Journal of Virological Methods 209 (2014) 126–135

Contents lists available at ScienceDirect

Journal of Virological Methods

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

A near-full length genotypic assay for HCV1b

Lize Cuypers[a,*], Joke Snoeck[a,1], Bram Vrancken[a], Lien Kerremans[a,2], Grégoire Vuagniaux[b], Jannick Verbeeck[a,3], Frederik Nevens[c], Ricardo J. Camacho[a,d], Anne-Mieke Vandamme[a,d], Sonia Van Dooren[a,4]

[a] Clinical and Epidemiological Virology, Rega Institute for Medical Research, Department of Microbiology and Immunology, KU Leuven, Minderbroedersstraat 10, 3000 Leuven, Belgium
[b] Debiopharm SA, Che. Messidor 5-7, P.O. Box 5911, 1002 Lausanne, Switzerland
[c] Department of Hepatology, University Hospitals Leuven, Herestraat 49, 3000 Leuven, Belgium
[d] Centro de Malária e Outras Doencas Tropicais and Unidade de Microbiologia, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Rua da Junqueira 100, 1349-008 Lisbon, Portugal

Article history:
Received 19 December 2013
Received in revised form 29 August 2014
Accepted 9 September 2014
Available online 22 September 2014

Keywords:
HCV1b
Genotyping
Full-genome sequencing
Drug resistance
NGS
Illumina

A B S T R A C T

A near-full genome genotypic assay for HCV1b was developed, which may prove useful to investigate antiviral drug resistance, given new combination therapies for HCV1 infection. The assay consists of three partially overlapping PCRs followed by Sanger population or Illumina next-generation sequencing. Seventy-seven therapy-naïve samples, spanning the entire diversity range of currently known HCV1b, were used for optimization of PCRs, of which ten were sequenced using Sanger and of these ten, four using Illumina. The median detection limits for the three regions, 5′UTR-NS2, E2-NS5A and NS4B–NS5B, were 570, 5670 and 56,670 IU/ml respectively. The number of Illumina reads mapped varied according to the software used, Segminator II being the best performing (81%). Consensus Illumina and Sanger sequencing results accord largely (0.013% major discordances). Differences were due almost exclusively to a larger number of ambiguities (presumably minority variants) scored by Illumina (1.50% minor discordances). The assay is easy to perform in an equipped laboratory; nevertheless, it was difficult to reach high sensitivity and reproducibility, due to the high genetic viral variability. This assay proved to be suitable for detecting drug resistance mutations and can also be used for epidemiological research, even though only a limited set of samples was used for validation.

1. Introduction

Chronic hepatitis caused by hepatitis C virus (HCV) is a serious public health problem causing severe disease burden in many regions of the world. To date, more than 170 million people are infected worldwide (Hanafiah et al., 2013). Epidemiological and phylogenetic studies of HCV identified six major genotypes, which originate and distribute differently around the world (Shepard et al., 2005; Simmonds et al., 2005; Magiorkinis et al., 2009). Few treatments are available, but their efficiency is often genotype dependent. The previous standard combination therapy of pegylated interferon-alpha with ribavirin is least effective in genotype 1, curing only 50–60% of the patients (Manns et al., 2001; Fried et al., 2002). Other treatment challenges were the numerous side effects of the drugs, and HIV/HCV co-infected patients. Until now there is no preventive vaccine available, but the search continues.

The need for novel potent therapeutic options, active on different targets, is high. Designing an effective treatment strategy was hampered for a long time by the lack of an *in vitro* cell culture system for HCV and the large genetic diversity of the virus

Abbreviations: HCV, hepatitis C virus; (c)DNA, (copy) deoxyribonucleic acid; PCR, polymerase chain reaction; HIV, human immunodeficiency virus; DAA, direct acting antiviral; RNA, ribonucleic acid; NGS, next-generation sequencing; IU/ml, international units per ml; RT, reverse transcriptase; dNTP, deoxyribonucleotide triphosphate; EtBr, ethidium bromide; bp, base pairs; OD, optical density; MAQ, mapping and assembly with quality; BWA, Burrows-Wheeler aligner; V-FAT, Viral Finishing and Annotation Tool; OF, outer forward; OR, outer reverse; IF, inner forward; IR, inner reverse; SF, sequencing forward; SR, sequencing reverse; nd, not determined; IQR, interquartile range.

