

World Journal of *Gastroenterology*

World J Gastroenterol 2019 August 28; 25(32): 4567-4795



**OPINION REVIEW**

- 4567** New Era: Endoscopic treatment options in obesity—a paradigm shift
Glass J, Chaudhry A, Zeeshan MS, Ramzan Z

REVIEW

- 4580** Chronic hepatitis delta: A state-of-the-art review and new therapies
Gilman C, Heller T, Koh C
- 4598** Eosinophilic esophagitis: Current concepts in diagnosis and treatment
Gómez-Aldana A, Jaramillo-Santos M, Delgado A, Jaramillo C, Lúquez-Mindiola A
- 4614** Locoregional treatments for hepatocellular carcinoma: Current evidence and future directions
Inchingolo R, Posa A, Mariappan M, Spiliopoulos S
- 4629** Review of current diagnostic methods and advances in *Helicobacter pylori* diagnostics in the era of next generation sequencing
Pohl D, Keller PM, Bordier V, Wagner K

MINIREVIEWS

- 4661** Exploring the hepatitis C virus genome using single molecule real-time sequencing
Takeda H, Yamashita T, Ueda Y, Sekine A
- 4673** Surgical management of Zollinger-Ellison syndrome: Classical considerations and current controversies
Shao QQ, Zhao BB, Dong LB, Cao HT, Wang WB
- 4682** Positron-emission tomography for hepatocellular carcinoma: Current status and future prospects
Lu RC, She B, Gao WT, Ji YH, Xu DD, Wang QS, Wang SB

ORIGINAL ARTICLE**Basic Study**

- 4696** Effect of mild moxibustion on intestinal microbiota and NLRP6 inflammasome signaling in rats with post-inflammatory irritable bowel syndrome
Bao CH, Wang CY, Li GN, Yan YL, Wang D, Jin XM, Wu LY, Liu HR, Wang XM, Shi Z, Wu HG
- 4715** Growth arrest-specific gene 2 suppresses hepatocarcinogenesis by intervention of cell cycle and p53-dependent apoptosis
Zhu RX, Cheng ASL, Chan HLY, Yang DY, Seto WK

- 4727 Integrative analysis of the inverse expression patterns in pancreas development and cancer progression
Zang HL, Huang GM, Ju HY, Tian XF

Retrospective Study

- 4739 Prognostic value of red blood cell distribution width for severe acute pancreatitis
Zhang FX, Li ZL, Zhang ZD, Ma XC

Observational Study

- 4749 Impact of national Human Development Index on liver cancer outcomes: Transition from 2008 to 2018
Shao SY, Hu QD, Wang M, Zhao XY, Wu WT, Huang JM, Liang TB

Prospective Study

- 4764 On-treatment monitoring of liver fibrosis with serum hepatitis B core-related antigen in chronic hepatitis B
Chang XJ, Sun C, Chen Y, Li XD, Yu ZJ, Dong Z, Bai WL, Wang XD, Li ZQ, Chen D, Du WJ, Liao H, Jiang QY, Sun LJ, Li YY, Zhang CH, Xu DP, Chen YP, Li Q, Yang YP

SYSTEMATIC REVIEWS

- 4779 Liver cirrhosis and left ventricle diastolic dysfunction: Systematic review
Stundiene I, Sarnelyte J, Norkute A, Aidietiene S, Liakina V, Masalaite L, Valantinas J

ABOUT COVER

Editorial board member of *World Journal of Gastroenterology*, Tamara Vorobjova, DA, PhD, Academic Research, Department of Immunology, Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu 51014, Estonia

AIMS AND SCOPE

World Journal of Gastroenterology (*World J Gastroenterol*, *WJG*, print ISSN 1007-9327, online ISSN 2219-2840, DOI: 10.3748) is a peer-reviewed open access journal. The *WJG* Editorial Board consists of 701 experts in gastroenterology and hepatology from 58 countries.

The primary task of *WJG* is to rapidly publish high-quality original articles, reviews, and commentaries in the fields of gastroenterology, hepatology, gastrointestinal endoscopy, gastrointestinal surgery, hepatobiliary surgery, gastrointestinal oncology, gastrointestinal radiation oncology, etc. The *WJG* is dedicated to become an influential and prestigious journal in gastroenterology and hepatology, to promote the development of above disciplines, and to improve the diagnostic and therapeutic skill and expertise of clinicians.

INDEXING/ABSTRACTING

The *WJG* is now indexed in Current Contents®/Clinical Medicine, Science Citation Index Expanded (also known as SciSearch®), Journal Citation Reports®, Index Medicus, MEDLINE, PubMed, PubMed Central, and Scopus. The 2019 edition of Journal Citation Report® cites the 2018 impact factor for *WJG* as 3.411 (5-year impact factor: 3.579), ranking *WJG* as 35th among 84 journals in gastroenterology and hepatology (quartile in category Q2). CiteScore (2018): 3.43.

RESPONSIBLE EDITORS FOR THIS ISSUE

Responsible Electronic Editor: Yan-Liang Zhang

Proofing Production Department Director: Yun-Xiaojuan Wu

NAME OF JOURNAL

World Journal of Gastroenterology

ISSN

ISSN 1007-9327 (print) ISSN 2219-2840 (online)

LAUNCH DATE

October 1, 1995

FREQUENCY

Weekly

EDITORS-IN-CHIEF

Subrata Ghosh, Andrzej S Tarnawski

EDITORIAL BOARD MEMBERS

<http://www.wjgnet.com/1007-9327/editorialboard.htm>

EDITORIAL OFFICE

Ze-Mao Gong, Director

PUBLICATION DATE

August 28, 2019

COPYRIGHT

© 2019 Baishideng Publishing Group Inc

INSTRUCTIONS TO AUTHORS

<https://www.wjgnet.com/bpg/gerinfo/204>

GUIDELINES FOR ETHICS DOCUMENTS

<https://www.wjgnet.com/bpg/GerInfo/287>

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH

<https://www.wjgnet.com/bpg/gerinfo/240>

PUBLICATION MISCONDUCT

<https://www.wjgnet.com/bpg/gerinfo/208>

ARTICLE PROCESSING CHARGE

<https://www.wjgnet.com/bpg/gerinfo/242>

STEPS FOR SUBMITTING MANUSCRIPTS

<https://www.wjgnet.com/bpg/GerInfo/239>

ONLINE SUBMISSION

<https://www.f6publishing.com>



Exploring the hepatitis C virus genome using single molecule real-time sequencing

Haruhiko Takeda, Taiki Yamashita, Yoshihide Ueda, Akihiro Sekine

ORCID number: Haruhiko Takeda (0000-0002-8954-9133); Taiki Yamashita (0000-0002-5465-534X); Yoshihide Ueda (0000-0003-3196-3494); Akihiro Sekine (0000-0003-3313-4331).

Author contributions: Takeda H and Yamashita T contributed to literature review and drafting of the manuscript; Ueda Y and Sekine A contributed to critical revision and editing of the manuscript; all authors approved the final version of the manuscript.

Conflict-of-interest statement: No potential conflicts of interest. No financial support was received for this work.

Open-Access: This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

Manuscript source: Invited manuscript

Received: May 17, 2019

Peer-review started: May 17, 2019

First decision: June 16, 2019

Revised: July 4, 2019

Accepted: July 19, 2019

Haruhiko Takeda, Taiki Yamashita, Akihiro Sekine, Department of Omics-based Medicine, Center for Preventive Medical Science, Chiba University, Chiba 260-0856, Japan

Haruhiko Takeda, Yoshihide Ueda, Department of Gastroenterology and Hepatology, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan

Corresponding author: Akihiro Sekine, PhD, Professor, Department of Omics-based Medicine, Center for Preventive Medical Science, Chiba University, 1-8-1, Inohana, Chuo-ku, Chiba 260-0856, Japan. sekine.akihiro@chiba-u.jp

Telephone: +81-226-2537

Abstract

Single molecular real-time (SMRT) sequencing, also called third-generation sequencing, is a novel sequencing technique capable of generating extremely long contiguous sequence reads. While conventional short-read sequencing cannot evaluate the linkage of nucleotide substitutions distant from one another, SMRT sequencing can directly demonstrate linkage of nucleotide changes over a span of more than 20 kbp, and thus can be applied to directly examine the haplotypes of viruses or bacteria whose genome structures are changing in real time. In addition, an error correction method (circular consensus sequencing) has been established and repeated sequencing of a single-molecule DNA template can result in extremely high accuracy. The advantages of long read sequencing enable accurate determination of the haplotypes of individual viral clones. SMRT sequencing has been applied in various studies of viral genomes including determination of the full-length contiguous genome sequence of hepatitis C virus (HCV), targeted deep sequencing of the HCV NS5A gene, and assessment of heterogeneity among viral populations. Recently, the emergence of multi-drug resistant HCV viruses has become a significant clinical issue and has been also demonstrated using SMRT sequencing. In this review, we introduce the novel third-generation PacBio RSII/Sequel systems, compare them with conventional next-generation sequencers, and summarize previous studies in which SMRT sequencing technology has been applied for HCV genome analysis. We also refer to another long-read sequencing platform, nanopore sequencing technology, and discuss the advantages, limitations and future perspectives in using these third-generation sequencers for HCV genome analysis.

