

# A Microarray Based Approach for the Identification of Common Foodborne Viruses

Mobolanle Ayodeji, Michael Kulka, Scott A. Jackson, Isha Patel, Mark Mammel, Thomas A. Cebula<sup>§</sup>, and Biswendu B. Goswami\*

Division of Molecular Biology, Office of Applied Research and Safety Assessment (OARSA), Food And Drug Administration, 8301 Muirkirk Road, Laurel, Maryland, 20708, USA

**Abstract:** An oligonucleotide array (microarray) incorporating 13,000 elements representing selected strains of hepatitis A virus (HAV), human coxsackieviruses A and B (CVA and CVB), genogroups I and II of Norovirus (NV), and human rotavirus (RV) gene segments 3,4,10, and 11 was designed based on the principle of tiling. Each oligonucleotide was 29 bases long, starting at every 5<sup>th</sup> base of every sequence, resulting in an overlap of 24 bases in two consecutive oligonucleotides. The applicability of the array for virus identification was examined using PCR amplified products from multiple HAV and CV strains. PCR products labeled with biotin were hybridized to the array, and the biotin was detected using a brief reaction with Cy3-labeled streptavidin, the array subjected to laser scanning, and the hybridization data plotted as fluorescence intensity against each oligonucleotide in the array. The combined signal intensities of all probes representing a particular strain of virus were calculated and plotted against all virus strains identified on a linear representation of the array. The profile of the total signal intensity identified the strain that is most likely represented in the amplified cDNA target. The results obtained with HAV and CV indicated that the hybridization profile thus generated can be used to identify closely related viral strains. This represents a significant improvement over current methods for virus identification using PCR amplification and amplicon sequencing.

## INTRODUCTION

Polymerase chain reaction (PCR) coupled to reverse transcription (RT) represents the most significant improvement in the area of RNA virus detection over classical cell culture based methods. In the classical culture based method, the principal mode of virus identification uses growth of the virus in permissive cells and observation of the morphological changes brought about by virus replication in the host cell [1]. Although it is possible to differentiate between cytopathic and non-cytopathic hepatitis A virus (HAV) strains due to a difference in the morphology of infected cells [2], in practice such morphological identification is of limited value because the morphological effects are cell-line specific, and many viruses in the same genus (e.g. *Enterovirus*) produce rapid and similar cytopathic changes in many of the cell-lines normally used for virus detection. Moreover, using multiple cell-lines for virus detection is also labor intensive and time consuming, and further confirmation and identification often requires the use of additional techniques such as serotyping [1].

Molecular methods based on viral RNA amplification by RT-PCR have evolved as rapid alternatives to cell culture for the detection and identification of viral strains [3]. For

example, the differential identification of strains within a species is possible based on the difference in the size of the amplified PCR product (amplicon) detectable by gel electrophoresis ([4] or single-strand conformational polymorphism (SSCP) [5]. Indeed, we have utilized agarose gel electrophoresis following RT-PCR using primer pairs straddling a 14 base insertion at the non-coding region of some HAV genomes to identify specific cytopathic strains from non-cytopathic strains of HAV [4, 6]. We also reported the use of SSCP analysis following *Alu* 1 or *Hinf* 1 digestion of amplicons generated from the 3' end of the viral genome to provide differential identification of multiple HAV strains [5]. However, SSCP is a multi-step procedure involving radiolabeling of restriction fragments prior to electrophoretic separation of individual DNA strands. Consequently, this procedure works best when the restriction fragments are small enough to provide sufficient single-stranded DNA separation for effective strain identification. For genetically well-conserved viruses such as HAV, the region to be amplified for SSCP analysis has to be carefully chosen in order to represent areas of reasonable diversity [7, 8]. Due to these considerations, it has been preferable to sequence the PCR amplified DNA fragment in order to specifically identify the genotypes or strains of the viruses. While sequencing amplified PCR products is considered a precise technique for identification, PCR amplification of a mixed population of target sequences may be biased in favor of a dominant (by copy number) target such that subsequent sequence analysis may not reveal the presence of other closely related target sequences in starting populations. Putative mixed virus populations (e.g. of the same or different species) can exist in isolates obtained from environmental and infected-host samples particularly those resulting from RNA virus replication that

\*Address correspondence to this author at the Division of Molecular Biology, Office of Applied Research and Safety Assessment (OARSA), Food and Drug Administration, 8301 Muirkirk Road, HFS-025, Laurel, MD, 20708, USA; Tel: +1-301-210-7812; E-mail: biswendu.goswami@fda.hhs.gov

<sup>§</sup>Present Address: Visiting Professor, Department of Biology, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA

is known to generate a sub-population of “quasi-species” [9]. Therefore, a threshold number of RNA molecules must have the same specific mutation in order to be unambiguously detectable by RT-PCR and sequencing, due to possible inhibition of amplification of a less abundant template by template competition [10]. Conversely, the dominant mutation present in a population may be preferentially amplified, and therefore, sequence analysis would represent the dominant mutant [11]. Therefore, while sequencing remains a “gold-standard” for target sequence identification, the identification of multiple viral species or tracking species mutations necessitated the development and application of a broader approach to identification prior to undertaking sequence analysis.

As an alternative to sequencing, Proudnikov *et al.* [12] applied a hybridization-based technique to the detection of genetic variants of poliovirus within a virus population or among viral strains. Oligonucleotide probes are synthesized and then immobilized on a solid surface. A target consisting of amplified viral complementary DNA (cDNA) then labeled and hybridized to the immobilized probes and the hybridization to the individual probes detected [12]. The presence of a change in the nucleotide sequence in the target is detected by the absence or the reduction of hybridization to the wild type probes around the change, or by the ratio of the signals generated by a mutant against a reference strain. Modifications of the above technique including the use of amplified viral complementary RNA (cRNA) were used to identify genetic variations arising during cultivation of a vaccine strain of poliovirus and the emergence of vaccine derived poliovirus in immunized patients showing signs of vaccine associated paralytic poliomyelitis [13, 14]. Application of this procedure was restricted, however, to identifying known mutations in specific virus strains.

Advances in microarray technology have allowed the identification of genetic variability over very long stretches of DNA in bacterial genomes [15]. These newly developed high density microarrays contain thousands to hundreds of thousands of oligonucleotide probes, instead of a few dozen, in a single array thereby expanding the power of identification [15-17]. In the current investigation we report the design and use of a high density oligonucleotide microarray for the identification of HAV and coxsackievirus (CV), both food-borne human pathogens. Our results indicate that the microarray hybridization technique can be applied to the identification of viruses of differing genus and species present in a sample and detect single nucleotide polymorphisms (SNP) to identify closely related viral strains belonging to the same species.

