Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.jzus.zju.edu.cn; www.springerlink.com
E-mail: jzus@zju.edu.cn

DNA sequences homologous to hepatitis C virus (HCV) in the extrachromosomal circular DNA in peripheral blood mononuclear cells of HCV-negative subjects*#

Reinhard H. DENNIN[†‡1], Jian-er WO[2]

[1]*Formerly Department of Infectious Diseases and Microbiology, University of Lübeck, University Hospital Schleswig-Holstein, 23538 Lübeck, Germany*

[2]*State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, Institute of Infectious Diseases, the First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China*

[†]E-mail: rh.dennin@t-online.de

Received Aug. 30, 2018; Revision accepted Feb. 27, 2019; Crosschecked June 10, 2019

Abstract: Objective: This study aimed to investigate DNA sequences that are substantially homologous to the corresponding RNA sequence sections of the hepatitis C virus (HCV). These DNA sequences are present in the whole DNA extracted from peripheral blood mononuclear cells (PBMCs) of HCV-negative subjects. We presumed that these experimentally proven 5'-noncoding region (5'-NCR) homologous DNA sequences could be contained in the extrachromosomal circular DNA (eccDNA) fraction as part of the whole cellular DNA. Methods: Home-made polymerase chain reaction (PCR) with whole cellular and isolated eccDNA, nucleotide basic local alignment search tool (BLASTn) alignments, and tests for patterns of methylation in selected sequence sections were performed. Results: The PCR tests revealed DNA sequences of up to 320 bp that broadly matched the corresponding sequence sections of known HCV genotypes. In contrast, BLASTn alignment searches of published HCV 5'-NCR sequences with human genome databases revealed only sequence segments of up to 36 bp of the 5'-NCR. The composition of these sequences shows missing base pairs, base pair mismatches as well as complete homology with HCV reference sequences. These short sequence sections are present in numerous copies on both the same and different chromosomes. The selected sequence region within the DNA sequences of the 5'-NCR revealed a broad diversity of individual patterns of methylation. Conclusions: The experimental results confirm our assumption that parts of the HCV 5'-NCR genomic RNA sequences are present at the DNA level in the eccDNA fraction of PBMCs. The tests for methylation patterns therein revealed individual methylomes which could represent an epigenetic feature. The respective sequence section might be subject to genetic regulation.

Key words: Hepatitis C virus (HCV); 5'-Non-coding region (5'-NCR); Human genome; Extrachromosomal DNA; Circular DNA; Pattern of methylation
https://doi.org/10.1631/jzus.B1800453 **CLC number:** R512.6

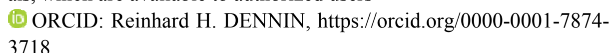
1 Introduction

The genomic information of the infectious hepatitis C virus (HCV) is encoded in a single-stranded, positive-sense RNA, but it is a non-retroviral RNA virus. It shows high sequence similarity with members of the Flaviviridae family and belongs to the genus *Hepacivirus* (Ashfaq et al., 2011). The HCV is primarily known for causing liver diseases such as

‡ Corresponding author

* Project supported by the Exchange Scholarship Programs of the Landesregierung Schleswig-Holstein, Germany, for Jian-er WO

Electronic supplementary materials: The online version of this article (https://doi.org/10.1631/jzus.B1800453) contains supplementary materials, which are available to authorized users



ORCID: Reinhard H. DENNIN, https://orcid.org/0000-0001-7874-3718

© Zhejiang University and Springer-Verlag GmbH Germany, part of Springer Nature 2019

cirrhosis, and is associated with hepatocellular carcinoma decades after infection (Axley et al., 2018). Furthermore, HCV can cause a broad range of extrahepatic manifestations. Due to its lymphotropism, HCV is associated with various immunological extrahepatic disorders including B cell lymphoproliferative disorders (Böhlig et al., 2014; Terrier et al., 2014; Mihăilă, 2016).

Previous results showed that whole DNA extracted from the peripheral blood mononuclear cells (PBMCs) of all HCV-negative individuals tested (more than 150) as well as HCV-positive patients harbors HCV homologous sequences. These sequences span from as short as 82 bp up to 320/341 bp of its 5'-noncoding region (5'-NCR), depending on the primers used for polymerase chain reaction (PCR) (Dennin and Wo, 2003). Checks of these and published sequences of the respective region of the 5'-NCR with the nucleotide basic local alignment search tool (BLASTn) alignments in human genomic DNA (Human Genome Project) never showed such sequences of the length obtained by PCR. Therefore, it appeared possible that these sequences may belong to a particular type of extrachromosomal DNA (ecDNA) co-extracted as part of the whole cellular DNA. We focused on the extrachromosomal circular DNA (eccDNA) which is part of the ecDNA and is present in eukaryotic cells. The size of eccDNA sequences ranges from some hundred base pairs (bp) to more than 100 kb, and their share of the total cellular DNA ranges from about 4% to 20%.

We used DNA extracted from PBMCs of HCV-negative tested subjects. The extracted whole DNA was subjected to a commercial protocol to isolate eccDNA. This study aimed to demonstrate that HCV homologous DNA sequences are contained in the fraction of eccDNA, and to provide experimental evidence of individual patterns of methylation of the HCV homologous DNA sequences in its 5'-NCR.

