Virology 448 (2014) 168–175

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Deamination intensity profiling of human APOBEC3 protein activity along the near full-length genomes of HIV-1 and MoMLV by HyperHRM analysis

Kasandra Bélanger [a], Mathieu Savoie [a], Halil Aydin [a], Tyler Milston Renner [a], Zahra Montazeri [b], Marc-André Langlois [a,*]

[a] Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
[b] Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada

ARTICLE INFO

Article history:
Received 22 August 2013
Returned to author for revisions
15 September 2013
Accepted 4 October 2013
Available online 25 October 2013

Keywords:
APOBEC3
Hypermutation
Cytidine deamination
HIV
Moloney murine leukemia virus

ABSTRACT

Enzymatic deamination of cytidines in DNA is an intrinsic component of antibody maturation and retroviral resistance, but can also be a source of HIV drug resistance and cancer-causing mutations. Here, we developed a high-throughput method based on high resolution melt (HRM) analysis called HyperHRM that can screen genomic DNA for rare hypermutated proviral sequences and accurately quantify the number of C-to-T or G-to-A mutations in each sequence. We demonstrate the effectiveness of the approach by profiling in parallel the intensity of the DNA mutator activity of all seven human APOBEC3 proteins on the near full-length sequence of HIV-1 and the Moloney murine leukemia virus. Additionally, HRM was successfully used to identify hypermutated proviral sequences in peripheral blood mononuclear cells from an HIV-1 patient. These results exemplify the effectiveness of HRM-based approaches for hypermutation quantification and for the detection of hypermutated DNA sequences potentially associated with disease or retroviral drug resistance.

Introduction

Deamination of cytosine (C) into uracil (U) in DNA is a potentially mutagenic process that is fundamental to both adaptive and innate immunity (Conticello et al., 2007). Deoxycytidine deaminases such as the activation-induced deaminase (AID) and APOBEC3 proteins (A3) are enzymes that extensively convert deoxycytidines in single-stranded DNA (ssDNA) into deoxyuridines, a process called hypermutation [see (Conticello, 2008) for a review]. Because human DNA polymerases cannot distinguish U from thymidine (T) in a DNA template, adenine (A) is usually incorporated opposite U during DNA replication, resulting in G-to-A/C-to-T mutations. The physiological roles of the mutator activities of AID and A3 proteins are to promote adaptive immunity through antibody affinity maturation and innate immunity by restricting retroviral infection, respectively. However, recent studies have established strong correlations between the enzymatic activity of A3 proteins and antiretroviral drug resistance (Fourati et al., 2012; Hache et al., 2006; Jern et al., 2009; Mulder et al., 2008), and also with certain forms of cancer (Alexandrov et al., 2013; Burns et al., 2013a, 2013b; Nik-Zainal et al., 2012; Pavri and Nussenzweig, 2011; Roberts et al., 2013). For these reasons, a simple, quick and high-throughput tool for detecting and quantifying hypermutation caused by the enzymatic activities of cytidine deaminases in both the laboratory and clinic is of significant interest.

Until now, current methods were mostly designed to either screen for hypermutated sequences or conduct quantitative analyses of mutation frequency, but not both. Direct DNA sequencing of cloned proviral DNA is an approach whereby PCR is first performed on DNA isolated from an infected cell population. Amplicons are then cloned, and a small selection of clones are analyzed by DNA sequencing [see (Bélanger et al., 2013) for an application example]. Although this is a relatively simple approach for quantifying mutations in a sequence and identifying which bases are mutated, the method is not suited for large-scale screening purposes, is labor intensive because it requires cloning and plasmid purification, and does not provide high-throughput capabilities.

Next generation sequencing (NGS) is another approach that is occasionally used to analyze hypermutations in integrated retroviral DNA. It can provide direct sequence information on mutated residues and the frequency at which these mutations occur in a heterogeneous sample. This method, however, is very costly, must usually be outsourced, and the data analysis is labor intensive and requires specialized software tools. Furthermore, NGS is not a suitable approach for hypermutation screening purposes or for routine

* Correspondence to: Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, Canada K1H 8M5. Tel.: +1 613 562 5800 x7110.
E-mail address: langlois@uottawa.ca (M.-A. Langlois).

0042-6822/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2013.10.008

hypermutation analyses *in vitro*. Finally, the read lengths of NGS technologies are relatively short and are thereby not amenable for the analysis of long contiguous stretches (> 9 kb) of proviral DNA as required for near full-length single genome analyses.

For hypermutation screening purposes, there exists a very simple and powerful technique called differential DNA denaturing PCR, or commonly called 3D-PCR (Suspene et al., 2005b). 3D-PCR involves performing a first round of PCR on the gDNA extracted from a cell population, followed by a narrow gradient PCR on a short nested region (generally less than 300 bp). Amplification of experimental samples at lower melting temperatures than unmutated controls suggests the presence of A/T-rich DNA sequences that are a hallmark of A3-induced hypermutation. The greatest advantage of 3D-PCR over other methods is that it can amplify very rare hypermutated sequences in a heterogeneous sample (Suspene et al., 2011, 2005a; Vartanian et al., 2008). However, the strength of this method is also its weakness because it is naturally biased towards amplifying only the most heavily hypermutated sequences. It cannot therefore provide quantitative information on the frequency of hypermutated sequences or the mean intensity of the mutations. Another shortcoming of the method is that it is not amenable to high-throughput applications because of the numerous tubes required for each experimental sample at the gradient PCR stage. Lastly, 3D-PCR only performs well on relatively short DNA amplicons containing high numbers of mutations.