Key words: Third generation sequencing; PacBio RSII; Single molecule real-time sequencing; Hepatitis C virus; Resistance-associated substitution; Nanopore sequencer

Article in press: July 19, 2019
Published online: August 28, 2019

P-Reviewer: Gao YT, Lei YC
S-Editor: Yan JP
L-Editor: A
E-Editor: Qi LL



©The Author(s) 2019. Published by Baishideng Publishing Group Inc. All rights reserved.

Core tip: Single molecular real-time (SMRT) sequencing, also called third-generation sequencing, is a novel sequencing technique capable of generating extremely long contiguous sequence reads. The advantages of long read sequencing enable accurate determination of the haplotypes of infected viral clones. We introduce the novel third-generation sequencers PacBio RSII/Sequel systems, compare them with conventional next-generation sequencers, and summarize previous studies in which SMRT sequencing technology has been applied for hepatitis C virus genome analysis.

Citation: Takeda H, Yamashita T, Ueda Y, Sekine A. Exploring the hepatitis C virus genome using single molecule real-time sequencing. *World J Gastroenterol* 2019; 25(32): 4661-4672
URL: <https://www.wjgnet.com/1007-9327/full/v25/i32/4661.htm>
DOI: <https://dx.doi.org/10.3748/wjg.v25.i32.4661>

INTRODUCTION

Anti-hepatitis C virus (HCV) therapy has drastically improved over the last decade^[1]. The development of oral direct-acting antivirals (DAAs) has enabled the majority of HCV-infected patients to achieve sustained virologic response (SVR)^[2-13]. However, drug resistance-associated substitutions (RASs) including NS5A-P32del have been reported as one of the major causes of DAA treatment failure^[14-28]. A subset of patients are difficult to treat with DAAs, such as patients with decompensated liver cirrhosis or immunosuppressed patients following liver transplantation. These patients are more likely to experience DAA treatment failure^[1,29,30]. Thus, to achieve complete eradication of HCV, a more detailed understanding is needed of the hepatitis virus genome, especially of genetic alterations related to multidrug resistance.

Sequencing technology has made drastic progress in recent years^[31,32]. Sanger sequencing of hepatitis viruses has been broadly applied in real-world clinical practice, mainly to predict the efficacy of antiviral therapy. Sanger sequencing can determine the major viral haplotypes present, but cannot detect low-abundance haplotypes which may have acquired RASs. By contrast, recently developed next-generation sequencing (NGS) instruments can generate sequence reads with much higher throughput compared to Sanger sequencing. These instruments can also detect rare nucleotide changes in variants at frequencies of less than 1% (Figure 1)^[33]. Many genetic analyses of the hepatitis virus genome have been conducted using NGS over the last decade, and the utility of these methods has been validated in multiple studies worldwide. NGS has enabled detection of rare viral variants in the sera of individuals infected with HCV, analysis of the dynamics of drug-resistant variants in chronically HCV-infected patients, and even prediction of clinical outcomes such as responses to anti-HCV drugs.

Conventional NGS instruments including the Illumina Miseq, Illumina Hiseq and Ion Torrent sequencers have several serious limitations such as short read lengths (approximately 400 bp) and amplification biases. These factors restrict our ability to understand the landscape of the HCV genome^[31]. One of the central limitations of conventional NGS techniques for genomic analysis of multi-drug resistant viral clones is its short-read nature. HCV has a single-stranded, 9-kbp RNA genome, and several variants associated with drug resistance are distributed over a 3-kbp region from the NS3 to the NS5A genes. Although conventional NGS can be used to evaluate the frequencies of variants present in a sample, short-read NGS of viral genomes cannot assess the linkage of nucleotide variants located at sites distant from one another in a single viral clone (haplotype). As a result, short-read NGS technologies cannot completely evaluate the population of multi-drug resistant viral clones, despite multi-drug resistance being closely related to relapse during or after anti-viral therapy.

Recently, third-generation sequencing (TGS) platforms based on single molecular real-time (SMRT) technology have been developed. These platforms can generate extremely longer DNA sequences with high accuracy^[31,34-36], providing us the means to obtain continuous long sequence reads from single viral clones. SMRT sequencing can be accomplished using the PacBio system (Pacific Biosciences), and several studies of viral genomes using this instrument have already been reported^[32,37]. In this article, we review previous reports of HCV genetic analysis using PacBio sequencing and summarize the advantages and promise of this instrument in comparison with other

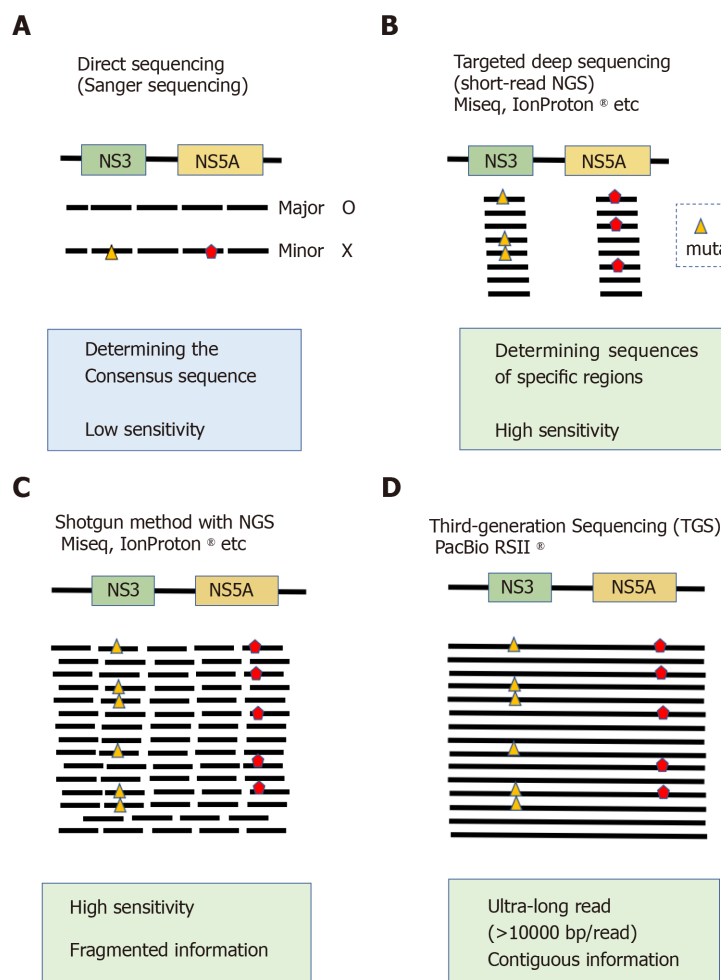


Figure 1 Comparison of sequencing platforms. A: Direct sequencing (Sanger sequencing). This conventional sequencing method determines the consensus sequence of target regions. Nucleotide variants with allele frequencies of approximately 15% can be detected; B: Targeted deep sequencing using conventional short-read next-generation sequencing (NGS) can detect low abundance variants making up approximately 1% of total mapped reads; C: When long PCR products are used as templates for conventional short-read NGS, they are first fragmented into 100-200 bp segments, ligated to sequence adapters, amplified and then sequenced. The sequenced reads are mapped to a reference sequence using the shotgun method. One of the limitations of this technique is a lack of information regarding whether two distant mutations co-exist on a single template molecule; D: Third-generation sequencing methods represented by single-molecular real-time sequencing can generate ultra-long reads of more than 10000 bp, and contiguous sequence information can be obtained. NGS: Next-generation sequencing.

NGS platforms.