2 Results

2.1 Computational approaches by alignment search with BLASTn and the 5'-NCR of HCV in human genomic DNA

Depending on the versions available, the human genome library shows HCV sequences of up to 36 bp containing 5'-NCR sections (Fig. 1). These are shorter than those detected by experimental PCR. However, according to representative samples of the HCV genome genotype (GT) 1 and GT6 by BLASTn alignment searches, large portions of the HCV genome appear to be contained in the human genome after summing up the short single-sequence sections; the distribution of sequences is not random in relation to HCV genomic regions. Incidentally, this also applies to the 5'-NCR in genomic libraries of non-human primates (Dennin et al., 2000).

Furthermore, we checked by BLASTn the sequence sections of the 5'-NCR from nucleotides (nt) 10 to nt 80 and from nt 150 to nt 300: the BLASTn alignments of the 5'-NCR sequence revealed several hits from nt 41 to nt 64, with the highest homology of 22/24 identities, and showed no significant similarity.

In addition, we checked the HCV GT1b complete genome without query subranges along the current human genome database. The alignment displayed the most extended HCV sequence nt 9390 to nt 9457 with 56/68 identities on *Homo sapiens* chromosome 17; this sequence section is part of the 3'-NCR that follows the HCV RNA-dependent RNA polymerase (non-structural protein 5B (NS5B)) sequence.

As suspected already from previous experiments (Dennin and Wo, 2003) using whole DNA extracted from PBMCs of HCV-negative subjects, we presumed from the divergence of computational alignments and the HCV homologous DNA sequences detectable by experimental PCR that these sequences must be present in the ecDNA fraction. We focused on the eccDNA contained therein.

2.2 Experimental approaches

2.2.1 Isolation of eccDNA from whole cellular DNA

To separate the eccDNA fraction from the chromosomal DNA (chrDNA) of whole DNA extracted from PBMCs, we used the Plasmid-Safe™ ATP-dependent DNase (PSAD) protocol (see Materials and methods (Data S1), Section 3). According to the information presented by the provider, the DNA fractions left over after PSAD digestion may contain circular nicked or closed double-stranded DNA (dsDNA) or supercoiled DNA, i.e., the PSAD-resistant eccDNA, whereas linear DNA, predominantly chrDNA, is digested. The PSAD-resistant eccDNA was analyzed for

Alignments of HCV sub-GT 1b, with the accession number AJ238799.1, 5'-NCR, query subrange nt 80–nt 150 versus:

i) *Homo sapiens* chromosome 2, GRCh38.p12 Primary Assembly

Features: 56 875 bp at 5' side: potassium voltage-gated channel subfamily F member 1 162 514 bp at 3' side: uncharacterized protein C2orf50 isoform X1

```
Query  111       CCTCCAGGAcccccccTCCCGGGAGAGCCATAGTGG  146
                 ||| |||||||||   ||| ||||||||||||| ||
Sbjct  10970786  CCTGCAGGACCCCTAGTCCAGGGAGAGCCATAGAGG  10970821
```

ii) *Homo sapiens* chromosome 13, GRCh38.p12 Primary Assembly

Features: uncharacterized protein LOC107983958

```
Query  103       TCGTGCAGCCTCCAGGAcccccccTCCCGGGAGA  136
                 ||| | |||||||||||||||  |||| ||| ||
Sbjct  111626361 TCGGGGAGCCTCCAGGACCCCATCTCCTGGGTGA  111626394
```

Fig. 1 Sequences from the 5'-NCR with query subrange from nt 80 to nt 150 of the HCV sub-genotype (GT) 1b for BLASTn alignment searches

Selected for slightly similar sequences (BLASTn). The selection of this sequence section was chosen to cover the sequences of the external downstream primers PTNC-E1 and PTNC-I1 for reasons outlined in the Materials and methods (Data S1), Section 7. The BLASTn HCV queries generated 12 hits on the chrDNA deposited in the human genome library; two selected ones are shown here: i) The hit with the longest sequence with 30/36 identities was chosen from a total of 12 matches, and hits were shown on both the same and different chromosomes including X; ii) The BLASTn alignment with the same query for the HCV 5'-NCR, nt 80 to nt 150, revealed one match with 28/34 identities on chromosome 13, and the shortest homologous sequence segment is located on chromosome 7 with 18/18 identities. For details, we used the HCV sequence databases (Lohmann et al., 1999; Kuiken et al., 2005). 5'-NCR: 5'-noncoding region; HCV: hepatitis C virus; BLASTn: nucleotide basic local alignment search tool; chrDNA: chromosomal DNA; nt: nucleotides

HCV homologous sequences by regular nested and semi-nested PCR. Fig. 2a illustrates how the PSAD protocol (right part) works on whole DNA (left part) extracted from PBMCs.

The following test was performed to determine the presence of HCV DNA sequences in both the fractions of (35±15) kb in undigested whole DNA extracts from PBMCs (Fig. 2a, left part) and after subjugation of the whole DNA with PSAD digestion and gel electrophoresis (Fig. 2a, right part). The fractions of about 35 kb isolated from total PBMC DNA prior to PSAD digestion contained HCV homologous DNA sequences (Fig. 2b, left part). The PSAD-resistant eccDNA fractions (Fig. 2b, right part) contained HCV homologous DNA sequences as well.