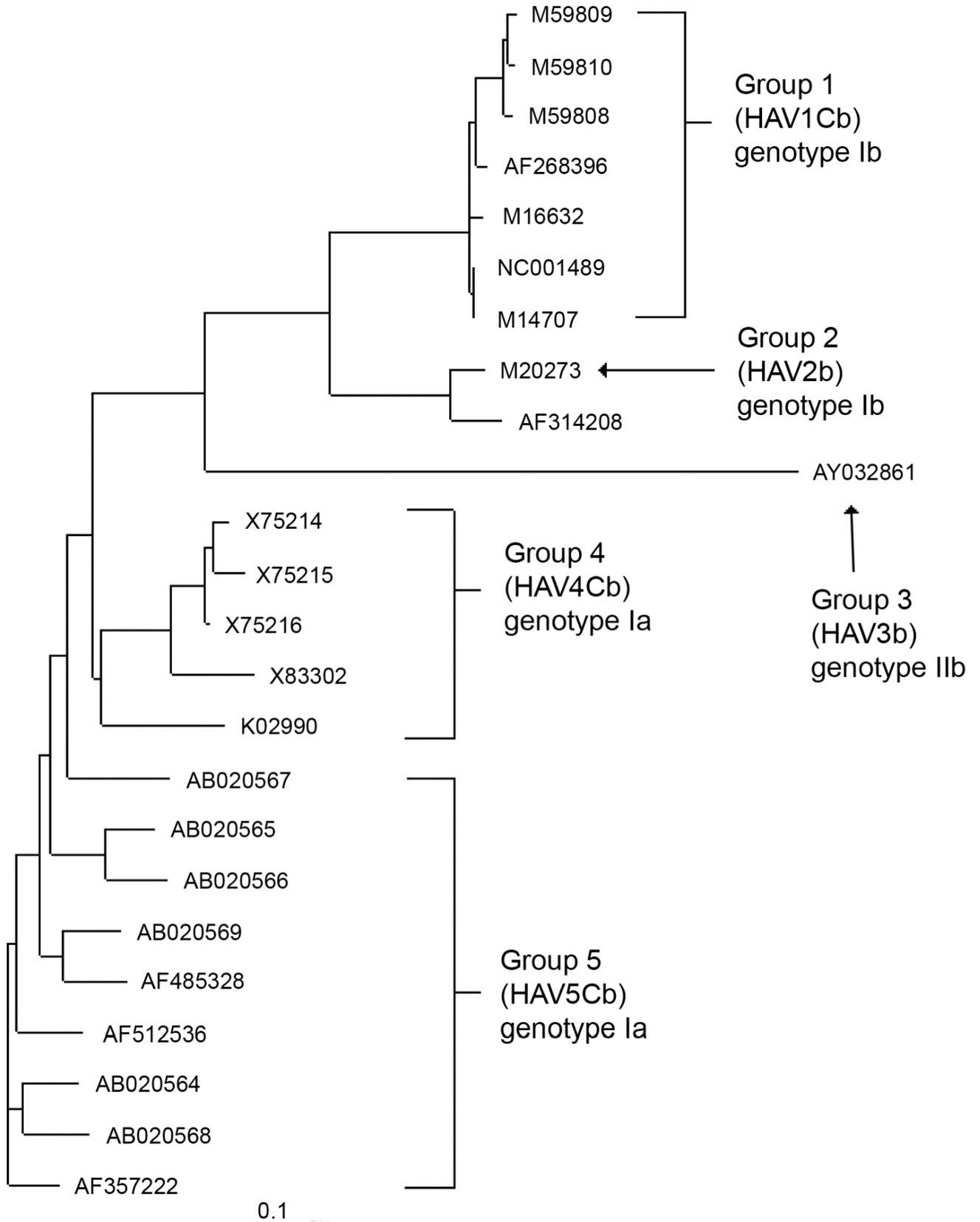
## MATERIALS AND METHODS

**Viruses and Plasmids.** Hepatitis A virus strains HM175/clone 1 and 18f, and coxsackievirus (CV) serotypes B1, A3 and A5 strains used in this study were obtained from ATCC (Manassas, VA) and further grown in FRhK4 cells [18]. The plasmid pHAV/7 contains a full length cDNA copy of wild type HAV strain HM175 cloned into the vector pGEM-1 [19] that was grown and purified as previously described [20]. HM175 clone1 and 18f are culture-adapted strains derived from continuous culture passage of the wild-type strain HAV HM175 [21].

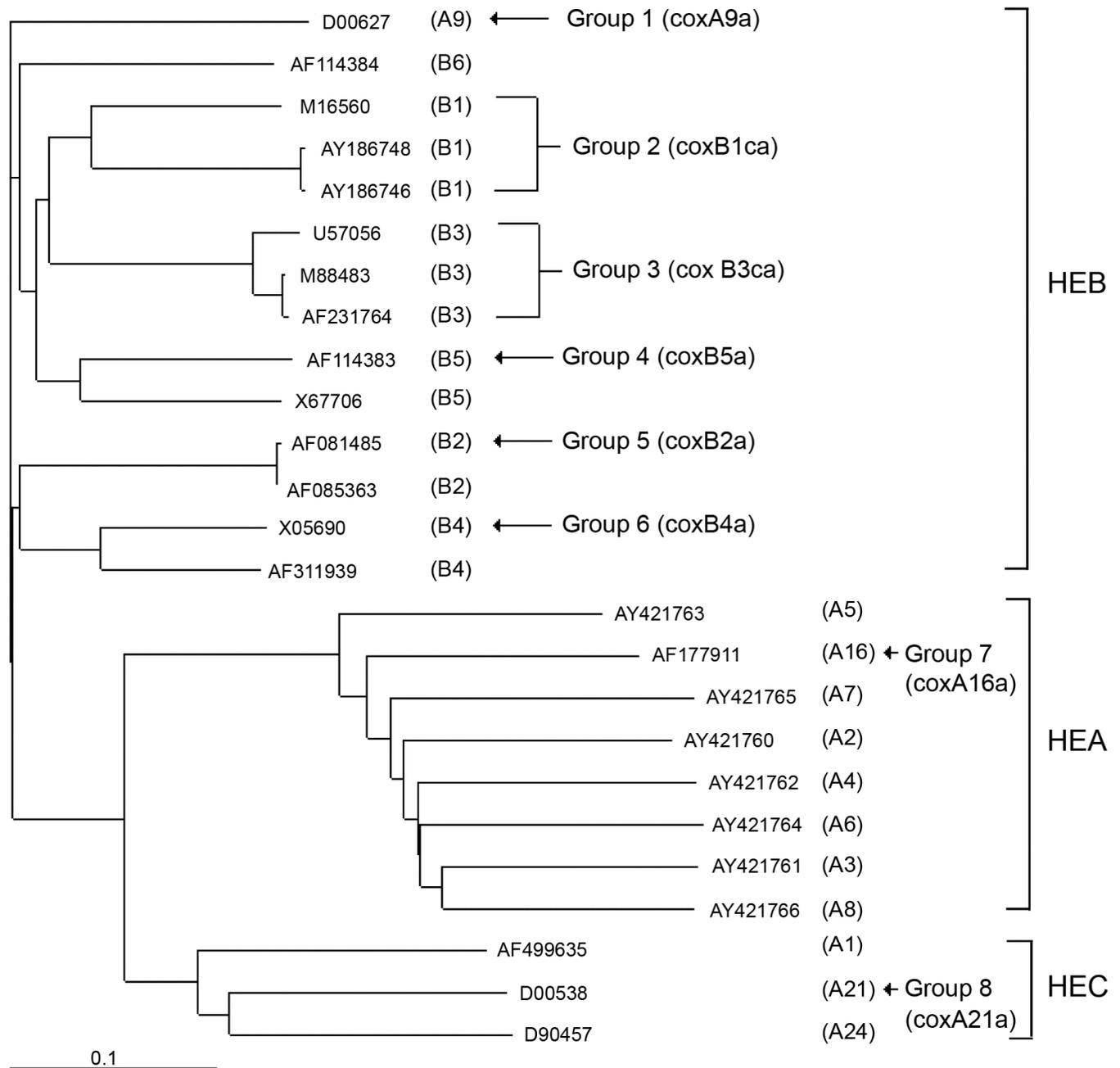
**Oligonucleotide Arrays.** All microarrays used in this study were manufactured by NimbleGen Systems Inc. (Madison, WI) using a maskless array synthesis (MAS) technology for *in situ* synthesis of DNA oligonucleotides directly onto glass microscopy slides [16, 17]. Oligonucleotide design was based on available complete viral genome sequences obtained from GenBank for CV (n=25), HAV (n=23), *Norovirus* genogroup I (n=4), *Norovirus* genogroup II (n=21), rotavirus (various species) segments 3 (n=11), 4 (n=19), 8 (n=11), and 11 (n=12) where n equals the number sequences obtained for each virus group. All genomic sequences within a virus group were aligned using CLUSTALX [22], and dendrograms were generated and consensus sequences constructed based on these analyses. Examples of these dendrograms are shown for HAV and CV (Figs. 1, 2, respectively). For the purpose of generating representative viral genomic sequences on which to base subsequent oligonucleotide designs, the HAV strains were clustered into 5 groups whose viral genome sequences were constructed as follows: i) a consensus sequence based on the seven genotype Ib (i.e. genotype I, subgenotype b) strains that clustered into group 1 which includes the HAV HM175/wt strain (M14707), ii) a sequence derived from M20273 based on the pairing of M20273 and AF314208 (genotype Ib sequences in group 2), iii) a sequence derived from the single HAV genotype II sequence (IIb) available (AY032861) and assigned as group 3, iv) two consensus sequences based on either cluster group 4 or 5 derived from fourteen genotype Ia sequences that were clustered into either of these two groups. The three consensus sequences representing cluster groups 1, 4 and 5 were obtained following a group sequence alignment and the assignment of the most frequently occurring nucleotide at positions containing nucleotide differences. The clustering of either one or two sequences within a group (as in groups 3 and 2, respectively) resulted in the selection of a single sequence representing that group. Due to the highly diverse (genetic) nature of the CV genome sequences, clustering of strains for generating a group consensus sequence was only done for serotype strains B1 and B3 (groups 1 and 2, respectively). Four additional unique strain sequences were selected as representative sequences for broadly clustered strains identified as groups 3-6. Viral genomic sequences (approximately 3000 bases) from either the 3' end of the HAV genome group sequences or the 5' end of the CV genome group sequences were submitted for design of a tiling oligonucleotide array consisting of oligonucleotides of length 29, starting at every 5<sup>th</sup> base in every sequence, resulting in an overlap of 24 bases in two consecutive oligonucleotides [15]. Similar methods were applied to the development and tiling of oligonucleotides as probes for norovirus and rotavirus sequences on the array. The resulting array contained approximately 13,000 viral probes.

## Reverse Transcription and PCR of Viral Genomes.

All reverse transcription (RT) reactions were completed using RNA templates obtained from linearized plasmid pHAV/7 transcribed *in vitro* with SP6 polymerase, total cellular RNA (1 µg) isolated from virus infected cells using the RNA AqueousKit (Ambion, Austin, TX), or viral genomic RNAs (equivalent to 5 x 10<sup>6</sup> infectious particles) isolated directly from clarified tissue culture supernatants using the



**Fig. (1).** Dendrogram showing the grouping of HAV strains based on their genetic relatedness for developing viral probe sequences to be used for oligonucleotide design. HAV strains are identified by their GenBank accession number and their respective complete sequences were used to generate the dendrogram. Brackets encompass strain sequences selected to derive a group consensus sequence while arrows identify group individual (i.e. non-consensus) strain sequences for probe set development. Group sequences are designated by numbers 1-5 followed by the probe set identifier (within parentheses), and the genotype (I or II) and subgenotype (a or b) designation.