SINGLE MOLECULE REAL-TIME SEQUENCING

PacBio sequencing

Recently, novel sequencing technologies capable of generating extremely long reads have been developed. As a group, these technologies have often been called TGS platforms^[31]. One of the major TGS technologies is SMRT sequencing using the PacBio RSII or Sequel sequencers (Pacific Biosciences)^[34,36]. SMRT sequencing is a sequencing-by-synthesis technology based on real-time imaging of fluorescently-tagged nucleotides simultaneously with the synthesis along individual DNA template molecules. Because the reaction is driven by a DNA polymerase, and because single molecules are imaged in this technology, there is no degradation of signal over time. The sequencing reaction continues until the template and polymerase dissociate. The average sequencing read length from the current PacBio RSII instrument is about 12 kbp, and the newly-released PacBio Sequel sequencer can generate reads longer than 20 kbp on average. The longest sequence reads produced by the instrument exceed 50 kbp. These reads are about 200 times longer than those generated by conventional

NGS instruments. Taking the advantage of ultra-long read length into consideration, SMRT sequencing has been applied to various research questions such as determination of the full-length genome sequences of bacteria, metagenome analyses of viruses or bacteria, haplotype determination of genes, transcriptome analyses of splicing variants, and determination of the full-length human genome sequence^[32,38-41].

Circular consensus sequencing

The single-pass error rate for raw long reads generated by PacBio sequencing is as high as 11%-15%, with indel errors dominating. Thus, several error correction methodologies have been developed. One of the most commonly-used error correction methods is circular consensus sequencing (CCS) (Figure 2)^[42,43].

The template for PacBio sequencing, called SMRTBell, is created by ligating hairpin adaptors to both ends of the prepared DNA templates, that is, double-stranded DNA molecules including PCR amplicons. The template then acts as a single-stranded closed circle loop. The enzyme initiates sequencing reaction at the specific region of hairpin adaptor (identical in all SMRTBells) and sequences the template until the polymerase loses activity. The enzyme proceeds around the hairpin on the other end of the SMRTBell and can traverse a single DNA template multiple times. Then, error-corrected consensus reads (CCS reads) are generated using data from a single template sequenced multiple times. For example, CCS reads resulting from five polymerase passes around closed loop SMRTbells are defined as "5-pass CCS reads". Based on the random error nature of SMRT sequencing, more passes result in higher accuracy of the consensus reads; the error rate of 5-pass CCS reads is as low as 0.1% per base and that of 10-pass CCS reads is less than 0.03%^[44].

To generate 10-pass CCS reads from a 3000-bp DNA template, raw reads longer than 30000 bp should be generated by SMRT sequencing. The average length of raw reads generated by the PacBio RSII instrument is approximately 12000 bp, and typically as few as 5% of the raw reads are longer than 30000 bp. Thus, most of the sequenced reads would be excluded from the final analysis. To avoid this limitation, decreasing the template length or the pass number cutoff for analysis can be considered, although the advantages of SMRT sequencing are limited in turn. Considering that longer sequencing can generate more accurate consensus reads, improvement of sequencing cells, reagents or instruments for SMRT sequencing is expected in the future.

Nanopore sequencing

Another technology for single-molecular real-time long read sequencing is nanopore sequencing (Oxford Nanopore Technologies, Oxford, United Kingdom), which is often compared with the PacBio sequencing platform^[45]. As the throughput of PacBio RS II has been somewhat limited and its running costs has been so high, many smaller laboratories have not been able to take advantage of this instrument. In 2014, the first device of a nanopore sequencer (the MinION) became available, and was immediately attractive to smaller laboratories due to its low costs and small size. Unlike other platforms, nanopore sequencers do not use base synthesis reaction in the sequencing process, which differs fundamentally from other sequencing technologies including PacBio sequencing. Instead, nanopore sequencers directly detect the sequence of the nucleotides composing a native single stranded DNA molecule *via* changes in electronic voltage as it passes through a protein pore.

This sequencing process generates 1D and 2D reads in which both "1D" strands can be aligned to create a consensus sequence "2D" read. The 1D raw reads have error rates of more than 10%, similar to PacBio raw reads. Although the error rates of 2D reads are somewhat improved, these are still higher than the consensus reads generated by PacBio sequencing. As a methodology for error correction of nanopore sequencing platform has not yet been established^[46], its utility to detect SNVs in viral genome is limited in its current form. One group recently used nanopore sequencing to detect HCV genomes in patient sera and demonstrated that despite an error rate as high as 20%, genotypes could still be determined. Despite these limitations, nanopore sequencing is expected to be used for clinical sequencing in the future because of its low costs, USB power requirements, handheld use, and real-time processing capacity. Thus, not only improvements of instruments but also establishment of bioinformatic error correction methods is expected.

HCV GENOME SEQUENCING USING NEXT-AND THIRD-GENERATION SEQUENCERS

Targeted deep sequencing of viral genomes using conventional next-generation

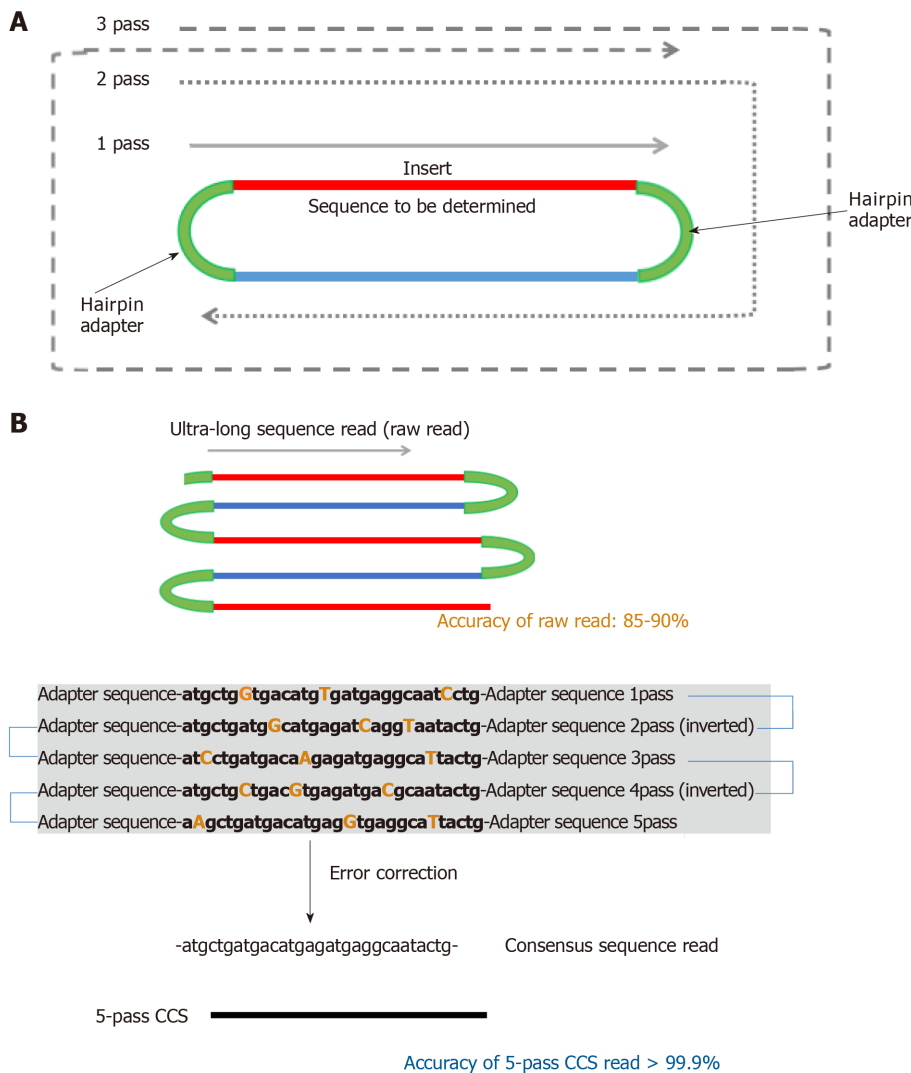


Figure 2 Generation of circular consensus sequences. A: The template for PacBio sequencing, called SMRTBell, is created by ligating hairpin adaptors to both ends of a double-stranded DNA molecule containing the sequence to be determined. This template then acts like a single-stranded closed circle. The polymerase initiates at the primer location and sequences the template until it falls off. The enzyme then proceeds around the hairpin on the other end of the SMRTBell, and can circle around the same template multiple times; B: Scheme for generation of 5-pass circular consensus sequences (CCS) reads. Ultra-long raw reads are generated by a polymerase. Although the accuracy of the raw read is 85%-90%, error-corrected consensus reads (CCS reads) can be generated using the data from a single template sequenced multiple times. The accuracy of 5-pass CCS reads is as high as 99.9%. CCS: Circular consensus sequences.

sequencing

Conventional NGS instruments, represented by the Illumina HiSeq or ThermoFisher IonTorrent systems, are characterized by their short reads. Generally, short DNA fragments 100-400 bp in length are sequenced in massively parallel fashion^[31,47]. One common NGS application is amplicon sequencing. In this application, a short region including sequences of interest (such as NS5A-aa93 in the HCV genome) is amplified by RT-PCR, followed by library generation and sequencing. Another application is based on the shotgun sequencing technique. For example, 9 kbp of the HCV genome is first amplified using long-range RT-PCR and then the long amplicons are sheared into shorter 100-200 bp fragments. Sequencing libraries are generated using the short fragments and reads are finally mapped to a reference sequence such as the full-length HCV genome. One of the advantages of conventional NGS is its ability to generate enormous amounts of sequenced nucleotide data in less time compared with Sanger sequencing. Therefore, conventional NGS has been widely applied to examine viral quasispecies and the dynamics of viral genomes^[48,49].