2.2.2 PCR tests

Two rounds of PCR with primers: AE1/AE2 and PTNC E1/AK3. The generated amplicons were cloned and sequenced (see Materials and methods, Sections 8 and 9). For the first check, ahead of the electrophoresis in the 0.6% agarose gel, both the whole DNA extracted from PBMCs and the DNA left over after PSAD digestion were checked by nested PCR with HCV 5'-NCR homologous primers. For the second check, additional steps were taken for the purification of the PSAD-resistant DNA from the gel ahead of PCR to generate amplicons shown in Fig. 2, For cloning and sequencing, the PSAD-digested PBMC DNA, i.e., the PSAD-resistant eccDNA prepared from the same subjects as shown in Fig. 2 (right part): after electrophoresis on a 0.6% agarose gel, the prominent DNA fractions left over after PSAD digestion present in the range of about 35 kb were excised; DNA was extracted from the gel blocks using a QIAEX II Gel Extraction kit (see Materials and methods, Section 4). After this additional purification step, the DNA extracted from the gel revealed amplicons after PCR of the expected size (186 bp) according to the location of the primers used along the 5'-NCR sequence. The specificity of the amplicons was confirmed by cloning and sequencing (see Materials and methods, Sections 8 and 9). If not otherwise stated, 100 ng of DNAs were used throughout for the PCR. Therefore, the DNA concentrations used per test were not standardized to uniform templates. The resulting amplicons had the expected size of 186 bp.

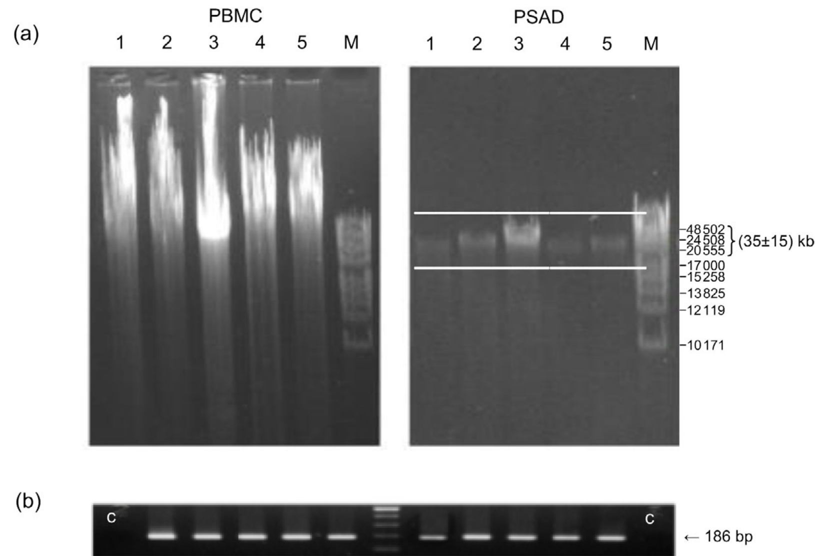


Fig. 2 Results after gel electrophoresis of the whole DNA extracted from PBMCs and the whole leftover DNA fraction resulting from the PSAD digestion of the PBMC DNA (a) and results of PCR with primers of the 5'-NCR/HCV applied to total PBMC DNA and the leftover DNA (eccDNA), respectively (b)

(a) Analysis of the whole DNA extracted from PBMCs from five subjects in Lanes 1 to 5 (for details, see Materials and methods, Data S1). Left part: undigested whole PBMC DNA, 500 ng/lane; Right part: the PSAD-resistant (leftover) eccDNA shown per lane results from 500 ng of PSAD-digested PBMC DNA. The range of (35±15) kb is marked by two white lines. M: 10171 to 48502 bp markers. (b) The first check: ahead of the electrophoresis in the 0.6% agarose gel, both the whole DNA extracted from PBMCs and the DNA left over after PSAD digestion were checked by nested PCR with HCV 5'-NCR homologous primers. An arrow targets the expected size: 186 bp of the amplicons. “C” in (b) indicates blank controls: consistently negative. PBMC: peripheral blood mononuclear cell; PSAD: Plasmid-Safe™ ATP-dependent DNase; HCV: hepatitis C virus; 5'-NCR: 5'-noncoding region

It is necessary to mention that depending on the HCV 5'-NCR gene-specific primers used for PCR with the whole DNA extracted from PBMCs, sometimes amplicons of longer than the expected sizes (LEXS) are generated from the whole DNA of PBMCs of both human subjects and non-human primates (Dennin et al., 2000). These amplicons of LEXS show certain kind of repeats of regular sequences, partly overlapping; upon pre-PCR digestion with restriction enzymes (following the same protocol as used below, Fig. 3), most of them are digested down to amplicons of the “expected” size (results not shown here). No HCV-DNA sequences downstream of the 5'-NCR were detected by PCR.