**Fig. (2).** Dendrogram showing the grouping of CV serotype strains based on their genetic relatedness for developing viral probe sequences to be used for oligonucleotide design. CV strains are identified by their GenBank accession number and their respective complete sequences were used to generate the dendrogram. CV serotypes are given in parenthesis following the accession numbers. Brackets encompass strain sequences selected to derive a group consensus sequence while arrows identify group individual (i.e. non-consensus) strain sequences for probe set development. Group sequences are designated by the numbers 1-8 followed by the probe set identifier (within parentheses). Brackets (outer right) indicate which human enterovirus species (HEA, HEB and HEC) are represented by the CV serotype strains used to develop the dendrogram.

RNeasy Micro Kit (Qiagen, Valencia, CA); a mixture of oligo(dT)<sub>15</sub> and random hexamers (pdN6) as primers; and AMV reverse transcriptase (Promega, Madison, WI) as previously described [4, 20]. *In vitro* transcribed and infected cell RNA templates represent *in vitro* and *in vivo* replication, respectively. PCR amplification with HAV or CV specific primers was carried out in 50µl reactions using 5µl of each RT reaction as template or 5ng of pHAV/7 plasmid DNA as previously described [20]. PCR products (5µl) were analyzed

by agarose gel electrophoresis to confirm authenticity of product formation (data not shown).

**PCR Primers.** Two primers, 3399 - 3423 (forward) and 7084 - 7105 (reverse), were used to amplify an approximately 3.7 kb region of the HAV genome [4, 6, 20]. Tables 1 and 2 show the sequences around the primer binding sites of selected HAV strains represented on the array. Tables 3 and 4 contain the sequence alignments at the forward and reverse

**Table 1. Genomic Sequence Alignment of Selected HAV Strains Around the Forward Primer**

Strain <sup>a</sup>	Sequence <sup>b</sup>
M14707	TGATTAGATTGAATGATGAGAAATGGACAGAAATGAAGGATGACA
M59808	TGATTAGATTGAATGATGAGAAATGGACAGAAATGAAGGATGACA
M59809	TGATTAGATTGAATGATGAGAAATGGACAGAAATGAAGGATGACA
M59810	TGATTAGATTGAATGATGAGAAATGGACAGAAATGAAGGATGACA
M16632	TGATTAGATTGAATGATGAGAAATGGACAGAAATGAAGGATGACA
X75214	TGGT <b>C</b> AGATTGAATGATGAGAAATGGACAGAAATGAA <b>A</b> GATGATA
X75215	TGGT <b>C</b> AGATTGAATGATGAGAAATGGACAGAAATGAA <b>A</b> GATGATA
X83302	TGGT <b>C</b> AGATTGAATGATGAGAAATGGACAGAAATGAA <b>A</b> GATGATA
K02990	TGAT <b>C</b> AGATTGAATGATGAGAAATGGACAGAAATGAAGGATGATA
AB020564	TGGT <b>C</b> AGATTGAATGATGAGAAATGGACAGAAATGAA <b>A</b> GATGATA
AB020565	TGGTTAGATTGAATGATGAGAAATGGACAGAAATGAAGGATGACA
AB020566	TGGT <b>C</b> AGATTGAATGATGAGAAATGGACAGAAATGAAGGATGACA
AB020567	TGAT <b>C</b> AGATTGAATGATGAGAAATGGACAGAAATGAAGGATGATA
AB020568	TGGTTAGATTGAATGATGAGAAATGGACAGAAATGAA <b>A</b> GATGATA
AB020569	TGATTAGATTGAATGATGAGAAATGGACAGAAATGAAGGATGATA
forward primer	GAATGATGAGAAATGGACAGAAATG

<sup>a</sup>Strains are identified by their GenBank accession number.

<sup>b</sup>HAV sequence alignment is presented as the positive (sense) genomic strand in 5' to 3' orientation. The primer sequence and nucleotide identity is based on the genomic sequence and nucleotide numbering of HM175 strain 18f (M59808) at nucleotide positions 3399 to 3423.

**Table 2. Genomic Sequence Alignment of Selected HAV Strains around the Reverse Primer**

Strain <sup>a</sup>	Sequence <sup>b</sup>
M14707	AAAGTCAATTCCGAAACTGGTTTCAGCTGAGGTACATTCTTG
M59808	AAAGTCAATTCCGAAACTGGTTTCAGCTGAGGTACATTCTTG
M59809	AAAGTCAATTCCGAAACTGGTTTCAGCTGAGGTACATTCTTG
M59810	AAAGTCAATTCCGAAACTGGTTTCAGCTGAGGTACATTCTTG
M16632	AAAGTCAATTCCGAAACTGGTTTCAGCTGAGGTACATTCTTG
X75214	AAAGTCAATTCT <b>G</b> AAAC <b>G</b> GGCTTCAGT <b>T</b> GGGG <b>C</b> ACATTT <b>T</b> TA
X75215	AAAGTCAATTCT <b>G</b> AAAC <b>G</b> GGCTTCAGT <b>T</b> GGGG <b>C</b> ACATTT <b>T</b> TA
X83302	AAAGTCAATTCCGAAACTGGTTTCAGCTGAGGTACATTCTTG
K02990	AAAGTCAATTCT <b>G</b> AAACT <b>G</b> GGCTTCAGT <b>T</b> GAGG <b>C</b> ACATTT <b>T</b> TA
AB020564	AAGGTCAATTCT <b>G</b> AAACT <b>G</b> GGCTTCAGT <b>T</b> GAGG <b>C</b> ACATTT <b>T</b> TA
AB020565	AAAGTCAATTCT <b>G</b> AAACT <b>G</b> GGCTTCAGT <b>T</b> GAGG <b>C</b> ACATTT <b>T</b> TA
AB020566	AAAGTCAATTCT <b>G</b> AAACT <b>G</b> GGCTTCAGT <b>T</b> GAGG <b>C</b> ACATTT <b>T</b> TA
AB020567	AAAGTCAATTCT <b>G</b> AAACT <b>G</b> GGCTTCAGT <b>T</b> GAGG <b>C</b> ACATTT <b>T</b> TA
AB020568	AAAGTCAATTCT <b>G</b> AAACT <b>G</b> GGCTTCAGT <b>T</b> GAGG <b>C</b> ACATTT <b>T</b> TA
AB020569	AAAGTCAATTCT <b>G</b> AAACT <b>G</b> GGCTTCAGT <b>T</b> GAGG <b>C</b> ACATTT <b>T</b> TA
reverse primer	CCGAAACTGGTTTCAGCTGAGG

<sup>a</sup>HAV strains are identified by their GenBank accession number.

<sup>b</sup>HAV sequence alignment is presented as the reverse complement (antisense) of the genomic strand in 5' to 3' orientation. The primer sequence and nucleotide identity is based on the genomic sequence and nucleotide numbering of HM175 strain 18f (M59808) at nucleotide positions 7105 to 7084.

primer binding sites for selected CV strains. The reverse primer for CV is degenerate owing to sequence differences among strains in this region [23]. These primers amplify a 746 bp fragment from several B and A strains (data not shown).

**Labeling of PCR Products and Hybridization.** PCR products were purified using a spin column procedure [Qiagen or Stratagene, (La Jolla, CA)]. One µg of each purified PCR product was labeled with biotin-dUTP in a primer extension reaction using random hexamers and Klenow po-

**Table 3. Genomic Sequence Alignment of Selected CV Strains Around the Forward Primer**

Strain <sup>a</sup>	Sequence <sup>b</sup>
X05690	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
AF311939	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
AF114384	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
AF081485	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
AF085363	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
AY186748	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
AY186746	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
M16560	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
M88483	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
AF231764	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
U57056	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
D00627	TCCTCCGGCCCCCTGAATGCGGCTAATCCAACTGCGGA
AF114383	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
X67706	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
D00538	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACCA CGGA
D90457	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACCAT GGA
AF177911	TCCTCCGGCCCCCTGAATGCGGCTAATCCTAACTGCGGA
forward primer	CCCCTGAATGCGGCTAATCC

<sup>a</sup>CV strains are identified by their GenBank accession number and includes both A and B serotype strains (refer to Fig. (2) for serotype identification).

<sup>b</sup>CV sequence alignment is presented as the positive (sense) genomic strand in 5' to 3' orientation. The primer sequence and nucleotide identity is based on the genomic sequence and nucleotide numbering of CVB1 (M16560) at nucleotide positions 454 to 473.

**Table 4. Genomic Sequence Alignment of Selected CV Strains around the Reverse Primer**

Strain <sup>a</sup>	Sequence <sup>b</sup>
X05690	CATTTCCTGACAATGCATCTGGGAACCTCCACCACCACCCC
AF311939	CATCTCTGACAGCGCATCTGGGAACCTCCACCACCACCCC
AF114384	CATATTGGACAGCGCATCAGGAAACTTCCACCACCACCT
AF081485	CATGTCGTAGTAGTGCCTGGGAAACTTCCACCACCAACCA
AF085363	CATGTCGTAGTAGTGCCTGGGAAACTTCCACCACCAACCA
AY186748	CATTTCGATAGCGCATCTGGCAGTTTCCACCACCACCCA
AY186746	CATTTCGATAGCGCATCTGGCAGTTTCCACCACCACCCA
M16560	CATTTCGTAAGTGCATCTGGTAACCTCCACCACCAACCT
M88483	TAAGTTCGACAAAGCATCAGGCAGCTTCCACCACCATCCT
AF231764	TAAGTTCGACAAAGCATCAGGCAGCTTCCACCACCATCCT
U57056	TAAGTTCGACAAAGCATCAGGCAGCTTCCACCACCATCCT
D00627	CATATCACTAAGCGCTTCTGGGAACCTCCACCACCACCCC
AF177911	TACTTCTGTCAAAAACATCAGGGAATTTCCAATACCATCCC
reverse primer	TCDGGNARYTTCCACCACCA

<sup>a</sup>CV strains are identified by their GenBank accession number.