* Corresponding author. Tel.: +32 16 332151; fax: +32 16 332131.
E-mail addresses: lize.cuypers@uzleuven.be, lize.cuypers@rega.kuleuven.be (L. Cuypers).
[1] Currently at Nuffield Department of Surgical Sciences, John Radcliffe Hospital, Headington, Oxford OX3 9DU, United Kingdom.
[2] Currently at Multiplicom N.V., Galileilaan 18; 2845 Niel, Belgium.
[3] Currently at Department of Microbiology and Immunology, KU Leuven, Minderbroedersstraat 10, 3000 Leuven, Belgium.
[4] Currently at Centre for Medical Genetics, Research Group Reproduction and Genetics, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, 1090 Brussels; Belgium and Centre for Medical Genetics, Universitair Ziekenhuis Brussel, Vrije Universiteit Brussel (VUB), Laarbeeklaan 101, 1090 Brussels, Belgium.

http://dx.doi.org/10.1016/j.jviromet.2014.09.009
0166-0934/

(Roelandt et al., 2012). Major progress was made with the discovery and development of direct acting antiviral (DAA) drugs, which specifically target HCV proteins. The first two approved protease inhibitors, boceprevir and telaprevir, have improved the treatment options in genotype 1 patients (Hézode et al., 2009; McHutchison et al., 2010; Bacon et al., 2011; Poordad et al., 2011). Recently the third protease inhibitor, simeprevir, and the first NS5B polymerase inhibitor, sofosbuvir, were approved. Several other DAAs, like NS5A inhibitors and host cyclophilin targeting agents have also entered late stage clinical trials, possibly increasing treatment options for HCV patients soon. Interferon-free regimens using a combination of DAAs are very promising (Welsch et al., 2012; Lawitz et al., 2013).

The effectiveness of DAAs is frequently hampered by the selection of viral drug-resistant mutants emerging due to the lack of proof reading by the RNA dependent RNA polymerase and a high level of replication of the virus. During therapy, and often already at baseline, resistant variants occur and raise the chance of failure to treatment. Combination therapy strategies with multiple DAAs are used to tackle drug resistance development (Gane et al., 2010; Zeuzem et al., 2012). To detect these resistance mutations simultaneously in the different viral target genes under selective pressure of a combination therapy and to study the epistatic effects of such mutations in different parts of the genome, a full-genome sequencing strategy is required. Traditionally such a genotypic test is performed by PCR amplification and capillary sequencing. However this approach is not able to reliably detect minority variants below 20%, while the more recent next-generation sequencing (NGS) methodology allows the reliable detection of variants at levels below 1% when sufficient template is available. Several platforms are already available to perform such so-called deep-sequencing, with the predominant ones being 454 sequencing (Roche Diagnostics), SOLiD sequencing (Life Technologies) and Illumina sequencing (Illumina), all with their own strengths and weaknesses (Radford et al., 2012; Loman et al., 2012). Deep sequencing is being used more and more in virus research (Daly et al., 2011), especially for the study of resistance under drug selective pressure (Ji et al., 2011).

HCV1b is one of the most prevalent subtypes and is less susceptible to the classical treatment with pegylated interferon-alpha and ribavirin. It is therefore a major target for development of antiviral drugs, and there is a need for genotypic resistance assays able to sequence the entire genome. Therefore, the goal of this study was to develop a near-full length genotypic assay suitable for Sanger and next-generation sequencing (NGS), that can be used for investigation of resistance patterns of drugs that target different viral proteins (Quiñones-Mateu et al., 2014). The assay was based on PCR amplification of three partially overlapping fragments, spanning the near full-genome, followed by sequencing. The assay was optimized to cover the currently known diversity range of HCV1b in order to ensure successful genotyping beyond the local epidemic. We wanted to evaluate and compare its performance both with conventional Sanger sequencing and with the Illumina sequencing (an NGS) methodology. The assay could be useful both for drug resistance and epidemiological studies, provided bioinformatics and financial support is available.