The relative proportion of PSAD-resistant eccDNA as part of the whole PBMC DNA after gel electrophoresis ranged from around 4% to 21% (details in Materials and methods, Section 11). This range of PSAD-resistant eccDNA as part of whole PBMC DNA is the same as that published by Schmidt et al. (2009).

We checked the PSAD-digested fraction (about 35 kb) for human circular mitochondrial DNA (mtDNA, about 16.5 kb) by PCR with mtDNA gene-specific primers. This procedure was used to verify the reliability of the PSAD protocol to retain a naturally existing circular dsDNA in the PSAD-resistant eccDNA. Although mtDNA is a smaller episomal molecule than the more prominent PSAD-resistant eccDNA, it should be partially contained within the (35±15) kb fraction of both the undigested and PSAD-digested DNAs from PBMCs. As expected, mtDNA was present in the respective samples (results not shown here).

2.2.3 Testing for methylation of the sequence section selected by the primers used

We previously showed that HCV-DNA homologous sequences found in the whole DNA extracted from PBMCs exhibit individual patterns of methylation (Dennin and Wo, 2003). Therefore, we investigated

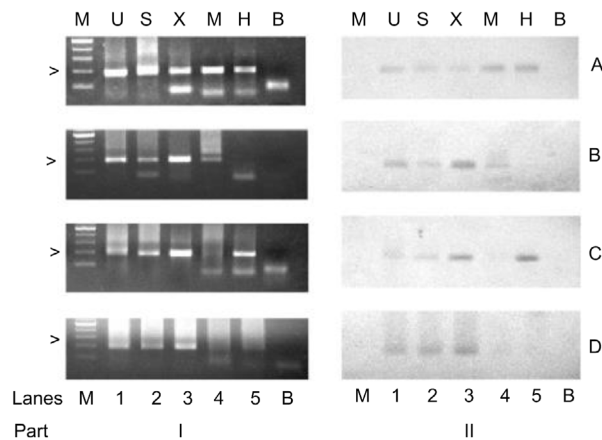


Fig. 3 Examination of the 5'-NCR of HCV on methylation patterns with restriction enzymes

Four healthy, HCV-negative subjects A to D: after exposing the whole DNA extracted from their PBMCs to the PSAD protocol, the remaining eccDNA fraction (about 35 kb) was subjected to pre-PCR digestion with four REs. Lane 1/U: no pre-PCR treatment (i.e., undigested eccDNA); Lane 2/S: digestion with *Sma*I; Lane 3/X: digestion with *Xma*II; Lane 4/M: digestion with *Msp*I; Lane 5/H: digestion with *Hpa*II; Lane B: blank control, a complete reaction mix, but without a DNA matrix; M: molecular weight marker, 100 bp ladder. The RE-digested products were used for nested PCR with the primers AE1/AE2 and PTNC E1/AK3. Lanes 1 to 5 after gel electrophoresis (Part I) and followed by hybridization with the PTNC probe (Part II). The arrowheads ">" mark the amplicons of the expected size resulting from target sequences not digested by the RE used. HCV: hepatitis C virus; PBMC: peripheral blood mononuclear cell; PSAD: Plasmid-Safe™ ATP-dependent DNase; eccDNA: extrachromosomal circular DNA; RE: restriction enzyme

eccDNA isolated from the PSAD-digested fraction (about 35 kb) of PBMCs (Fig. 2a, right part) for the presence of methylated HCV-DNA sequences too.

We applied the same pre-PCR digestion protocol (Dennin and Wo, 2003) with four methylation-sensitive restriction enzymes (MSREs) *Sma*I, *Xma*CI, *Msp*I, and *Hpa*II. These MSREs are inhibited when methylation is present at specific cytosine residues within their targeting sites. Their cutting sequences are included in the selected sequence portion of the 5'-NCR of the HCV. This pre-PCR digestion assay can reveal individual differences in methylation patterns. Such an epigenetic phenomenon has been named a methylome (Lister et al., 2009), here characterizing certain sequence sections of HCV homologous DNA in PBMCs. It can also exclude possible contamination from previous amplicons that are unmethylated and therefore digested by the MSREs.

Additional hints concern the different sizes of amplicons generated by PCR after the pre-PCR RE digest of the eccDNA. Depending on the subject, significantly shorter and longer (partly smears) amplicons of "not the expected sizes" (NEXPs) are present in addition to the amplicons of the expected size. The presence of NEXP amplicons indicates that after restriction enzyme digestion, portions of sequences of the same or different eccDNA molecules are present. These residual DNA sequences must harbor various arrangements of particular complementary target sequences that are suitable for amplification by PCR with the primers used. However, the resulting NEXP amplicons contain no sequences complementary to the PTNC probe: no signals for hybridization, but positive signals for amplicons of the expected size. The results show that the HCV-DNA sequence stretches contained in the eccDNA (about 35 kb) subjected to pre-PCR digestion with four restriction enzymes exhibit patterns of methylation characteristic of each of the four individuals tested. Whether the respective methylation patterns belong to the same species of eccDNA molecules requires further investigation. Checks by BLASTn of sequences from the amplicons gained from the eccDNA fractions reveal: the cutting sequences of the restriction enzymes used are present at least once with their genuine sequences in the amplicons obtained with the primer sets used. The BLASTn alignments allow excluding general changes by substitutions or gaps within. Therefore, the presence of amplicons of the expected size after prePCR digest with the restriction enzymes used rather identifies these amplicon sequences in the eccDNA as being resistant to the four restriction enzymes due to methylation of their cutting sequences. The PTNC probe used for hybridization contains the respective sequences for *Msp*I and *Hpa*II.