<sup>b</sup>CV sequence alignment is presented as the reverse complement (antisense) of the genomic strand in 5' to 3' orientation. The primer sequence and nucleotide identity is based on the genomic sequence and nucleotide numbering of CVB1 (M16560) at nucleotide positions 1199 to 1180. The primer is degenerate where D equals A, G, or T; N equals A, G, C, or T; R equals A or G; and Y equals C or T, respectively.

lymerase (Exo<sup>o</sup>). Labeled products were purified by spin column chromatography, and concentrated by centrifugation through Microcon<sup>®</sup> (Millipore, Billerica, MA) filters. Biotin-labeled DNA was denatured in a total volume of 20µl of hybridization solution containing 5XSSC, 0.1%SDS, 5µg poly A, and 5µg human Cot-1 DNA and 6 µl used per hybridization reaction per well of a 12 well sample pod (NimbleGen Systems, Inc.). The microarray slide (NimbleGen) was laid on top (oligonucleotide side down) of the sample

pod and held in place in a metal cassette provided by the manufacturer. Hybridization was carried out for 12h at 42 °C. The slides were washed sequentially with 2XSSC/0.1%SDS, and 0.1XSSC/0.1%SDS at 42 °C then distilled-deionized water at room temperature. The slides were then stained with a Cy3-streptavidin conjugate (Amersham Biosciences, Piscataway, NJ) as described in Jackson *et al.* [15].

**Data Extraction and Analysis.** Hybridized, Cy3-stained microarrays were scanned using an Axon GenePix<sup>®</sup> 4200A scanner at 5  $\mu\text{m}$  resolution using a 532 nm laser. Fluorescence intensities of each feature (oligonucleotide probe) were extracted utilizing NimbleScan<sup>™</sup> software (NimbleGen Systems Inc), and all subsequent data analyses were performed using MS Excel. Data were analyzed independent of comparison to a reference strain assuming that each virus strain is unique. Following normalization for background fluorescence, the fluorescent intensity of each probe (normalized probe intensity) was plotted against the genome position of each probe to generate a hybridization profile for each viral strain [15, 17]. To generate the average probe intensity for each probe set per hybridized virus strain, the sum of all normalized probe intensities for individual probes within a probe set (i.e. set of probes derived from an individual strain or group sequence) was divided by the number of probes within that set [15].

## RESULTS

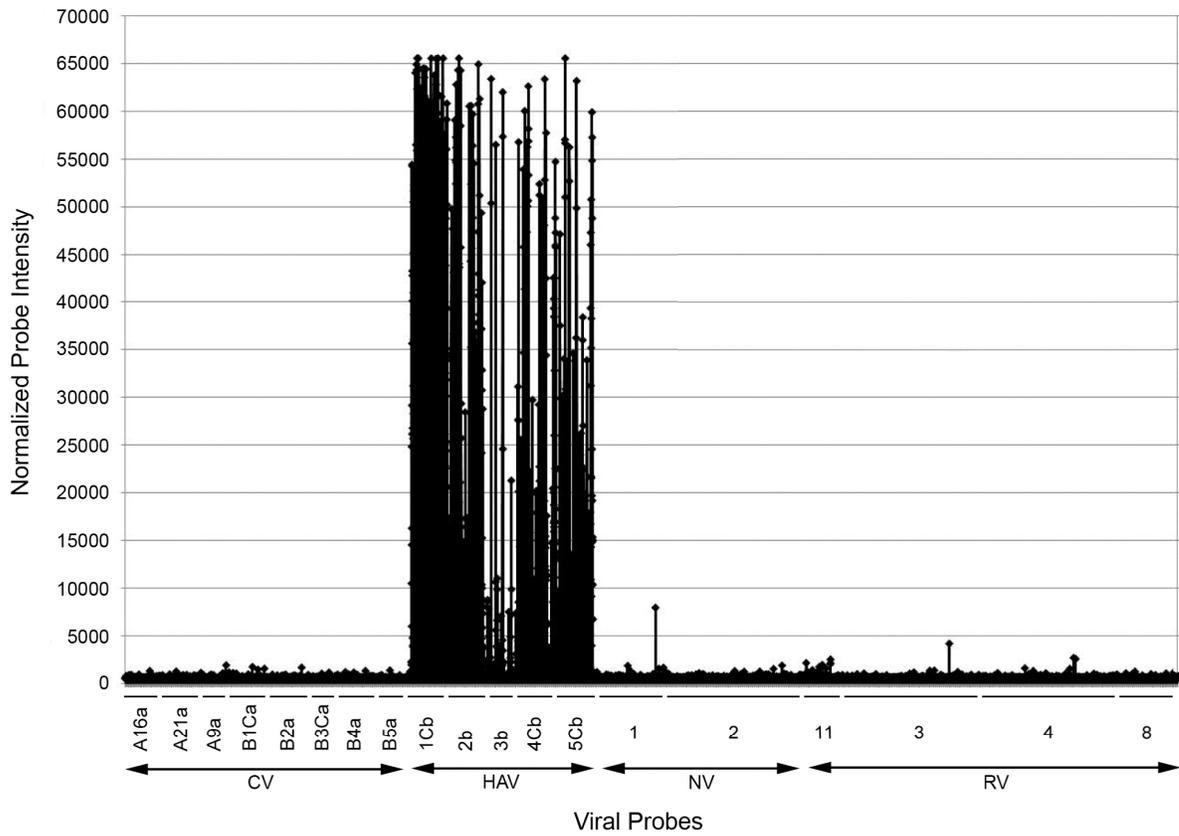
As members of the positive-stranded RNA virus family *Picornaviridae*, Hepatitis A virus and coxsackievirus belong to the genera *Hepatovirus* and *Enterovirus*, respectively. At the nucleotide level there is substantial genetic diversity between these two groups with a greater within group diversity observed for coxsackieviruses than for hepatitis virus strains. Indeed, the coxsackievirus genomes are much less conserved and mutations are distributed throughout the genome [23]. Sequence analyses of small segments of different strains of HAV have led to the recognition of six genotypes (I to VI) of this virus [24]. Genotypes I, II and III have been further subdivided into two subgenotypes, a and b [24, 25]. Within each genotype the strains have greater than 85% sequence homology, whereas subgenotypes may differ from each other in up to 7.5% of nucleotide positions. Genotype I is the most prevalent HAV genotype world wide [25]. For the present investigation, complete HAV genome sequences available at the time of chip design and construction belonging to genotypes I and II were clustered into five groups (Fig. 1). Two of these groups contain subgenotype Ia (groups 4 and 5) while sequences belonging to subgenotype Ib were clustered into two groups (1 and 2) and a single sequence of genotype II was designated group 3 (Fig. 1).

As members of the genus *Enterovirus*, the Human enterovirus A and Human enterovirus C virus species comprise most of the species assigned to CV serotype A strains with the exception of serotype A9 [23]. Coxsackievirus serotypes B1-6, coxsackievirus serotype A9 (CVA9), enteroviruses 69 and 73 and the majority of echoviruses are classified within the Human enterovirus B species that is one of the largest species group in the family *Picornaviridae* [26]. Fig. (2) shows the CLUSTALX analysis of the different coxsackievirus serotype strains used for designing the tiling array. For B1 and B3 strains a consensus sequence was developed based on the sequences available for individual members of these groups, whereas single sequences were used for the strains B4, B5, A16 and A21, due to extensive sequence variations between individual members of these groups [23, 26]. Since sequence diversity among the CV strains was distributed over the entire length of the genome, the first 3000 nucleotides were determined to contain sufficient sequence diversity to identify a strain without ambiguity.