High resolution melting (HRM) is a post-PCR method that detects small nucleotide variations in a given DNA sequence in comparison to a reference. Base pair variations are measured as a function of their impact on the melting temperature of the DNA segment (Vossen et al., 2009; Wittwer, 2009). Accordingly, increasing the A/T content of a short DNA sequence, albeit by only a single base pair substitution, will decrease the melting temperature of that DNA fragment. HRM relies on dyes that specifically emit fluorescence only when bound to dsDNA. HRM analysis is performed following a quantitative PCR (qPCR) reaction by slowly melting dye-saturated PCR amplicons. Fluorescence is then measured at regular temperature intervals until the DNA strands are completely denatured and the fluorescence level reaches background. HRM is most commonly used in single nucleotide polymorphism (SNP) genotyping and DNA methylation analyses (Reed and Wittwer, 2004; Wojdacz and Dobrovic, 2007).

Here, we adapted HRM analysis combined with colony PCR and a custom analysis algorithm to identify hypermutated DNA sequences and directly evaluate the number of G-to-A/C-to-T mutations, without the need for plasmid purification or DNA sequencing. We call this approach HyperHRM. We provide application examples for our methods by analyzing the near full-length genomes of both HIV-1 and MoMLV that have been individually hypermutated by each of the seven human A3 family members. This analysis has allowed us to identify very weakly deaminated zones in the genomes of both viruses and to characterize the specific deamination profile of each A3 protein. Finally, we demonstrate the effectiveness of HRM-based approaches for identifying hypermutated proviral sequences bearing hallmarks of A3 mutator activity in the gDNA of peripheral blood mononuclear cells (PBMCs) from an HIV-1 infected patient.

Results and discussion

Here we propose two strategies that employ HRM to detect sequences with lower melting temperatures due to C-to-T/G-to-A transition mutations (Fig. 1A and Note S1). The first strategy is a direct HRM analysis performed using a single round of qPCR on the gDNA extracted from a cell population. This method is useful for detecting hypermutated sequences within a large pool of unmutated sequences. The second strategy, called HyperHRM, offers more sensitivity and resolution because sequences of interest are first cloned and then their HRM fluorescence output is measured to controls with known numbers of mutations using the algorithm defined in the Materials and Methods Section.

To validate the accuracy of mutation quantification by HyperHRM, we first cloned eGFP-derived DNA sequences containing 1 and 8G-to-A mutations in defined segments of 100 bp to 1 kb in length (Fig. 1B and Table S1). We then resolved the mathematical relationship between the fluorescence measurements taken during the HRM step and the number of mutations in each DNA amplicon of increasing length (see Materials and Methods Section). We established that there exists a robust linear correlation between fluorescence and mutation number at the temperature that maximizes the distance between the melting curves of the controls. R-square values above 0.94 were obtained for every DNA amplicon length tested below 1 kb (Fig. 1B). Furthermore, the approach is also very sensitive; a single mutation could be accurately detected in amplicons of up to 648 bp in length, while the 900 bp amplicon had an accurate detection limit of three mutations (Table 1). Considering that hypermutated sequences contain, by definition, several more C-to-T/G-to-A transition mutations than any other type of mutation, the technique is therefore sensitive enough to identify hypermutation in amplicons up to 900 bp in length. We then examined if HRM is sufficiently sensitive to serve as an effective screening tool for detecting rare hypermutated sequences among a majority of unmutated sequences. To investigate this, we performed serial dilutions of plasmid DNA (containing either 1, 3 or 8 mutations) in unmutated plasmid (Fig. 1C). Our experiment shows that fluorescence was detected above background in all conditions for dilutions up to 1/1000 (Fig. 1C and Table S3).

For comparison purposes, we challenged the sensitivity of our HRM-based approached with that of 3D-PCR (Fig. S1) (Suspene et al., 2005b). Here we tested the sensitivity of 3D-PCR on amplicons of 279 bp and 648 bp in length, containing 1, 3 or 8 G-to-A mutations (Fig. S1). Although 3D-PCR was able to detect all mutations in a short 279 bp amplicon, the method failed to provide a positive signal when 1 or 3 mutations were present in the larger 648 bp amplicon.

Having established that HyperHRM is both sensitive and accurate, we next tested its high-throughput capabilities in a standard retroviral hypermutation assay by analyzing the DNA deamination intensity of each of the seven human A3 proteins along the near full-length (~8.3 kb) proviral DNA sequence of MoMLV. We produced MoMLV reporter virus in cells that transiently expressed one of the seven human A3 proteins. For the mutation analysis, we designed an approach in which the genome of MoMLV is amplified in 5 large PCR amplicons (A–E) (Fig. 2A). HRM is then performed on nested regions within each amplicon (R1–15) (Fig. 2B). Relative fluorescence is calculated as the difference between experimental sample DNA and the unmutated control sequence. HRM hypermutation profiles revealed that APOBEC3A (A3A) had very little activity on MoMLV, as expected, and that APOBEC3G (A3G) displayed the highest mutator activity on the virus.

Considering that many cell types simultaneously express more than one A3 protein, we plotted the combined effect of all seven A3 proteins on the virus to provide a hypothetical hypermutation profile if all the A3 proteins were expressed (Fig. 2C). An intriguing result emerged with region 9 (R9), which appears to be much less targeted for deamination by all members of the A3 family. This type of deamination pattern has been observed before for A3G acting on HIV-1, where the dead zones were identified at the central polypurine tract (cPPT) and LTR proximal polypurine tract (3'PPT) of the virus (Hu et al., 2010; Suspene et al., 2006; Wurtzer et al., 2006). Although the 3'PPT is located immediately upstream of the 3'LTR for all retroviruses, a cPPT has never been described for MoMLV or for retroviruses other than lentiviruses. Sequence analysis of MoMLV R9 reveals the presence of a