2.2.4 Test for human endogenous retrovirus type K (HERV-K) sequences in the eccDNA fraction

The PSAD-resistant eccDNA fractions of three subjects were tested for sequences of the HERV-K by PCR with primers for their long terminal repeats and the polymerase. Both tests showed positive results including amplicons of the expected size as well as

shorter ones with individual variations (results not shown here). Correlated results have been demonstrated by Schmidt et al. (2009).

3 Discussion

3.1 HCV genomic sequences contained in the human genome

The RNA genome of HCV does not encode for an enzyme acting as reverse transcriptase. However, the amplicons generated by regular PCR with HCV gene-specific primers (without a reverse transcriptase step) using the whole cellular DNA extracted from PBMCs of HCV-negative subjects revealed that these amplicons contain DNA sequences broadly homologous to the corresponding genomic RNA of the HCV 5'-NCR. Such sequences are present in the whole DNA extracted from PBMCs of almost all healthy and HCV-negative subjects tested from Europe, Asia, and indigenous people from Argentina, as well as non-human primates (Dennin et al., 2000).

These PCR-generated sequences are longer than those detectable by computational alignment searches using BLASTn and human chrDNA datasets/libraries (Human Genome Project), the longest ones being up to 36 bp. The alignments of the selected 5'-NCR sequence stretches with query subranges from nt 80 to nt 150 showed multiple hits on different chromosomes. Two hits with the most extended sequence area and the highest identities "30/36" and "28/34" are shown in Fig. 1. Note that the first match with the same "30/36" identity was present in previous BLASTn alignments on chromosome 2, sometimes with different loci; the others, however, changed with the timing of alignments.

Upon repetition, the alignments revealed a different number of hits at the same and different positions on different chromosomes. These differences may be due to newer versions of the chrDNAs available in the Human Genome databanks, but they may also indicate the variation of the human genome. Repeated sequences of similar sizes are considered a category of "Human Genome Structural Variation." The results described here may reflect such a type of structural variation. Sequences of the detected lengths represent uniqueness, i.e., not "…by chance." As a supplement, one BLASTn alignment without a query subrange revealed several hits, and the one with the longest sequence showed 56/68 identities. The embeddedness of such a long sequence predominantly homologous to HCV in a human gene may raise issues of evolutionary categories.

As far as tested so far, the BLASTn alignments with the PCR-generated HCV homologous sequences show similarities with known HCV genotype/subtype. However, the differences rule out contamination by currently circulating HCV genotypes. Additional studies on occult hepatitis C of the type "no HCV RNA detectable in serum/plasma, anti-HCV test negative, but HCV RNA positive in PBMCs" should make clear-cut distinctions against possible intermingling derived from HCV DNA sequences present in the eccDNA (Halfon et al., 2008; Austria and Wu, 2018). The HCV-encoding sequences isolated from the 5'-NCR of HCV-infected patients show "almost no diversity" (Wang et al., 2010). The predominant homologous HCV sequences in different parts of several chromosomes detected by BLASTn suggest a high resistance to evolutionary selection.

3.2 Extrachromosomal DNA: eccDNA and its molecular features

The next step was to use commercial protocols to isolate the eccDNA fraction from whole DNA extracted from non-sorted PBMCs, resulting in the PSAD-resistant eccDNA (also termed small polydispersed circular DNA (spcDNA)).

The eccDNA molecules as part of the ecDNA fraction are well-known components of eukaryotic cells. On average there can be from about 200 to 1000 molecules per cell that cover a great variety of polydispersed eccDNA. They range in size from 100 bp up to >35 kb, apart from mtDNA (Yamagishi, 1986; Gaubatz, 1990; Barreto et al., 2014; Dillon et al., 2015). mtDNA is not considered here.

The PBMC fraction comprises multitudes of cells with various tasks including pluripotent stem cells and native as well as activated B- and T-lymphocytes, which may result in the induction of different sets of eccDNA molecules in the respective cells. Therefore, the present PSAD-resistant eccDNA fraction might contain a significant number of diverse eccDNA species with different patterns of methylation, for example in response to external or metabolic stress. The eccDNA species harboring HCV homologous

sequences may represent a small fraction of the PSAD-resistant eccDNA fraction.

3.3 Facts and suppositions about the genesis of eccDNA

Reports from sequence analyses of ec/eccDNAs of various sizes demonstrate that these types of molecules might have been generated via certain kinds of chrDNA rearrangements (Cohen et al., 1997). Because they contain sequences similar or even identical to chromosomal sequences, it is reasonable to accept them as being of chromosomal origin. In effect, the HCV DNA sections of up to 320 bp containing its 5'-NCR shown here as a part of certain eccDNAs could have been assembled by incremental acquisition from these numerous short HCV homologous sequence stretches present in different regions of the same or different chromosomes. The results may point to a situation in which, at the genomic level, a mechanism operates to pick up short sequences from the "chromosomal based genomic reservoir" and piece them together like the model of 'genes in pieces' (Smithers et al., 2019) into longer sequence sections as part of the HCV 5'-NCR. Whether this is an ongoing process has not yet been determined.