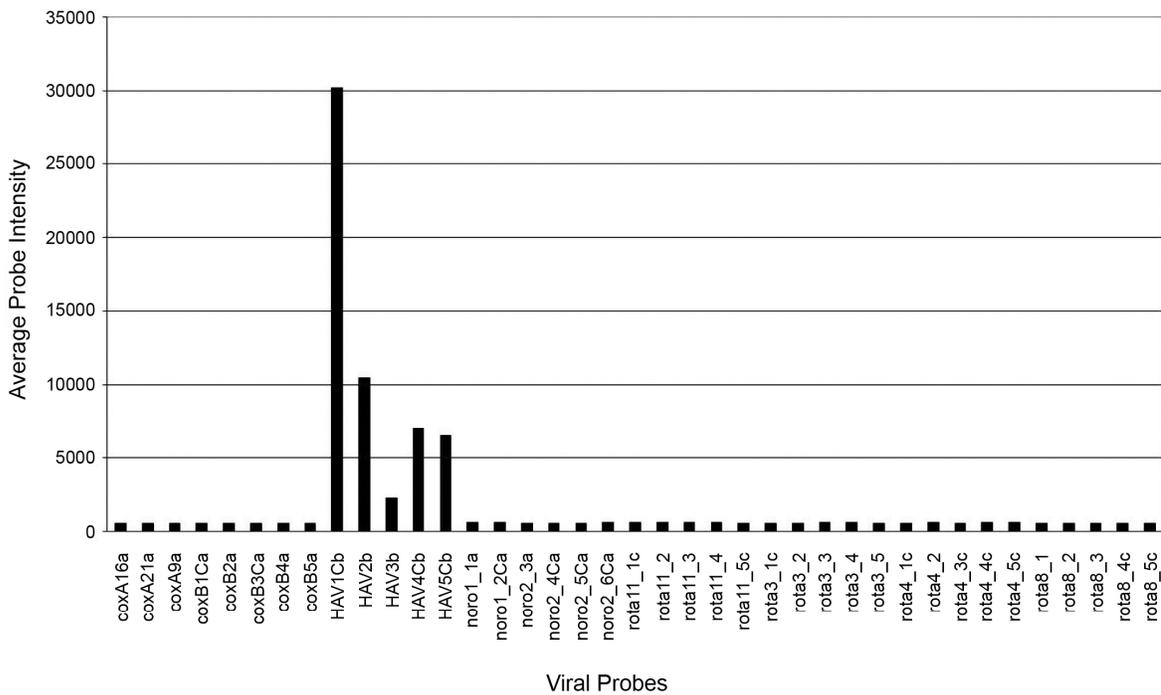
**Identification of HAV Genotype by Microarray Hybridization.** Fig. (3) shows the hybridization profile obtained with a target synthesized by PCR amplification of the plasmid pHAV/7. This plasmid contains a copy of the entire HAV sequence of wild-type HAV strain HM175 [19, 27] that originated from an Australian outbreak, and was designated as genotype Ib by subsequent sequence analysis [24, 28]. The hybridization signals (normalized probe intensities) produced a profile indicating areas of intense hybridization at the position where the HAV sequences are clustered in the array. However, variations in the intensity of hybridization can be observed within these sequences, where the target hybridization intensities against group 1 probes (hav1Cb) differ from probes derived from groups 2 through 5 (hav2b, hav3b, hav4Cb and hav5Cb) sequences. This is more clearly observed in Fig. (4), where the normalized probe intensities for individual probes within each group sequence present in the array were converted to average probe intensities and plotted for the target. The plot reveals that the HAV genotype 1b (HM175 wild-type) target hybridized most efficiently to probes from genotype Ib, group 1 consensus sequence (hav1cb). These results are consistent with the fact that the viral genome sequence for HAV HM 175 wt strain (14707) is a member of, and therefore most closely related to, group 1 derived probe sequences. Probes representing a closely related HAV Ib strain from group 2 (hav2b) hybridize the target about two thirds as efficiently, while probes from the more genetically distant genotype II virus (hav3b) hybridize with the least intensity. The other probe groups (hav4cb and 5cb) both representing genotype Ia consensus sequences (Fig. 1, cluster groups 4 and 5) hybridize less efficiently than genotype Ib. Given the readily observable differences in both the normalized and average signal intensities among the genotype group sequence probes (groups 1-5) following genotype Ib target hybridization, and the fact that viruses belonging to different subgenotypes can differ by as much as 7.5% in sequence [21, 24, 28], the data in Fig. (4) indicate that it is possible to identify HAV strains at the level of both genotype and subgenotype with this type of array.

To further explore genotype/subgenotype differentiation, different HAV strains belonging to the same subgenotype Ib sequence (Fig. 1, group 1) were hybridized to the array. As shown in Fig. (5), HAV strains HM175 wt, clone 1 and 18f hybridize most efficiently to genotype Ib (consensus group 1) probes (hav1Cb). Lower efficiencies of hybridization are observed for all three targets against all other probe sets. These results reflect a greater target specificity for the probe set that contains target member sequences than for the other genotype Ib derived probe set (hav2b) that does not contain target member sequences (group 2 in Fig. 1). For all target strains, the remaining probe sets yielded signal intensities equivalent to or less than intensities for probe set hav2b. Thus, in support of the interpretation of results of Fig. (4), this array has the potential to discriminate viral targets at the level of both their genotype and subgenotype.

**Differential Analysis of Two Target (HAV) Strain Hybridization Profiles Reveals a Correlation with Known Nucleotide Differences.** It is important to note that despite the variation in average probe intensities for the individual strains against probe set hav1Cb (Fig. 5), the information as presented cannot be used to identify actual target nucleotide



**Fig. (3).** Hybridization profile of wild-type HAV HM175 strain target: normalized probe intensities. PCR amplified HAV target (pHAV/7) sequences were labeled prior to array hybridization as described in Materials and Methods. The probe hybridization values (intensities) were plotted as normalized probe intensities vs all probe sequences. The graphical output was compressed to include all probes represented on the array. As a result, individual probe sequence identifiers are not listed on the plot. Probe set identifiers (under solid line) are given for each of the virus species (under double-headed arrow) represented on the array where CV, HAV, NV and RV represent coxsackievirus, hepatitis A virus, norovirus and rotavirus, respectively.

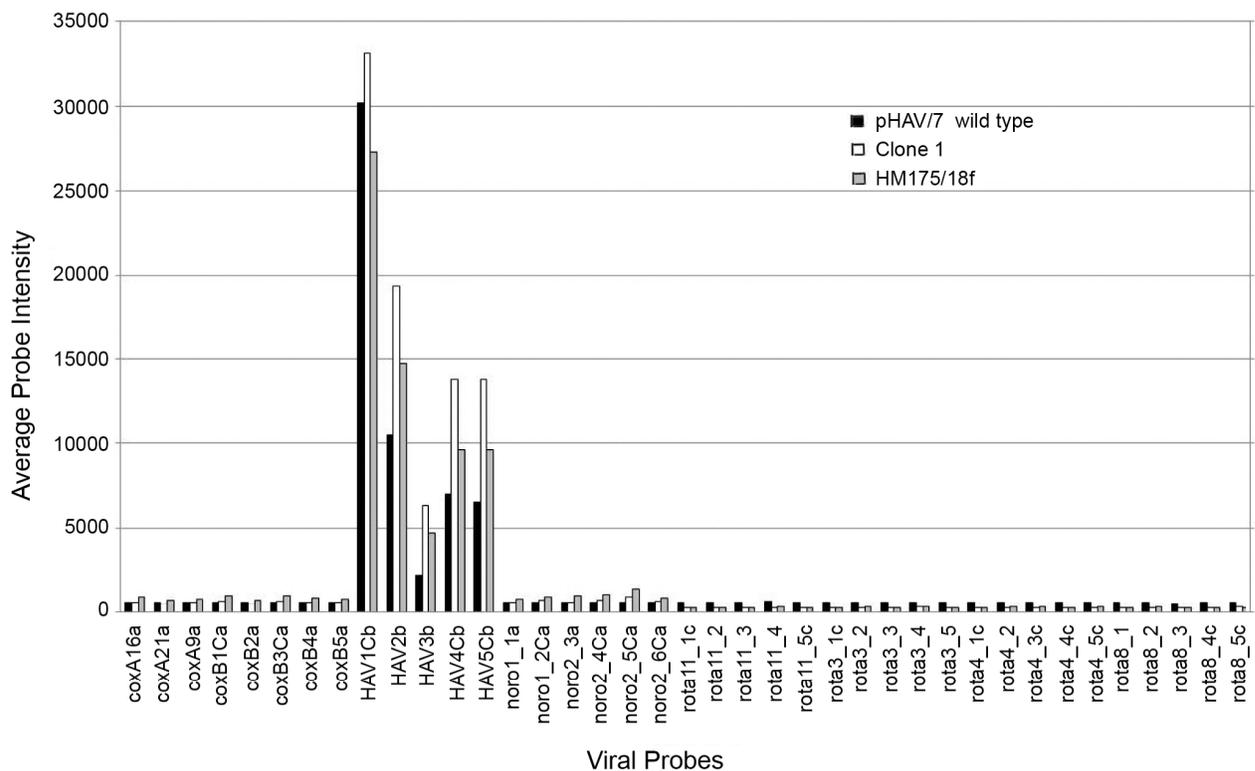


**Fig. (4).** Hybridization profile of wild-type HAV HM175 strain target: average probe intensity. The hybridization data (normalized probe intensities) for HAV HM175 from Fig. (3) was converted to average probe intensities [15] and plotted vs each individual probe set. A given probe set represents all probes derived from their respective sequence group.