Using a conventional NGS method, Nasu *et al.*^[48] detected various sorts of low-abundance viral clones associated with drug resistance and characterized their dynamics in a variety of clinical settings in patients infected with HCV. This technique also enabled the discovery that various resistance-associated nucleotide alterations naturally pre-existed in treatment-naïve HCV positive patients. Sato *et al.*^[49]

conducted targeted deep sequencing of the NS3 region of HCV using serum samples obtained before and after anti-HCV therapies. They compared the sequences of an approximately 450 bp segment of the HCV NS3 region and evaluated the evolution of variants resistant to interferon-based protease inhibitor therapy using phylogenetic analysis. Teraoka *et al*^[50] compared serum samples collected from chronic HCV patients before and after oral DAA therapy and demonstrated that multidrug resistant viral clones frequently emerge at the point of treatment failure. In this manner, targeted deep sequencing of the HCV genome using conventional NGS has been widely applied and its ability to detect rare variants has been well established. After validating the reliability of variant detection by deep sequencing, targeted deep sequencing has even been applied in clinical trials of anti-HCV drugs^[9].

As described above, targeted sequencing using conventional short-read NGS yields only fragmented information such as the relative frequencies of particular mutations in NS3 or NS5A. Conventional NGS cannot establish linkage between distant mutations in the NS3 to NS5A region in individual viral clones due to its short reads. Thus, conventional NGS platforms are limited in their ability to provide information regarding multi-drug resistance (Figure 2).

Targeted deep sequencing using SMRT sequencing

HCV has a single--stranded RNA genome encoding a total of 10 proteins. Nonstructural proteins including NS3, NS5A and NS5B are essential for viral replication and have been identified as the targets of DAAs. Although the therapeutic effects of DAAs are excellent, several nucleotide changes within NS regions are associated with drug resistance. In particular, nucleotide substitutions in the NS5A region are clinically important and a number of studies have used various NGS platforms for targeted deep sequencing of the NS5A region^[18,19,23,24,26,50].

Targeted deep sequencing of the NS5A region of the HCV genome using the PacBio RSII platform was first reported in 2015. Bergfors *et al*^[51] generated CCS reads from 626-bp PCR amplicons covering the NS5A region (including aa25 to aa93). The templates were prepared from 10 sera, including seven GT 1a samples, three GT 3a samples and a control plasmid.

The authors first examined the error rate by sequencing the H77 GT 1a control plasmid at NS5A aa25-95 sites and analyzing copy number, and found a mean error rate of approximately 0.05%-0.25%. The pass number of CCS reads was not noted in this analysis. They found that PacBio SMRT sequencing permitted detection of very low frequencies (as low as 0.24%) of potentially resistant HCV variants in the NS5A region. These data suggested that the detection rate of rare mutations by PacBio SMRT sequencing might be similar or even superior to that of conventional short-read sequencers.

Full-length genome determination by SMRT sequencing

Bull *et al*^[52] described SMRT sequencing of amplicons nearly spanning the full-length HCV genome. In this report, CCS reads generated by PacBio RS II sequencing with a minimum of two passes of the full-length HCV amplicon (reads longer than 18 kb) were selected for analysis. The authors compared the sensitivity of PacBio sequencing to detect low frequency mutations with data generated by Illumina sequencing of the same amplicon. Both sequencing platforms detected all SNVs at frequencies of > 7%. However, PacBio reads only detected 4.2% of the SNVs with frequencies < 7% detected using the Illumina platform. This finding is consistent with a previous report by Jiao *et al*^[42]. These authors conducted a benchmark study of the accuracy of CCS reads generated by the PacBio sequencer and found that the Phred-like Quality Value of 2-pass CCS reads was quite low. Thus, for accurate detection of rare SNVs, a higher pass number is needed. In order to generate 5-pass CCS reads of the full-length HCV genome (9.2 kbp), ultra-long raw reads at least 46 kb in length (five times 9.2 kb) should be sequenced. Such enormously long reads can only be rarely generated even by PacBio sequencing. Thus, accurate contiguous sequencing of the full-length HCV genome is considered a major challenge. The recently launched PacBio Sequel platform, which is reported to generate longer reads than the PacBio RSII, or improvements in sequencing polymerases might overcome this problem in the near future.

Evaluation of viral heterogeneity

To date, heterogeneity of various viral populations including HCV has been evaluated using conventional short-read sequencers. Recently, PacBio sequencing has been also applied for analysis of viral quasispecies. For example, Ho *et al*^[53] evaluated the sequence diversity of a 1680 nucleotide-long HCV envelope genome region in individuals belonging to a cluster of sexually-transmitted cases using PacBio sequencing. Using 7-pass CCS reads, they reported an error rate of 0.37%.

We evaluated heterogeneity within the NS regions of the HCV genome in treatment-naïve HCV patients using the PacBio RS II platform^[44]. For this purpose, we first performed control sequencing using a plasmid containing HCV genome as a template. We amplified its NS3/4 and NS5A regions and generated the double-stranded DNA templates for SMRT sequencing. To ensure high accuracy of sequence reads, 10-pass CCS reads were strictly selected. The average mismatch error rate of 10-pass CCS reads was 0.0287% per bp, indicating that the SMRT sequencing platform achieved extremely high accuracy sequence reads. Using these high-quality CCS reads, we applied this sequencing platform to clinical serum samples. When sequence reads were aligned to each reference sequence of HCV genome, the coverage curve showed a uniform distribution across every nucleotide position compared with the coverage curve obtained from short-read sequencing platforms. Thus, the SMRT sequencing platform can provide information on heterogeneity at each nucleotide position without positional bias (Figure 3). We performed phylogenetic analysis and found that HCV clones from chronically-infected individuals were widely distributed and that individual viral clones identified in each sample showed sequence diversity. Long-read sequencing revealed that none of the viral clones present in each individual's serum had completely identical sequences through the NS3, 4A/B, and 5A regions^[44]. This finding is reasonable considering that the HCV genome is replicated by the error-prone NS5B polymerase and thus HCV clones can easily and quickly accumulate genetic mutations in their RNA genomes.

Dynamics of multi-drug resistant HCV clones

One of the most important advantages of long-read sequencing with the PacBio RSII platform is its ability to determine the haplotypes of individual viral clones (Figure 4). In the era of DAA therapy for HCV, several RASs have been identified worldwide. These RASs are mainly present in the genes encoding NS3 and NS5A. For example, NS5A-Y93H is associated with ledipasvir or daclatasvir resistance and NS3-D168V is associated with simeprevir resistance. Co-occurrence of some of these RASs in single viral clones has been reported to be associated with high rates of DAA treatment failure. Thus, evaluating the co-occurrence of RASs in NS3 and NS5A is critically important. The distance between aa168 of NS3 and aa93 of NS5A is approximately 3 kbp, and conventional short-read sequencing cannot determine linkage between RASs at NS3-aa168 and NS5A-aa93 within a single viral genome. In contrast, ultra-long read sequencing using the SMRT sequencing platform with PacBio RSII instrument can generate long contiguous sequence reads and overcome this limitation of short-read sequencers.