The mechanism(s) underlying this phenomenon to generate eccDNAs may be through the activity of mobile genetic elements (MGEs) (Smit, 1999; Wessler, 2006) or transposable elements (TEs) (Seberg and Petersen, 2009; Hua-Van et al., 2011). A common feature of MGEs and TEs is that they generate compositions of new and longer sequences made up from shorter sequence sections contained in the genuine human target chrDNA sequences (Cordaux and Batzer, 2009).

Kanduc (2011) compared HCV polyprotein and human proteins at the polypeptide level and found a high degree of peptide sharing. Sequences coding for these proteins must be contained at the nucleotide level of both the HCV and the human genome.

Whether circular micro eccDNAs belong to some kinds of early precursor molecular forms that may develop into larger ones as described here, although with different potentials, remains an open issue (Cohen et al., 2010; Shibata et al., 2012).

3.4 Patterns of methylation and epigenetics

The results of the pre-PCR restriction enzyme digestion protocol hint at an epigenetic phenomenon being imminent with the HCV 5'-NCR DNA contained in these eccDNAs. The results reveal unique methylation patterns for each of the subjects analyzed and hint at an epigenetic modification, which is possibly subject to metabolic impacts; therefore, they may represent individual DNA-based methylomes. In general, DNA methylation of cytosines in different contexts is known for silencing (highly methylated) at the transcriptional level (Lister et al., 2009).

The PBMC population used for DNA extraction contained cells with different tasks. In this context, the different individual concentration of the PSAD leftover DNA (about 35 kb) and the diversity of the individual methylation patterns at the HCV homologous sequence stretches contained in eccDNA may be an indication of different metabolic activities or processes involved in differentiation. Due to the known impacts of methylated cytosines on the regulation of gene activities, this should be viewed in the context of possible biological functions (Phillips, 2008). We suppose that there is a continuous production of eccDNA of different sizes, numbers, and quality subjected to extracellular influences, which may induce cryptic programmed pressures by intracellular signaling cascades, such as '…physiological or pathological stress…' (Komosa et al., 2015).

The presence of methylations within these HCV homologous sequences supports the notion that at least these sections of the 5'-NCR at the DNA level have gained a quality/status in the cell that coerces the cellular machinery to bring them under control by a regulatory mechanism; this relates to possible promoter traits contained therein (see below). Subjects followed for some months or several years showed the same methylomes as well as variations in the pattern of methylation in the same sequence section during the observation period (Dennin and Chen, 1997). The fact that large portions of the HCV 5'-NCR of the eccDNA could be detected in almost all human subjects tested suggests that it has proved to be a stable sequence during evolution. In particular, in the context of the individual pattern of methylation (Fig. 3), the sequences with high homology to the HCV 5'-NCR must have proved resistant against purifying selection in humans (Ward and Kellis, 2012) or even have turned out to be beneficial. The individual patterns of methylation of cytosines within the nt 80 to nt 170 range observed in the eccDNA of the

HCV-negative subjects studied thus far may point to regulative effects in transcription control.

This situation has to be considered in the context of the known presence of methylated cytosines in promoter regions. Dumas et al. (2003) reported that a part of the HCV 5'-NCR at the DNA level possesses a promoter activity able to drive the expression of genes inserted downstream; it starts at nt 67, which is close to the sequence region from nt 80 to nt 170 of the 5'-NCR. The presence of a promoter domain as part of the 5'-NCR at the DNA level with an entirely different function from internal ribosomal entry site (IRES) at the RNA level suggests diverse evolutionary pathways.

4 Conclusions

As HCV homologous sequences detected by regular PCR in whole cellular DNA could not be aligned along their full length with corresponding sequences in human genomic chrDNA databases/ libraries, we designated them as belonging to ecDNA.

The findings presented here support the assignment of the HCV homologous DNA sequences to within the PSAD-resistant eccDNA, i.e., they are an integral part of this eccDNA. Whether these sequences are present temporarily or persistently requires further investigation.

It is disconcerting that the eccDNA fraction has not yet been considered to be part of the total DNA used for whole genome sequencing (WGS) regarding the Human Genome Project (Dennin, 2018).

All information provided for the BLASTn alignments gives no hint of ecDNA when preparing human whole cellular DNA for fragmentation, cloning, etc. for the Human Genome Project (Waterston et al., 2002; Hood and Rowen, 2013). These circumstances indicate that before fragmentation and cloning, no protocol has been included to separate chromosomes from partly large-size eccDNA. Depending on the protocols used for fragmentation, sections of the eccDNA may also have been cloned and incorporated into the alignments for chromosomes. To the best of our knowledge, no reports exclude any possible impact of this nature. This objection is independent of the DNA source for sequencing, e.g., whether DNA is extracted from PBMCs or isolated from human plasma (Zhu et al., 2017).