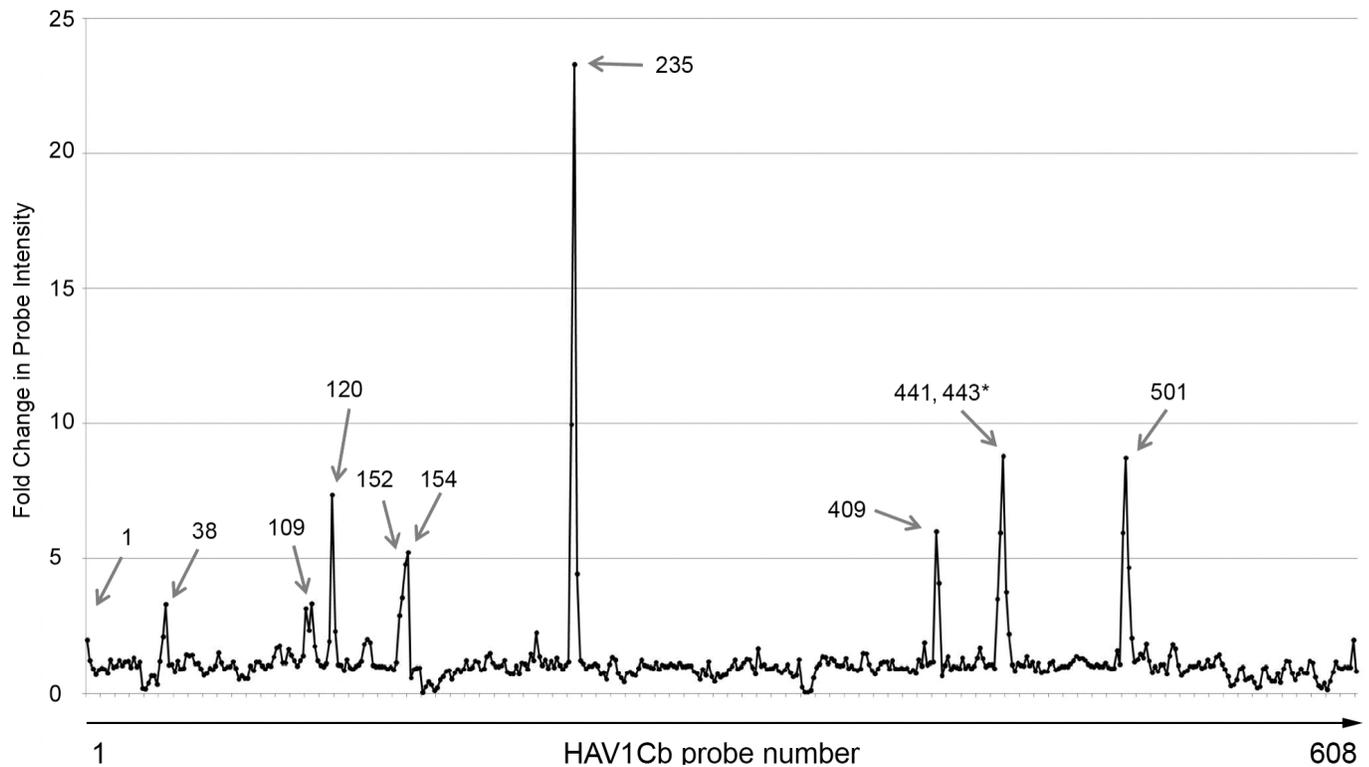
differences. For example, differences in signal height could be attributed to differing hybridization efficiencies between two different experiments. Indeed, a target derived from *in vitro* synthesized RNA from pHAV/7 (representing *in vitro* replication of the viral genome) was indistinguishable from plasmid derived target, or the virus following several rounds of replication in culture except for the peak height (data not shown). We pursued, therefore, an alternative method of analysis because the tiling array design offers the potential to distinguish between these closely related strains following hybridization by i) determining the normalized probe intensities for each target, and ii) plotting the change in signal intensity of hybridization by each target to the same probe set as the ratio (fold-change in probe intensity) vs the individual probes. As discussed by Jackson *et al.* [15], this method of analysis can reveal distinct peaks with defined slopes (above background/signal noise) where changes in signal strength would occur with probes tiled further up or down stream of the nucleotide change. The presence of a mutation in the genome causes a destabilization of a number of probes around the mutation, which can be identified by the appearance of well defined peaks. Therefore, this method of analysis offers the potential to differentiate closely related strains of virus belonging to subgenotype Ib at the level of individual nucleotide differences, thereby producing data that can be used to tell them apart.

In order to complete this analysis, the two different HM175 strains designated clone 1 and the cytopathic 18f strain were again subjected to hybridization and the total normalized intensities of all probes belonging to the different HAV probe groups were plotted as in Fig. (3). Again, we

found no overall differences in the hybridization profile but rather found peaks of hybridization intensities with the strongest hybridization intensities for the group 1 (HAV1Cb) consensus sequence following calculation of average probe intensity (data not shown). The fold-change in intensity between clone 1 and 18f targets was calculated for each probe in the probe set HAV1Cb. As shown in Fig. (6), ten well defined peaks were observed over the range of the HAV1Cb probe set and the probe number that corresponds to each peak was identified. It is important to note that due to the initial size of the graphical analysis output, it was necessary to compress the scale of the x-axis (HAV1Cb probe number) in order to fit all data points within a smaller graph. As a result, analysis of the hybridization (signal) values revealed two features not readily discernable on the graph; i) a probable single peak at probe 109 rather than what appears as two adjacent (overlapping) peaks, and ii) a possible second overlapping peak adjacent to probe 441. Since the HAV1Cb probe set (group 1) is a consensus sequence developed from the alignment of seven strains assigned to this group (Fig. 1), there are nucleotide differences between each group member and the consensus sequence. Plotting the fold-change in intensity between clone 1 and 18f would potentially identify nucleotide sequences in a probe that are identical to clone 1 but not identical to 18f. Indeed, upon comparative analysis of clone 1 and 18f amplified target sequences with HAV1Cb probes set sequence synonymous with the target sequences, one would predict a total of 11 peaks to occur by this method of analysis. We then sought to determine whether the “peak” probes contained nucleotide differences that could be mapped to nucleotide differences [e.g. single-nucleotide po-



**Fig. (5).** Comparison of hybridization profiles for three HAV genotype Ib strain targets. Average signal probe intensities were calculated and plotted following hybridization of targets generated as PCR products from reverse transcription of RNA derived from either *in vitro* transcribed pHAV/7 (black bar), HAV HM175 clone 1 infected cells (white bar) or clarified supernatant from HAV 18f infected cells (grey bar).



**Fig. (6).** Detection of nucleotide differences between two genetically related HAV strains. Average signal probe intensities were calculated following hybridization of targets generated as PCR products from reverse transcription of RNA derived from HAV HM175 clone 1 infected cells or clarified supernatant from HAV 18f infected cells. The amplified targets (3.7 kb) derived from both clone 1 and 18f contain nucleotide sequences synonymous with the first 2.7 kb (probes 1- 543) of the 3.1 kb group 1 (HAV1Cb; Fig. 1) consensus sequence used to develop the HAV1Cb probe set that is comprised of 608 probes. The individual points on the graph represent specific probe numbers; however, due to graphical compression of the original data, there are iterative probes not represented by individual points on the graph. Arrows identify the probe number having the peak intensity difference between clone 1 and 18f where a nucleotide(s) present in the consensus sequence is identical to nucleotide(s) in clone 1 but not identical to nucleotide(s) in 18f.

lymorphisms (SNPs, deletions, or insertions) that exist between clone 1 and 18f (and the probe set). As shown in Table 5, we were able to conservatively detect 10 out of 11 predicted nucleotide changes in the 18f genome identifiable by this method of analysis. It is important to note that these nucleotide changes represent mutations arising in the 18f virus during its emergence as a cytopathic strain from the HM175 noncytopathic strain which were identified by direct sequencing [21]. These results demonstrate a strong correlation between results obtained by direct sequencing and array hybridization and strongly suggest that tiling arrays can be used to detect nucleotide changes instead of sequencing amplified PCR products over a much longer span of the genome in a single experiment.

**Identification of CV Serotype by Microarray Hybridization.** Unlike HAV strains, there is tremendous genetic diversity between CV strains, even within the same species as observed, for example, among serotype B strains although they are all members of HEV species [23, 26]. We sought, therefore, to determine whether this array hybridization technique could be used to identify a CV serotype strain target. A typical hybridization profile with a 746 bp segment amplified from CV strains is shown for CVB1 in Fig. (7, panel A) where the data is presented as average probe intensity for all probes derived from the same group sequence, i.e. probe set. Similar to the results obtained following hybridization

with HAV targets, CVB1 targets hybridized very efficiently and with greatest intensity to probes (coxB1Ca) derived from a consensus sequence based on its own sequence, i.e. serotype B1 strains (Fig. 2, group 2). As indicated by the significantly lower probe intensities, minimal hybridization was observed among the remaining 7 CV probe sets indicating a lower efficiency of hybridization to non-CVB1 sequences represented on the array. In fact, hybridization to probes representing all other (non-CV) viruses was essentially at background signal intensity. The results are consistent with the extensive sequence heterogeneity that exists between the CV serotype A and B virus strains, the members within a serotype (A or B), as well as the probe sets derived from these strains. Importantly, these results demonstrate that even with highly (genetically) diverse viruses, such as coxsackieviruses, this array design can discriminate between strains of the same (or different) virus species. We next sought to determine whether discrimination between virus strains or species was possible when the viral target contains sequences not represented by either an individual or a consensus probe set on the array. To complete this experiment, a 746 bp targets derived from coxsackievirus serotype A3 and A5 strains were hybridized to the array. CVA3 and CVA5 serotype strains are both members of HEA species, however the probes' sequence (group 7, coxA16a) for the species was derived from CVA16 (Fig. 2). Analysis of normalized probe intensities reveal a striking reduction in the overall level of