Long-read sequencing using the PacBio RSII can be used to evaluate not only linkage of RASs but also to analyze all synonymous nucleotide changes. Using paired serum samples collected before and after DAA treatment, we compared the haplotypes of individual viral clones and assessed the clonal evolution of HCV during DAA therapy. For this purpose, we focused on synonymous nucleotide changes linked with a given RAS such as NS5A-Y93H^[44]. First, long contiguous sequences for individual viral clones present in 12 serum samples from 6 non-SVR patients (a total of more than 3000 clones) were sequenced using SMRT sequencing technology. Subsequently, all nucleotide substitutions in each viral clone before and after treatment were identified and then compared, and we found significant linkage between several synonymous nucleotide changes and major RASs. For example, several synonymous mutations were linked to NS5A-Y93H, one of the major RASs, in a subpopulation of pre-existing viral clones at baseline, and these synonymous mutations were shared by multi-drug resistant viral clones at viral breakthrough. Phylogenetic analyses revealed that pre-existing low-abundance drug-resistant clones and multi-drug resistant viral clones at viral breakthrough were genetically close each other. In addition, linkage analysis demonstrated that multiple RASs developed *de novo* based on pre-existing drug-resistant clones following DAA treatment in non-SVR cases. Long-read sequencing using the PacBio platform enabled us to compare the haplotypes of individual HCV clones and to estimate the origins and evolution of multi-drug resistant HCV clones during anti-HCV treatment.

PacBio long-read sequencing was also applied for analysis of multi-drug resistant clones of other virus species. Huang *et al*^[37] examined linkage between six loci related to drug resistance of human immunodeficiency virus (HIV). They compared the drug resistance profiles of each HIV clone at two time points. The study examined a patient infected with HIV whose plasma viral load of HIV increased suddenly within one month of treatment from approximately 3000 copies/mL to approximately 30000 copies/mL. The authors found that rare viral populations with multi-drug resistant haplotypes identical to those of the major clones at the time point of relapse were already present at the pretreatment time point. They hypothesized that drug-resistant haplotypes had already existed as minor species in the viral population at the

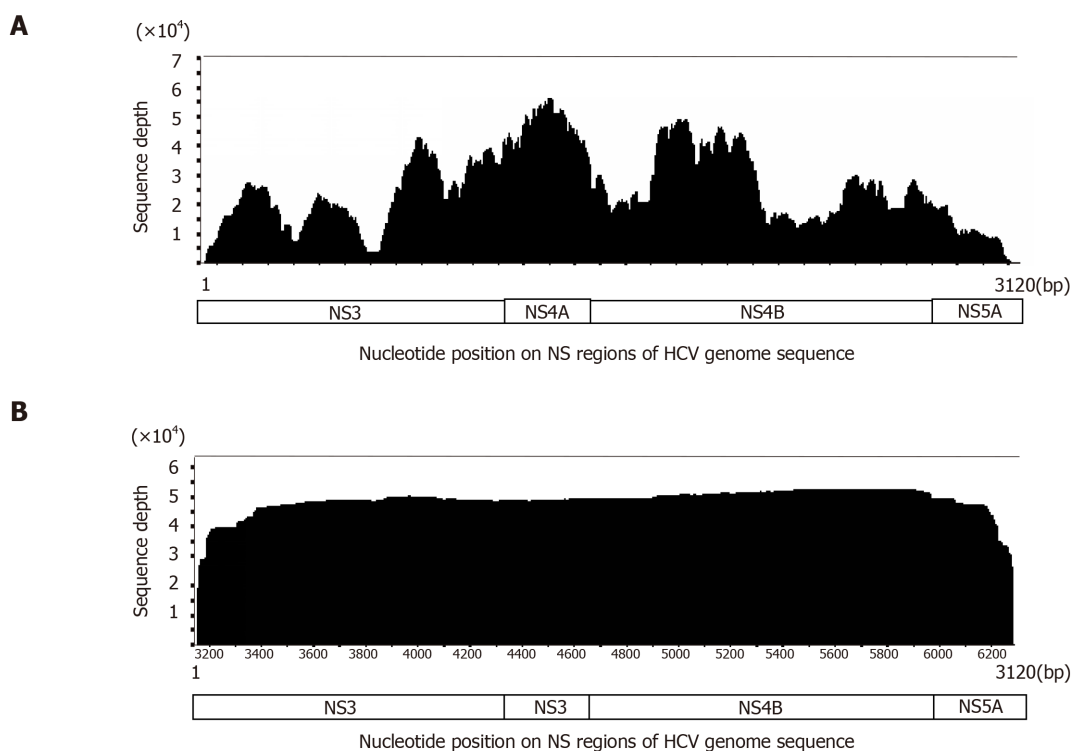


Figure 3 Comparison of coverage curves generated by short-read next-generation sequencing and long-read single-molecular real-time sequencing. A: A coverage curve generated by an IonProton sequencer. Approximately 3120 bp from the NS3 to NS5A region of the hepatitis C virus (HCV) genome from an HCV-infected patient was amplified and long-PCR products were subjected to short-read sequencing. The sequencing depth varies according to genomic location; B: When the same template was sequenced using PacBio RSII sequencer, the coverage curve demonstrates uniform coverage through the NS3 to NS5A regions. HCV: Hepatitis C virus.

pretreatment time point and that under the selective pressure of anti-viral therapy, were quickly selected for and became dominant. Thus, long CCS reads generated by PacBio RSII sequencing can reliably provide data for HIV quasispecies-level analysis.

CONCLUSION

SMRT long read sequencing technologies represented by the PacBio platform have opened a new era for genetic analysis of viruses. Using these sequencers, long contiguous sequences have been determined, linkage between distant SNVs can be analyzed and viral quasispecies can be analyzed in more detail than permitted by previous sequencing methods. Haplotype data generated by PacBio sequencing can be used to analyze clonal evolution of viral genomes and viral dynamics in clinical settings. Error correction methods for long reads of the HCV genome should be applicable for analysis of other viruses including hepatitis B virus, HIV or other pandemic viruses. In addition, the newly-launched PacBio Sequel System, which reportedly has seven-fold higher throughput than the RS II, is expected to further enable long-read sequencing analyses of viral genomes.

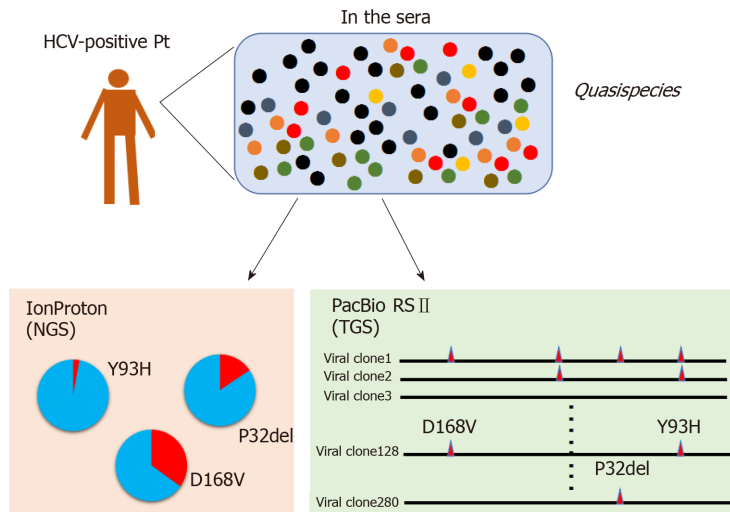


Figure 4 Comparison of short-read and long-read sequencing for analysis of viral quasispecies. Conventional short-read sequencing, such as the IonProton sequencer, generates bulk information on viral clones. However, only fragmented information can be obtained such as the frequency of viral clones bearing the NS3-D168V or NS5A-P32del variants. By contrast, PacBio RSII sequencing can determine the contiguous genome sequence of each template, permitting analysis of linkage between several nucleotide changes through the NS3 to NS5A regions for individual viral clones. TGS: Third-generation sequencing; NGS: Next-generation sequencing; HCV: Hepatitis C virus.

ACKNOWLEDGEMENTS

We thank Drs. Marusawa H, Takai A, Takahashi K, Ohtsuru S, Matsumoto T, Inuzuka T, Nakamura F and Arasawa S for helpful advice.