2. Materials and methods

2.1. Samples

The study was approved by the Leuven ethical committee (number ML9219, subproject of ML8635). For the validation of this assay 90 clinical samples from two hospitals were processed. One batch of 65 samples was obtained from treatment-naïve patients attending the University Hospital in Leuven. Another batch of 25 samples was obtained from therapy-naïve patients attending Centro Hospitalar de Lisboa Ocidental, Lisbon, Portugal. The viral load was determined using the HCV Viral Load COBAS AMPLICOR system, from Roche Molecular Diagnostics (Basel, Switzerland), or a quantitative real-time PCR assay (Beuselinck et al., 2005). For the Portuguese samples, anonymized leftovers from routine clinical practice were used to optimize this assay. Almost all samples had high viral loads, ranging from 3490 IU/ml to >10,000,000 IU/ml. In both hospitals, genotype was determined using the InnoLipa 1.0 test (Innogenetics, Zwijnaarde, Belgium). HCV1b subtype confirmation was performed by amplification and sequencing of a small region of the NS5B gene (Arrais et al., 2008), with subsequent automated HCV subtyping using the Oxford tool (De Oliveira et al., 2005; http://www.bioafrica.net/rega-genotype/html/subtypinghcv.html). Seventy-seven samples were confirmed to be HCV1b, and thus, 13 samples were excluded for the analysis of the performance of this assay, most of these proved to be HCV1a.

All 77 HCV1b samples were used to evaluate the performance of the PCR protocol. For Sanger sequencing, a subset of 10 samples were randomly selected provided they fulfilled the following selection criteria: 100% reproducible in triplicate testing, viral load above 100,000 IU/ml and sufficient plasma available. Of these 10 samples selected for Sanger sequencing, four samples with a viral load above 500,000 IU/ml, and sufficient plasma available, were selected for NGS. All sequences obtained in this study, were submitted to GenBank (accession numbers KM277568–KM277581).

Based on the viral load of the plasma samples, the units per ml (IU/ml) of HCV present in each PCR reaction can be calculated (1 IU/ml = 3 RNA copies/ml (Franciscus and Highleyman, 2012)). To assess the sensitivity of the assay, for 15 samples, previously confirmed as subtype 1b of HCV, dilution series were prepared in HCV-negative blood donor plasma, to achieve a viral load of 1,700,000 IU/ml, 170,000 IU/ml, 17,000 IU/ml, 1700 IU/ml and 170 IU/ml for all samples. The median detection limit was assessed by gel electrophoresis and expressed in IU/ml original plasma for each of the three regions.

2.2. PCR and primer design

In order to cover the full-length HCV genome, three partially overlapping assays were designed. Aligned near-full genome sequences were downloaded from the Los Alamos HCV sequence database (http://hcv.lanl.gov/content/sequence/HCV/ToolsOutline.html). Primers for cDNA synthesis, PCR and population sequencing were developed in areas with low viral diversity, and based on standard compatibility features using the OLIGO software (Medprobe, Oslo, Norway). The three nested PCRs encompassed the 5′UTR to NS2 region (2520 bp fragment), from E2 to NS5A (3787 bp fragment) and from NS4B to NS5B (3241 bp fragment) (Fig. 1). In order to increase reliability and detection of quasispecies variants, sequencing primers for both strands were developed. The primer sequences are given in Supplement Table S1. The primers were synthesized by Applied Biosystems (California, USA).

Supplementary Table S1 related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jviromet.2014.09.009.

2.3. RNA-extraction

RNA-extraction was performed using the QIAamp viral RNA mini kit from Qiagen (Westburg, Leusden, The Netherlands), using 140 *mu*l of plasma, according to the manufacturer's protocol.