**Table 5. Detection of Clone 1 to 18f Nucleotide Changes by Hybridization<sup>a</sup>**

Probe <sup>b</sup>			Clone 1 > 18f Nucleotide Change	18f <sup>c</sup> Nucleotide Position
Number	Sequence	Nucleotide Range		
1	GTTATTATATGGCAAAGAGGGGAGGA	4394 - 4422	T > C	4402
38	ACAGATGAGGATGGTCAGATTTTGTCA	4579 - 4607	T > C	4590
109	TTGAAATTAGAAAACAAAACATGACTGAA	4934 - 4962	A > T	4938
120 <sup>d</sup>	TTCAGATGATGATAATGATAGTGCAGTAG	4989 - 5014	GAT > del	4998/4999
152	GGATGGTTTGTGTATAAGCATTCTCCCG	5146 - 5174	T > G	5152
154	TGTATAAGCATTCTCCCGCAAAGAGGAA	5156 - 5184	G > A	5174
235	GAGATATTACTCAGCATTTTATTAAGAAA	5561 - 5589	C > G	5572
409	ATGAATTGAGACCATTAGAGAAAGTGTG	6431 - 5559	A > G	6441
441	TCCTGATAGACAGTGGGATGAATTATTTA	6591 - 6619	G > A	6599
443 <sup>e</sup>	CAGTGGGATGAATTATTTAAACAATGAT	6601 - 6629	T > C	6613
501	TGTGTTTTCCAAGATATTTGGAAAGTCTC	6891 - 6919	C > T	6900

<sup>a</sup>Detection of (putative) nucleotide differences by array hybridization between clone 1 and 18f was initially based on the hybridization profile in Fig. (6).

<sup>b</sup>Probe numbers are from Fig. (6) and represent the oligonucleotide probes whose sequence contains nucleotide change(s) between clone 1 and 18f when the clone 1 target sequence is identical to the probe sequence. Nucleotide changes were identified (grey boxed) based on comparison of the clone 1 and 18f GenBank sequences (accession numbers in Fig. (1)) used to develop the HAV1Cb group 1 consensus probe set (Fig. 1). The probe nucleotide range numbering is defined by the 29-mer probe and corrected to 18f nucleotide numbering from Lemon *et al.* [21].

<sup>c</sup>The nucleotide change and position between clone 1 and 18f as reported by Lemon *et al.* [21].

<sup>d</sup>Probe 120 defines a 26 nucleotide base region of 18f due to the three-base GAT deletion.

<sup>e</sup>This probe identified as a potential peak overlapping with the peak at probe 441.

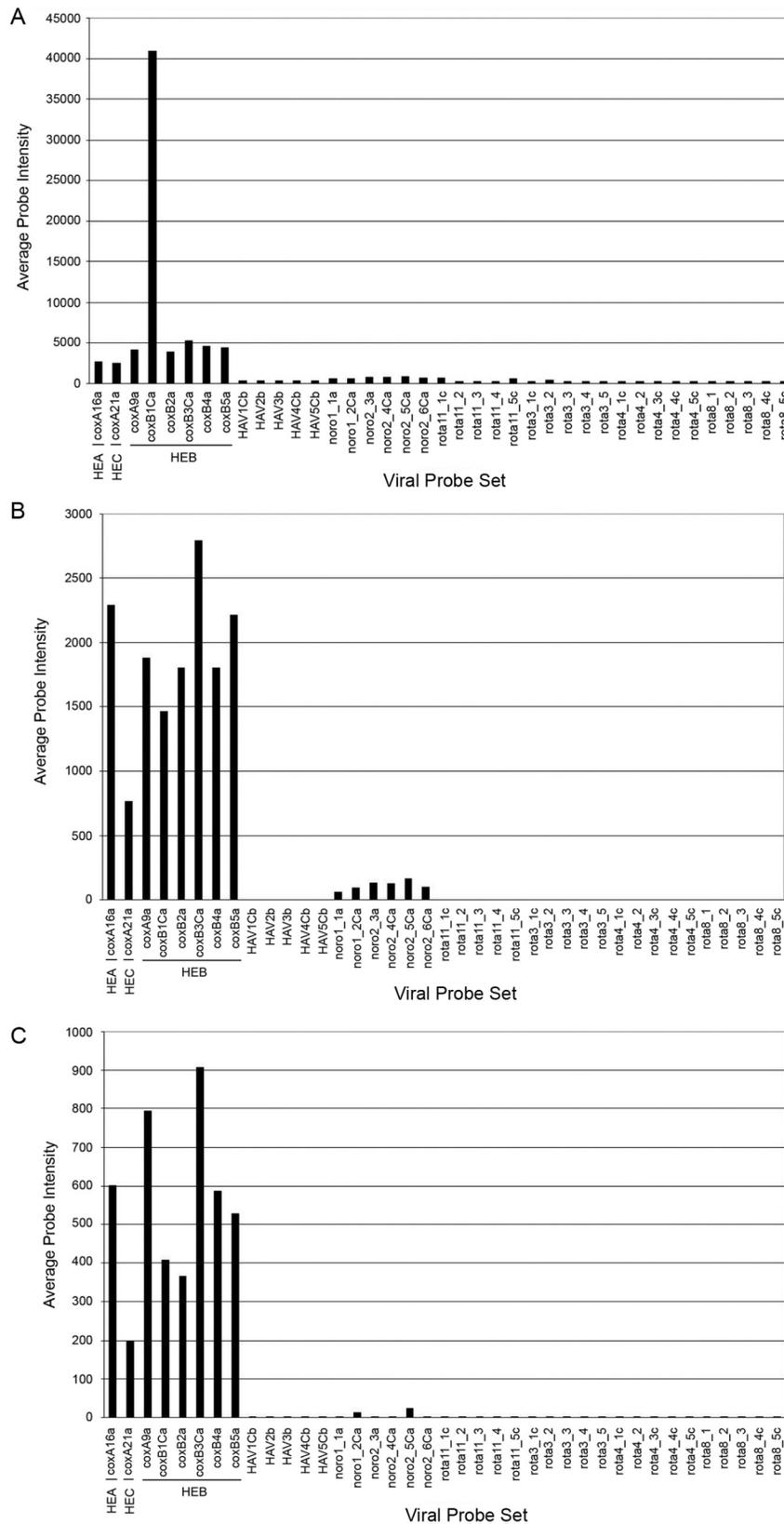
probe hybridization (normalized) intensities for CVA3 and CVA5 derived targets compared to those values obtained following hybridization with a CVB1 target (data not shown). As shown in Fig. (7, panels B and C), this is also observed following conversion to average probe intensity. The peak average probe intensity for these hybridizations is approximately 2750 units and 900 units with CVA3 and CVA5 targets, respectively. The results indicate that in the absence of matching probe sets on the array the sequence heterogeneity between these CV targets and the existing probe sets precludes the establishment of any strong or efficient hybridization to a single probe set. It is important to note, however, that neither of these targets hybridizes with any significance to non-CV probe sets suggesting that the genetic diversity between CV targets and probe sets does not prevent or obscure virus target group (i.e. CV) identification. In addition, these hybridization profiles are not only distinct from B1 (Fig. 7) but also from one another suggesting the possibility that unique hybridization profile patterns (calculated as normalized and/or averaged probe intensity) could be used for CV serotype target identification. The results from Fig. (7) also suggest that in a single experiment it is possible to identify whether a virus belongs to group A or group B. Indeed, identification of coxsackieviruses at the level of serotype strain may be possible without single nucleotide polymorphism (SNP) analysis and limited only by the number of probe sequences/sets present on the array.

## DISCUSSION

Currently, RT-PCR is the most widely used molecular method for the detection and identification of viruses in biological and environmental sources [27, 28]. Identification of genotypes of virus strains are based on the amplification of specific regions of the viral genome using gene specific

primers followed by sequencing of the amplicon by standard procedures [28]. In some instances a preliminary identification is possible using the techniques of single strand conformational polymorphisms (SSCP) or restriction fragment length polymorphisms (RFLP) [5]. Multiplex PCR allows the detection of more than one species of virus in a single analyte [29]. However, these techniques have limitations on sensitivity and versatility, and require extensive prior knowledge of the sequences to be amplified. The requirement of size differences in the amplicons to be analyzed by gel electrophoresis following amplification by multiplex PCR also limits its utility.

Different strains of HAV and many enteric viruses show variable sequence diversity [23, 24, 26]. This allows easy identification of a virus at the genotype level by sequencing discrete segments of the viral genome amplified by RT-PCR. Ideally, sequencing should be done on amplicons that are known to have multiple nucleotide differences between strains. However, designing PCR primers that will capture a significant number of members of that group requires significant sequence homology, and therefore, a relatively variable region flanked by conserved regions is needed for sequence based identification. While for some virus groups such as HAV it is relatively easy to find PCR primers that can capture many members, it is much more difficult with CV genomes due to extreme sequence diversity. The length of the amplified region is another constraint for sequence based identification. Sequencing an amplicon larger than 500 bp generally will require designing multiple primers for sequence walking. Although automated sequencing techniques currently available can be used for rapid sequencing of a moderate sized amplicon, the process is still too time consuming to be used on a routine basis where a quick identification is needed.