REFERENCES

- 1 **Falade-Nwulia O**, Suarez-Cuervo C, Nelson DR, Fried MW, Segal JB, Sulkowski MS. Oral Direct-Acting Agent Therapy for Hepatitis C Virus Infection: A Systematic Review. *Ann Intern Med* 2017; **166**: 637-648 [PMID: 28319996 DOI: 10.7326/M16-2575]
- 2 **Chayama K**, Takahashi S, Toyota J, Karino Y, Ikeda K, Ishikawa H, Watanabe H, McPhee F, Hughes E, Kumada H. Dual therapy with the nonstructural protein 5A inhibitor, daclatasvir, and the nonstructural protein 3 protease inhibitor, asunaprevir, in hepatitis C virus genotype 1b-infected null responders. *Hepatology* 2012; **55**: 742-748 [PMID: 21987462 DOI: 10.1002/hep.24724]
- 3 **Kumada H**, Chayama K, Rodrigues L, Suzuki F, Ikeda K, Toyoda H, Sato K, Karino Y, Matsuzaki Y, Kioka K, Setze C, Pilot-Matias T, Patwardhan M, Vilchez RA, Burroughs M, Redman R. Randomized phase 3 trial of ombitasvir/paritaprevir/ritonavir for hepatitis C virus genotype 1b-infected Japanese patients with or without cirrhosis. *Hepatology* 2015; **62**: 1037-1046 [PMID: 26147154 DOI: 10.1002/hep.27972]
- 4 **Kumada H**, Suzuki Y, Ikeda K, Toyota J, Karino Y, Chayama K, Kawakami Y, Ido A, Yamamoto K, Takaguchi K, Izumi N, Koike K, Takehara T, Kawada N, Sata M, Miyagoshi H, Eley T, McPhee F, Damokosh A, Ishikawa H, Hughes E. Daclatasvir plus asunaprevir for chronic HCV genotype 1b infection. *Hepatology* 2014; **59**: 2083-2091 [PMID: 24604476 DOI: 10.1002/hep.27113]
- 5 **Hayashi N**, Izumi N, Kumada H, Okanoue T, Tsubouchi H, Yatsuhashi H, Kato M, Ki R, Komada Y, Seto C, Goto S. Simeprevir with peginterferon/ribavirin for treatment-naïve hepatitis C genotype 1 patients in Japan: CONCERTO-1, a phase III trial. *J Hepatol* 2014; **61**: 219-227 [PMID: 24727123 DOI: 10.1016/j.jhep.2014.04.004]
- 6 **Lok AS**, Gardiner DF, Hézode C, Lawitz EJ, Bourlière M, Everson GT, Marcellin P, Rodriguez-Torres M, Pol S, Serfaty L, Eley T, Huang SP, Li J, Wind-Rotolo M, Yu F, McPhee F, Grasela DM, Pasquinelli C. Randomized trial of daclatasvir and asunaprevir with or without PegIFN/RBV for hepatitis C virus genotype 1 null responders. *J Hepatol* 2014; **60**: 490-499 [PMID: 24444658 DOI: 10.1016/j.jhep.2013.10.019]
- 7 **Manns M**, Marcellin P, Poordad F, de Araujo ES, Buti M, Horsmans Y, Janczewska E, Villamil F, Scott J, Peeters M, Lenz O, Ouwerkerk-Mahadevan S, De La Rosa G, Kalmeijer R, Sinha R, Beumont-Mauviel M. Simeprevir with pegylated interferon alfa 2a or 2b plus ribavirin in treatment-naïve patients with chronic hepatitis C virus genotype 1 infection (QUEST-2): A randomised, double-blind, placebo-controlled phase 3 trial. *Lancet* 2014; **384**: 414-426 [PMID: 24907224 DOI: 10.1016/S0140-6736(14)60538-9]
- 8 **Manns M**, Pol S, Jacobson IM, Marcellin P, Gordon SC, Peng CY, Chang TT, Everson GT, Heo J, Gerken G, Yoffe B, Townner WJ, Bourlière M, Metivier S, Chu CJ, Sievert W, Bronowicki JP, Thabut D, Lee YJ, Kao JH, McPhee F, Kopit J, Mendez P, Linaberry M, Hughes E, Novello S; HALLMARK-DUAL Study Team. All-oral daclatasvir plus asunaprevir for hepatitis C virus genotype 1b: A multinational, phase 3, multicohort study. *Lancet* 2014; **384**: 1597-1605 [PMID: 25078304 DOI: 10.1016/S0140-6736(14)61059-X]
- 9 **Afdhal N**, Reddy KR, Nelson DR, Lawitz E, Gordon SC, Schiff E, Nahass R, Ghalib R, Gitlin N, Herring R, Lalezari J, Younes ZH, Pockros PJ, Di Bisceglie AM, Arora S, Subramanian GM, Zhu Y, Dvory-Sobol