Contributors

Conceived and designed the experiments: Reinhard H. DENNIN, also partly Jian-er WO; performed the experiments: Jian-er WO, both when present at the Department of Infectious Diseases and Microbiology, University of Lübeck, Germany and at the State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, Institute of Infectious Diseases, the First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China; wrote the paper: Reinhard H. DENNIN; analyzed the data: Reinhard H. DENNIN and Jian-er WO.

Acknowledgments

We would like to thank Prof. Dr. Werner SOLBACH, formerly of the Institute of Medical Microbiology and Hygiene, University of Lübeck, Germany, for research funds for Prof. Jian-er WO.

Parts of this manuscript have been presented at the European Molecular Biology Organization-European Molecular Biology Laboratory (EMBO-EMBL) Symposium—the Mobile Genome, Heidelberg, Germany, September 16–19, 2015.

Compliance with ethics guidelines

Reinhard H. DENNIN and Jian-er WO declare that they have no conflict of interest.

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008 (5). Informed consent was obtained from all subjects for being included in the study.

References

Ashfaq UA, Javed T, Rehman S, et al., 2011. An overview of HCV molecular biology, replication and immune responses. *Virol J*, 8:161. https://doi.org/10.1186/1743-422X-8-161

Austria A, Wu GY, 2018. Occult hepatitis C virus infection: a review. *J Clin Transl Hepatol*, 6(2):155-160. https://doi.org/10.14218/JCTH.2017.00053

Axley P, Ahmed Z, Ravi S, et al., 2018. Hepatitis C virus and hepatocellular carcinoma: a narrative review. *J Clin Transl Hepatol*, 6(1):79-84. https://doi.org/10.14218/JCTH.2017.00067

Barreto SC, Uppalapati M, Ray A, 2014. Small circular DNAs in human pathology. *Malays J Med Sci*, 21(3):4-18.

Böhlig A, Puchner KP, Berg T, 2014. Extrahepatic manifestations. *In*: Mauss S, Berg T, Rockstroh J, et al. (Eds.), Short Guide to Hepatitis C. Flying Publisher, Wuppertal, Germany, p.139-147.

Cohen S, Regev A, Lavi S, 1997. Small polydispersed circular DNA (spcDNA) in human cells: association with genomic instability. *Oncogene*, 14(8):977-985. https://doi.org/10.1038/sj.onc.1200917

Cohen S, Agmon N, Sobol O, et al., 2010. Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in

human cells. *Mob DNA*, 1(1):11.
https://doi.org/10.1186/1759-8753-1-11

Cordaux R, Batzer MA, 2009. The impact of retrotransposons on human genome evolution. *Nat Rev Genet*, 10(10): 691-703.
https://doi.org/10.1038/nrg2640

Dennin RH, 2018. Overlooked: extrachromosomal DNA and their possible impact on whole genome sequencing. *Malays J Med Sci*, 25(2):20-26.
https://doi.org/10.21315/mjms2018.25.2.3

Dennin RH, Chen Z, 1997. Hepatitis C virus (HCV) specific sequences are demonstrable in the DNA fraction of peripheral blood mononuclear cells from healthy, anti-HCV antibody-negative individuals and cell lines of human origin. *Eur J Clin Chem Clin Biochem*, 35(12):899-905.
https://doi.org/10.1515/cclm.1997.35.12.899

Dennin RH, Wo J, 2003. Human DNA contains sequences homologous to the 5'-non-coding region of hepatitis C virus: characterization with restriction endonucleases reveals individual varieties. *Chin Med J (Engl)*, 116(7): 1092-1098.

Dennin RH, Wo J, Chen Z, et al., 2000. DNA of nonhuman primates harbors hepatitis C-virus-specific sequences of its 5'-non-coding region (5'-NCR). *Z Gastroenterol*, 38(12): 925-931.
https://doi.org/10.1055/s-2000-10021

Dillon LW, Kumar P, Shibata Y, et al., 2015. Production of extrachromosomal microDNAs is linked to mismatch repair pathways and transcriptional activity. *Cell Rep*, 11(11):1749-1759.
https://doi.org/10.1016/j.celrep.2015.05.020

Dumas E, Staedel C, Colombat M, et al., 2003. A promoter activity is present in the DNA sequence corresponding to the hepatitis C virus 5' UTR. *Nucleic Acids Res*, 31(4): 1275-1281.
https://doi.org/10.1093/nar/gkg199

Gaubatz JW, 1990. Extrachromosomal circular DNAs and genomic sequence plasticity in eukaryotic cells. *Mutat Res*, 237(5-6):271-292.
https://doi.org/10.1016/0921-8734(90)90009-G

Halfon P, Bourlière M, Ouzan D, et al., 2008. Occult hepatitis C virus infection revisited with ultrasensitive real-time PCR assay. *J Clin Microbiol*, 46(6):2106-2108.
https://doi.org/10.1128/JCM.00345-08

Hood L, Rowen L, 2013. The Human Genome Project: big science transforms biology and medicine. *Genome Med*, 5(9):79.
https://doi.org/10.1186/gm483

