478 418	470 034		
	SUBSTANTIA NIGRA							
Polymorphism	**K65 K**	**D67 D**	**L74 L**	**A98 A**	**L100 L**	**K101 K**	**K103 K**	**Y115 Y**
Nucleotide substitution (position)	A-G (660)	C-T (666)	T-C (685)	A-G (759)	T-C (763)	A-G (768)	A-G (774)	T-C (810)
Coverage (number of reads)	61 055	60 803	60 047	101 772	101 687	101 451	100 209	55 050
% Reads carrying mutation	1,5	97,5	95,7	94,9	96,2	88,9	95,6	88,5
Number of reads carrying mutation	916	59 283	57 465	96 582	97 823	90 190	95 800	48 719