- H, Yang JC, Pang PS, Symonds WT, McHutchison JG, Muir AJ, Sulkowski M, Kwo P; ION-2 Investigators. Ledipasvir and sofosbuvir for previously treated HCV genotype 1 infection. *N Engl J Med* 2014; **370**: 1483-1493 [PMID: [24725238](#) DOI: [10.1056/NEJMoa1316366](#)]
- 10 **Afdhal N**, Zeuzem S, Kwo P, Chojkier M, Gitlin N, Puoti M, Romero-Gomez M, Zarski JP, Agarwal K, Buggisch P, Foster GR, Bräu N, Buti M, Jacobson IM, Subramanian GM, Ding X, Mo H, Yang JC, Pang PS, Symonds WT, McHutchison JG, Muir AJ, Mangia A, Marcellin P; ION-1 Investigators. Ledipasvir and sofosbuvir for untreated HCV genotype 1 infection. *N Engl J Med* 2014; **370**: 1889-1898 [PMID: [24725239](#) DOI: [10.1056/NEJMoa1402454](#)]
- 11 **Jacobson IM**, McHutchison JG, Dusheiko G, Di Bisceglie AM, Reddy KR, Bzowej NH, Marcellin P, Muir AJ, Ferenci P, Flisiak R, George J, Rizzetto M, Shouval D, Sola R, Terg RA, Yoshida EM, Adda N, Bengtsson L, Sankoh AJ, Kieffer TL, George S, Kauffman RS, Zeuzem S; ADVANCE Study Team. Telaprevir for previously untreated chronic hepatitis C virus infection. *N Engl J Med* 2011; **364**: 2405-2416 [PMID: [21696307](#) DOI: [10.1056/NEJMoa1012912](#)]
- 12 **Zeuzem S**, Foster GR, Wang S, Asatryan A, Gane E, Feld JJ, Asselah T, Bourlière M, Ruane PJ, Wedemeyer H, Pol S, Flisiak R, Poordad F, Chuang WL, Stedman CA, Flamm S, Kwo P, Dore GJ, Sepulveda-Arzola G, Roberts SK, Soto-Malave R, Kaita K, Puoti M, Vierling J, Tam E, Vargas HE, Bruck R, Fuster F, Paik SW, Felizarta F, Kort J, Fu B, Liu R, Ng TI, Pilot-Matias T, Lin CW, Trinh R, Mensa FJ. Glecaprevir-Pibrentasvir for 8 or 12 Weeks in HCV Genotype 1 or 3 Infection. *N Engl J Med* 2018; **378**: 354-369 [PMID: [29365309](#) DOI: [10.1056/NEJMoa1702417](#)]
- 13 **Takeda H**, Takai A, Inuzuka T, Marusawa H. Genetic basis of hepatitis virus-associated hepatocellular carcinoma: Linkage between infection, inflammation, and tumorigenesis. *J Gastroenterol* 2017; **52**: 26-38 [PMID: [27714455](#) DOI: [10.1007/s00535-016-1273-2](#)]
- 14 **Krishnan P**, Tripathi R, Schnell G, Reisch T, Beyer J, Irvin M, Xie W, Larsen L, Cohen D, Podsadecki T, Pilot-Matias T, Collins C. Resistance analysis of baseline and treatment-emergent variants in hepatitis C virus genotype 1 in the AVIATOR study with paritaprevir-ritonavir, ombitasvir, and dasabuvir. *Antimicrob Agents Chemother* 2015; **59**: 5445-5454 [PMID: [26100711](#) DOI: [10.1128/AAC.00998-15](#)]
- 15 **McPhee F**, Friborg J, Levine S, Chen C, Falk P, Yu F, Hernandez D, Lee MS, Chaniewski S, Sheaffer AK, Pasquinelli C. Resistance analysis of the hepatitis C virus NS3 protease inhibitor asunaprevir. *Antimicrob Agents Chemother* 2012; **56**: 3670-3681 [PMID: [22508297](#) DOI: [10.1128/AAC.00308-12](#)]
- 16 **Pilot-Matias T**, Tripathi R, Cohen D, Gaultier I, Dekhtyar T, Lu L, Reisch T, Irvin M, Hopkins T, Pithawalla R, Middleton T, Ng T, McDaniel K, Or YS, Menon R, Kempf D, Molla A, Collins C. In vitro and in vivo antiviral activity and resistance profile of the hepatitis C virus NS3/4A protease inhibitor ABT-450. *Antimicrob Agents Chemother* 2015; **59**: 988-997 [PMID: [25451053](#) DOI: [10.1128/AAC.04227-14](#)]
- 17 **Sarrazin C**, Dvory-Sobol H, Svarovskaia ES, Doehle BP, Pang PS, Chuang SM, Ma J, Ding X, Afdhal NH, Kowdley KV, Gane EJ, Lawitz E, Brainard DM, McHutchison JG, Miller MD, Mo H. Prevalence of Resistance-Associated Substitutions in HCV NS5A, NS5B, or NS3 and Outcomes of Treatment With Ledipasvir and Sofosbuvir. *Gastroenterology* 2016; **151**: 501-512.e1 [PMID: [27296509](#) DOI: [10.1053/j.gastro.2016.06.002](#)]
- 18 **Abdelrahman T**, Hughes J, Main J, McLauchlan J, Thursz M, Thomson E. Next-generation sequencing sheds light on the natural history of hepatitis C infection in patients who fail treatment. *Hepatology* 2015; **61**: 88-97 [PMID: [24797101](#) DOI: [10.1002/hep.27192](#)]
- 19 **Kai Y**, Hikita H, Tatsumi T, Nakabori T, Saito Y, Morishita N, Tanaka S, Nawa T, Oze T, Sakamori R, Yakushiji T, Hiramatsu N, Suemizu H, Takehara T. Emergence of hepatitis C virus NS5A L31V plus Y93H variant upon treatment failure of daclatasvir and asunaprevir is relatively resistant to ledipasvir and NS5B polymerase nucleotide inhibitor GS-558093 in human hepatocyte chimeric mice. *J Gastroenterol* 2015; **50**: 1145-1151 [PMID: [26208695](#) DOI: [10.1007/s00535-015-1108-6](#)]
- 20 **Karino Y**, Toyota J, Ikeda K, Suzuki F, Chayama K, Kawakami Y, Ishikawa H, Watanabe H, Hernandez D, Yu F, McPhee F, Kumada H. Characterization of virologic escape in hepatitis C virus genotype 1b patients treated with the direct-acting antivirals daclatasvir and asunaprevir. *J Hepatol* 2013; **58**: 646-654 [PMID: [23178977](#) DOI: [10.1016/j.jhep.2012.11.012](#)]
- 21 **Sarrazin C**. The importance of resistance to direct antiviral drugs in HCV infection in clinical practice. *J Hepatol* 2016; **64**: 486-504 [PMID: [26409317](#) DOI: [10.1016/j.jhep.2015.09.011](#)]
- 22 **Yoshimi S**, Imamura M, Murakami E, Hiraga N, Tsuge M, Kawakami Y, Aikata H, Abe H, Hayes CN, Sasaki T, Ochi H, Chayama K. Long term persistence of NS5A inhibitor-resistant hepatitis C virus in patients who failed daclatasvir and asunaprevir therapy. *J Med Virol* 2015; **87**: 1913-1920 [PMID: [25954851](#) DOI: [10.1002/jmv.24255](#)]
- 23 **Kosaka K**, Imamura M, Hayes CN, Abe H, Hiraga N, Yoshimi S, Murakami E, Kawaoka T, Tsuge M, Aikata H, Miki D, Ochi H, Matsui H, Kanai A, Inaba T, Chayama K. Emergence of resistant variants detected by ultra-deep sequencing after asunaprevir and daclatasvir combination therapy in patients infected with hepatitis C virus genotype 1. *J Viral Hepat* 2015; **22**: 158-165 [PMID: [24943406](#) DOI: [10.1111/jvh.12271](#)]
- 24 **Mizokami M**, Dvory-Sobol H, Izumi N, Nishiguchi S, Doehle B, Svarovskaia ES, De-Oertel S, Knox S, Brainard DM, Miller MD, Mo H, Sakamoto N, Takehara T, Omata M. Resistance Analyses of Japanese Hepatitis C-Infected Patients Receiving Sofosbuvir or Ledipasvir/Sofosbuvir Containing Regimens in Phase 3 Studies. *J Viral Hepat* 2016; **23**: 780-788 [PMID: [27196675](#) DOI: [10.1111/jvh.12549](#)]
- 25 **Yoshimi S**, Ochi H, Murakami E, Uchida T, Kan H, Akamatsu S, Hayes CN, Abe H, Miki D, Hiraga N, Imamura M, Aikata H, Chayama K. Rapid, Sensitive, and Accurate Evaluation of Drug Resistant Mutant (NS5A-Y93H) Strain Frequency in Genotype 1b HCV by Invader Assay. *PLoS One* 2015; **10**: e0130022 [PMID: [26083687](#) DOI: [10.1371/journal.pone.0130022](#)]
- 26 **Chayama K**, Hayes CN. HCV Drug Resistance Challenges in Japan: The Role of Pre-Existing Variants and Emerging Resistant Strains in Direct Acting Antiviral Therapy. *Viruses* 2015; **7**: 5328-5342 [PMID: [26473914](#) DOI: [10.3390/v7102876](#)]
- 27 **Osawa M**, Imamura M, Teraoka Y, Uchida T, Morio K, Fujino H, Nakahara T, Ono A, Murakami E, Kawaoka T, Miki D, Tsuge M, Hiramatsu A, Aikata H, Hayes CN, Chayama K; Hiroshima Liver Study Group. Real-world efficacy of glecaprevir plus pibrentasvir for chronic hepatitis C patient with previous direct-acting antiviral therapy failures. *J Gastroenterol* 2019; **54**: 291-296 [PMID: [30334096](#) DOI: [10.1007/s00535-018-1520-9](#)]
- 28 **Uemura H**, Uchida Y, Kouyama JI, Naiki K, Tsuji S, Sugawara K, Nakao M, Motoya D, Nakayama N, Imai Y, Tomiya T, Mochida S. NS5A-P32 deletion as a factor involved in virologic failure in patients receiving glecaprevir and pibrentasvir. *J Gastroenterol* 2019; **54**: 459-470 [PMID: [30612205](#) DOI: [10.1007/s00535-018-01543-9](#)]