Hua-Van A, le Rouzic A, Boutin TS, et al., 2011. The struggle for life of the genome's selfish architects. *Biol Direct*, 6:19.
https://doi.org/10.1186/1745-6150-6-19

Kanduc D, 2011. HCV: written in our DNA. *Self/Nonself*, 2(2):108-113.
https://doi.org/10.4161/self.2.2.15795

Komosa M, Root H, Meyn MS, 2015. Visualization and quantitative analysis of extrachromosomal telomere-repeat DNA in individual human cells by Halo-FISH. *Nucleic Acids Res*, 43(4):2152-2163.
https://doi.org/10.1093/nar/gkv091

Kuiken C, Yusim K, Boykin L, et al., 2005. The Los Alamos hepatitis C sequence database. *Bioinformatics*, 21(3):379-384.
https://doi.org/10.1093/bioinformatics/bth485

Lister R, Pelizzola M, Dowen RH, et al., 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462(7271):315-322.
https://doi.org/10.1038/nature08514

Lohmann V, Körner F, Koch JO, et al., 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science*, 285(5424):110-113.
https://doi.org/10.1126/science.285.5424.110

Mihăilă RG, 2016. Hepatitis C virus-associated B cell non-Hodgkin's lymphoma. *World J Gastroenterol*, 22(27): 6214-6223.
https://doi.org/10.3748/wjg.v22.i27.6214

Phillips T, 2008. The role of methylation in gene expression. *Nat Educ*, 1(1):116.

Schmidt H, Taubert H, Lange H, et al., 2009. Small polydispersed circular DNA contains strains of mobile genetic elements and occurs more frequently in permanent cell lines of malignant tumors than in normal lymphocytes. *Oncol Rep*, 22(2):393-400.
https://doi.org/10.3892/or_00000450

Seberg O, Petersen G, 2009. A unified classification system for eukaryotic transposable elements should reflect their phylogeny. *Nat Rev Genet*, 10(4):276.
https://doi.org/10.1038/nrg2165-c3

Shibata Y, Kumar P, Layer R, et al., 2012. Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. *Science*, 336(6077):82-86.
https://doi.org/10.1126/science.1213307

Smit AF, 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev*, 9(6):657-663.

Smithers B, Oates M, Gough J, 2019. 'Why genes in pieces?'—revisited. *Nucleic Acids Res*, 47(10):4970-4973.
https://doi.org/10.1093/nar/gkz284

Terrier B, Chaara W, Dufat L, et al., 2014. Serum biomarker signature identifies patients with B-cell non-Hodgkin lymphoma associated with cryoglobulinemia vasculitis in chronic HCV infection. *Autoimmun Rev*, 13(3):319-326.
https://doi.org/10.1016/j.autrev.2013.11.001

Wang GP, Sherrill-Mix SA, Chang KM, et al., 2010. Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. *J Virol*, 84(12):6218-6228.
https://doi.org/10.1128/JVI.02271-09

Ward LD, Kellis M, 2012. Evidence of abundant purifying selection in humans for recently acquired regulatory functions. *Science*, 337(6102):1675-1678.

https://doi.org/10.1126/science.1225057
Waterston RH, Lander ES, Sulston JE, 2002. On the sequencing of the human genome. *Proc Natl Acad Sci USA*, 99(6):3712-3716.
https://doi.org/10.1073/pnas.042692499
Wessler SR, 2006. Transposable elements and the evolution of eukaryotic genomes. *Proc Natl Acad Sci USA*, 103(47): 17600-17601.
https://doi.org/10.1073/pnas.0607612103
Yamagishi H, 1986. Role of mammalian circular DNA in cellular differentiation. *BioEssays*, 4(5):218-221.
https://doi.org/10.1002/bies.950040508
Zhu J, Zhang F, Du MJ, et al., 2017. Molecular characterization of cell-free eccDNAs in human plasma. *Sci Rep*, 7(1):10968.
https://doi.org/10.1038/s41598-017-11368-w

List of electronic supplementary materials

Data S1 Materials and methods, supporting online information for protocols

中文概要

题　目：**HCV 阴性者外周血单核细胞的染色体外环状 DNA 中与 HCV 同源的 DNA 序列**
目　的：研究与丙型肝炎病毒（HCV）基因 RNA 序列同源的人染色体外环状 DNA（eccDNA）序列。
创新点：首次从 HCV 阴性者 eccDNA 中检测到 HCV 5'-非编码区（5'-NCR）基因组 RNA 序列，验证了我们的假设：HCV 同源 DNA 序列存在于人的外周血单核细胞的 eccDNA 组分中。
方　法：用分离的 eccDNA 进行 HCV 特异的聚合酶链反应（PCR），采用核苷酸序列同源性搜索分析软件（BLASTn）对测序结果进行比对分析，并检测其甲基化模式。
结　论：实验结果证实了我们的假设：即部分 HCV 5'-NCR 基因组 RNA 序列存在于外周血单核细胞的 eccDNA 组分。同时，甲基化分析结果显示了个体间的甲基化模式所代表的受遗传调控的表观遗传特征。
关键词：丙型肝炎病毒（HCV）；5'-非编码区；人类基因组；染色体外 DNA；环状 DNA；甲基化模式