- 29 Ikegami T, Ueda Y, Akamatsu N, Ishiyama K, Goto R, Soyama A, Kuramitsu K, Honda M, Shinoda M, Yoshizumi T, Okajima H, Kitagawa Y, Inomata Y, Ku Y, Eguchi S, Taketomi A, Ohdan H, Kokudo N, Shimada M, Yanaga K, Furukawa H, Uemoto S, Maehara Y. Asunaprevir and daclatasvir for recurrent hepatitis C after liver transplantation: A Japanese multicenter experience. *Clin Transplant* 2017; 31 [PMID: 28881052 DOI: 10.1111/ctr.13109]
- 30 Ueda Y, Ikegami T, Akamatsu N, Soyama A, Shinoda M, Goto R, Okajima H, Yoshizumi T, Taketomi A, Kitagawa Y, Eguchi S, Kokudo N, Uemoto S, Maehara Y. Treatment with sofosbuvir and ledipasvir without ribavirin for 12 weeks is highly effective for recurrent hepatitis C virus genotype 1b infection after living donor liver transplantation: A Japanese multicenter experience. *J Gastroenterol* 2017; 52: 986-991 [PMID: 28138756 DOI: 10.1007/s00535-017-1310-9]
- 31 Goodwin S, McPherson JD, McCombie WR. Coming of age: Ten years of next-generation sequencing technologies. *Nat Rev Genet* 2016; 17: 333-351 [PMID: 27184599 DOI: 10.1038/nrg.2016.49]
- 32 Archer J, Weber J, Henry K, Winner D, Gibson R, Lee L, Paxinos E, Arts EJ, Robertson DL, Mimms L, Quiñones-Mateu ME. Use of four next-generation sequencing platforms to determine HIV-1 coreceptor tropism. *PLoS One* 2012; 7: e49602 [PMID: 23166726 DOI: 10.1371/journal.pone.0049602]
- 33 Inuzuka T, Ueda Y, Morimura H, Fujii Y, Umeda M, Kou T, Osaki Y, Uemoto S, Chiba T, Marusawa H. Reactivation from occult HBV carrier status is characterized by low genetic heterogeneity with the wild-type or G1896A variant prevalence. *J Hepatol* 2014; 61: 492-501 [PMID: 24798622 DOI: 10.1016/j.jhep.2014.04.033]
- 34 Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner K, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. Real-time DNA sequencing from single polymerase molecules. *Science* 2009; 323: 133-138 [PMID: 19023044 DOI: 10.1126/science.1162986]
- 35 Guo X, Lehner K, O'Connell K, Zhang J, Dave SS, Jinks-Robertson S. SMRT Sequencing for Parallel Analysis of Multiple Targets and Accurate SNP Phasing. *G3 (Bethesda)* 2015; 5: 2801-2808 [PMID: 26497143 DOI: 10.1534/g3.115.023317]
- 36 Roberts RJ, Carneiro MO, Schatz MC. The advantages of SMRT sequencing. *Genome Biol* 2013; 14: 405 [PMID: 23822731 DOI: 10.1186/gb-2013-14-6-405]
- 37 Huang DW, Raley C, Jiang MK, Zheng X, Liang D, Rehman MT, Highbarger HC, Jiao X, Sherman B, Ma L, Chen X, Skelly T, Troyer J, Stephens R, Imamichi T, Pau A, Lempicki RA, Tran B, Nissley D, Lane HC, Dewar RL. Towards Better Precision Medicine: PacBio Single-Molecule Long Reads Resolve the Interpretation of HIV Drug Resistant Mutation Profiles at Explicit Quasispecies (Haplotype) Level. *J Data Mining Genomics Proteomics* 2016; 7: pii: 182 [PMID: 26949565 DOI: 10.4172/2153-0602.1000182]
- 38 Chaisson MJ, Huddleston J, Dennis MY, Sudmant PH, Malig M, Hormozdiari F, Antonacci F, Surti U, Sandstrom R, Boitano M, Landolin JM, Stamatoyannopoulos JA, Hunkapiller MW, Korlach J, Eichler EE. Resolving the complexity of the human genome using single-molecule sequencing. *Nature* 2015; 517: 608-611 [PMID: 25383537 DOI: 10.1038/nature13907]
- 39 Schirmer M, Sloan WT, Quince C. Benchmarking of viral haplotype reconstruction programmes: An overview of the capacities and limitations of currently available programmes. *Brief Bioinform* 2014; 15: 431-442 [PMID: 23257116 DOI: 10.1093/bib/bbs081]
- 40 Seo JS, Rhie A, Kim J, Lee S, Sohn MH, Kim CU, Hastie A, Cao H, Yun JY, Kim J, Kuk J, Park GH, Kim J, Ryu H, Kim J, Roh M, Baek J, Hunkapiller MW, Korlach J, Shin JY, Kim C. De novo assembly and phasing of a Korean human genome. *Nature* 2016; 538: 243-247 [PMID: 27706134 DOI: 10.1038/nature20098]
- 41 Smith CC, Wang Q, Chin CS, Salerno S, Damon LE, Levis MJ, Perl AE, Travers KJ, Wang S, Hunt JP, Zarrinkar PP, Schadt EE, Kasarskis A, Kuriyan J, Shah NP. Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature* 2012; 485: 260-263 [PMID: 22504184 DOI: 10.1038/nature11016]
- 42 Jiao X, Zheng X, Ma L, Kutty G, Gogineni E, Sun Q, Sherman BT, Hu X, Jones K, Raley C, Tran B, Munroe DJ, Stephens R, Liang D, Imamichi T, Kovacs JA, Lempicki RA, Huang DW. A Benchmark Study on Error Assessment and Quality Control of CCS Reads Derived from the PacBio RS. *J Data Mining Genomics Proteomics* 2013; 4: pii: 16008 [PMID: 24179701 DOI: 10.4172/2153-0602.1000136]
- 43 Travers KJ, Chin CS, Rank DR, Eid JS, Turner SW. A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic Acids Res* 2010; 38: e159 [PMID: 20571086 DOI: 10.1093/nar/gkq543]
- 44 Takeda H, Ueda Y, Inuzuka T, Yamashita Y, Osaki Y, Nasu A, Umeda M, Takemura R, Seno H, Sekine A, Marusawa H. Evolution of multi-drug resistant HCV clones from pre-existing resistant-associated variants during direct-acting antiviral therapy determined by third-generation sequencing. *Sci Rep* 2017; 7: 45605 [PMID: 28361915 DOI: 10.1038/srep45605]
- 45 Weirather JL, de Cesare M, Wang Y, Piazza P, Sebastiano V, Wang XJ, Buck D, Au KF. Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. *FL1000Res* 2017; 6: 100 [PMID: 28868132 DOI: 10.12688/fl1000research.10571.2]
- 46 Rang FJ, Kloosterman WP, de Ridder J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol* 2018; 19: 90 [PMID: 30005597 DOI: 10.1186/s13059-018-1462-9]
- 47 Mardis ER. New strategies and emerging technologies for massively parallel sequencing: Applications in medical research. *Genome Med* 2009; 1: 40 [PMID: 19435481 DOI: 10.1186/gm40]
- 48 Nasu A, Marusawa H, Ueda Y, Nishijima N, Takahashi K, Osaki Y, Yamashita Y, Inokuma T, Tamada T, Fujiwara T, Sato F, Shimizu K, Chiba T. Genetic heterogeneity of hepatitis C virus in association with antiviral therapy determined by ultra-deep sequencing. *PLoS One* 2011; 6: e24907 [PMID: 21966381 DOI: 10.1371/journal.pone.0024907]
- 49 Sato M, Maekawa S, Komatsu N, Tatsumi A, Miura M, Muraoka M, Suzuki Y, Amemiya F, Takano S, Fukasawa M, Nakayama Y, Yamaguchi T, Uetake T, Inoue T, Sato T, Sakamoto M, Yamashita A, Moriishi K, Enomoto N. Deep sequencing and phylogenetic analysis of variants resistant to interferon-based protease inhibitor therapy in chronic hepatitis induced by genotype 1b hepatitis C virus. *J Virol* 2015; 89: 6105-6116 [PMID: 25810555 DOI: 10.1128/JVI.03127-14]
- 50 Teraoka Y, Uchida T, Imamura M, Osawa M, Tsuge M, Abe-Chayama H, Hayes CN, Makokha GN,

- Aikata H, Miki D, Ochi H, Ishida Y, Tateno C, Chayama K; Hiroshima Liver Study Group. Prevalence of NS5A resistance associated variants in NS5A inhibitor treatment failures and an effective treatment for NS5A-P32 deleted hepatitis C virus in humanized mice. *Biochem Biophys Res Commun* 2018; **500**: 152-157 [PMID: [29621544](#) DOI: [10.1016/j.bbrc.2018.04.005](#)]
- 51 **Bergfors A**, Leenheer D, Bergqvist A, Ameer A, Lennerstrand J. Analysis of hepatitis C NS5A resistance associated polymorphisms using ultra deep single molecule real time (SMRT) sequencing. *Antiviral Res* 2016; **126**: 81-89 [PMID: [26707078](#) DOI: [10.1016/j.antiviral.2015.12.005](#)]
- 52 **Bull RA**, Eltahla AA, Rodrigo C, Koekkoek SM, Walker M, Pirozyan MR, Betz-Stablein B, Toepfer A, Laird M, Oh S, Heiner C, Maher L, Schinkel J, Lloyd AR, Luciani F. A method for near full-length amplification and sequencing for six hepatitis C virus genotypes. *BMC Genomics* 2016; **17**: 247 [PMID: [26988550](#) DOI: [10.1186/s12864-016-2575-8](#)]
- 53 **Ho CKY**, Raghwan J, Koekkoek S, Liang RH, Van der Meer JTM, Van Der Valk M, De Jong M, Pybus OG, Schinkel J, Molenkamp R. Characterization of Hepatitis C Virus (HCV) Envelope Diversification from Acute to Chronic Infection within a Sexually Transmitted HCV Cluster by Using Single-Molecule, Real-Time Sequencing. *J Virol* 2017; **91**: pii: e02262-16 [PMID: [28077634](#) DOI: [10.1128/JVI.02262-16](#)]



Published By Baishideng Publishing Group Inc
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-2238242
Fax: +1-925-2238243
E-mail: bpgoffice@wjgnet.com
Help Desk: <http://www.f6publishing.com/helpdesk>
<http://www.wjgnet.com>