**Fig. (7).** Comparison of hybridization profiles of three CV strain targets: average probe intensity. Viral genomic RNAs isolated directly from clarified tissue culture supernatants of infected cells were used for RT followed by PCR amplification and labeling prior to array hybridization as described in Materials and Methods. The hybridization data (normalized probe intensities) were converted to average probe intensities [15] and plotted vs each individual probe set following hybridization with either CVB1 (panel A), CVA3 (panel B) or CVA5 (panel C). The underlined identifies the human enterovirus species (HEA, HEB and HEC) represented by a CV probe set.

We investigated whether hybridization of fluorescently labeled amplified DNA (target) to a microarray containing many oligonucleotide probes representing many different viral genomes can identify a virus without sequencing. Unlike sequencing, these arrays can interrogate thousands of bases of a viral genome in a single experiment [15-17]. We determined the feasibility of this approach by using labeled targets amplified from either the DNA (i.e. as recombinant plasmid) or RNA from several strains of HAV and CV. As shown in Figs. (4-6), a single hybridization experiment using a multi-well array with different samples loaded in different wells of a 12-well sample pod can identify HAV and CVB by the unique profile generated with no ambiguity or cross-hybridization to oligonucleotides representing an unrelated virus. Within the broad genus of *hepatovirus* of which HAV is the only species member, different genotypes which differ from each other by 5% to 8% of base positions (Fig. 1) can be identified (Figs. 4-6). Within the same subgenotype 1b, strains such as wild type HM175 and the cell culture adapted variants including the cytopathic 18f strain differ by only 0.5% of base positions. We have shown that differentiation of these strains is possible by analyzing the ratio of the signal probe intensities generated by the isolates when hybridized to the probe sets present on this tiling array (Fig. 6 and Table 5). A sequence based identification of the same 3.7 kb amplicon would require several sequencing reactions with multiple primers in order to identify nucleotide differences. In addition, mutations accumulating in the HM175 genome during its evolution into the cytopathic 18f strain can be identified by ratio analysis (Fig. 6 and Table 5). Thus, the present array design is suitable for identification of species (e.g. CV and HAV) and HAV subgenotypes since in the latter case the nucleotide differences are very few.

We have also demonstrated that it is possible to distinguish between CVB and CVA strains by virtue of their hybridization profiles. In addition, individual members of A and B groups show distinct and characteristic hybridization patterns. Thus, members of CVA strains such as A3 and A5 can be easily identified not only as belonging to group A CV, but also a genotype A3 or A5. More virus strains need to be examined to determine if the method is applicable to other members of this group.

The more closely the target sequence matches the probe set, the stronger the hybridization signal. Diversity between and among probe sets representing virus strains within a group such as CV increases the power of discrimination (due to heterogeneity) particularly when the target is highly similar to one of the probe sets. This enables discrimination even at least at the strain level (e.g. among strains within the same serotype group such as CV group B serotypes). This level of discrimination is lost when a target whose sequence is not represented by a probe set is not present on the array. Again, this has been shown to be problematic with highly (genetically) diverse viruses such as CV. However, despite the loss of serotype discrimination, the diverse nature of such viruses does still enable the differentiation between virus groups as shown between CVA and HAV, NV, and rotavirus.

## CONCLUSION

Our results show that an oligonucleotide array incorporating thousands of probes representing genomes of multiple

foodborne RNA viruses including multiple hepatitis A virus genotype strains and multiple coxsackievirus serotype (A and B) strains can be used to differentiate between virus members of either genus to identify the genotype/serotype of these viruses by array hybridization assay. Because the large number of probes can bind and detect labeled targets over a much larger area of the viral genomes, producing distinctive signal patterns for each genotype/serotype, the need for large scale sequencing is eliminated for this level of discrimination.

## ACKNOWLEDGEMENTS

We thank Drs. Joseph E. Leclerc and Samantha Wales for critical review of the manuscript and Dr. Wales for assistance with the figures.

## REFERENCES

- [1] Hollinger FB, Emerson SU. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE, Eds. *Fields Virology*. 4th ed. Philadelphia, Lippincott Williams & Wilkins 2001; 685-722.
- [2] Cromeans T, Sobsey MD, Fields HA. Development of plaque assay for a cytopathic, rapidly replicating isolate of hepatitis A virus. *J Med Virol* 1987; 22: 45-56.
- [3] Goswami BB, Kulka M. In: Potter M, Ed. *Food consumption and disease risk. Consumer pathogen interactions*. Boca Raton, CRC press 2006; 343-392.
- [4] Goswami BB, Kulka M, Ngo D, Istafanos P, Cebula TA. A polymerase chain reaction based method for the detection of hepatitis A virus in produce and shellfish. *J Food Prot* 2002; 65: 393-402.
- [5] Goswami BB, Burkhardt W, Cebula TA. Identification of genetic variants of hepatitis A virus. *J Virol Methods* 1997; 65: 95-103.
- [6] Bhattacharya SS, Kulka M, Lampel KA, Cebula TA, Goswami, BB. Use of reverse transcription and PCR to discriminate between infectious and non-infectious hepatitis A virus. *J Virol Methods* 2004; 116: 181-7.
- [7] Hongyo T, Buzard GS, Calvert RJ, Weghorst CM. Cold SSCP: A simple rapid and non-radioactive method for optimized single strand conformation polymorphism analysis. *Nucleic Acids Res* 1993; 21: 3637-42.
- [8] Liu Q, Sommers SS. Restriction endonuclease fingerprinting (REF): A sensitive method for screening mutations in long contiguous segments of DNA. *BioTechniques* 1995; 18: 470-7.
- [9] Domingo E, Martinez-Salas E, Sobrino F, *et al*. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance - a review. *Gene* 1985; 40: 1-8.
- [10] Goswami BB, Koch WH, Cebula TA. Competitor template RNA for detection and quantitation of hepatitis A virus by PCR. *BioTechniques* 1994; 16: 114-21.
- [11] Fakhrai-Rad H, Pourmand N, Ronaghi M. Pyrosequencing: an accurate detection platform for single nucleotide polymorphisms. *Human Mutat* 2002; 19: 479-85.
- [12] Proudnikov D, Kirillov E, Chumakov K, Donlon J, Rezapkin G, Mirzabekov A. Analysis of mutations in oral poliovirus vaccine by hybridization with generic oligonucleotide microchips. *Biologicals* 200; 28: 57-66.
- [13] Cherkasova E, Laassri M, Chizhikov V, *et al*. Microarray analysis of evolution of RNA viruses: evidence of circulation of virulent highly divergent vaccine-derived polioviruses. *Proc Natl Acad Sci USA* 2003; 100: 9398-403.
- [14] Lassri M, Dragunsky E, Enterline J, *et al*. Genomic analysis of vaccine-derived poliovirus strains in stool specimens by combination of full-length PCR and oligonucleotide microarray hybridization. *J Clin Microbiol* 2005; 43: 2886-94.
- [15] Jackson SA, Mammel MK, Patel IR, *et al*. Interrogating genetic diversity of *E.coli* 0157:H7 using DNA tiling arrays. *Forensic Sci Internat* 2007; 168: 183-99.
- [16] Albert TJ, Norton J, Ott M, *et al*. Light-directed 5' → 3' synthesis of complex oligonucleotide microarrays. *Nucleic Acids Res* 2003; 31: e35.

- [17] Wong CW, Albert TJ, Vega VB, *et al.* Tracking the evolution of sars coronavirus using high-throughput, high density resequencing arrays. *Genome Res* 2004; 14: 398-405.
- [18] Kulka M, Chen A, Ngo D, Bhattacharya SS, Cebula TA, Goswami BB. The cytopathic 18f strain of hepatitis A virus induces RNA degradation in FrhK4 cells. *Arch Virol* 2003; 148: 1275-300.
- [19] Ticehurst JR, Racaniello VR, Baroudy BM, Baltimore D, Purcell RH, Feinstone SM. Molecular cloning and characterization of hepatitis A virus cDNA. *Proc Natl Acad Sci USA* 1983; 80: 5885-9.
- [20] Goswami BB, Koch WH, Cebula TA. Detection of hepatitis A virus in *Mercenaria mercenaria* by coupled reverse transcription and polymerase chain reaction. *Appl Environ Microbiol* 1993; 59: 2765-70.
- [21] Lemon SM, Murphy PC, Shields PA, *et al.* Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination. *J Virol* 1991; 65: 2056-65.
- [22] Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997; 25: 4876-82.
- [23] Oberste MS, Penaranda S, Maher K, Pallansch MA. Complete genome sequences of all members of the species human enterovirus A. *J Gen Virol* 2004; 85: 1597-607.
- [24] Costa-Mattioli M, Di Napoli A, Ferre V, Billaudel S, Perez-Bercoff R, Cristina J. Genetic variability of hepatitis A virus. *J Gen Virol* 2003; 84: 3191-201.
- [25] Cohen JI, Ticehurst JR, Purcell RH, Buckler-White A, Baroudy BM. Complete nucleotide sequence of wild type hepatitis A virus: comparison with different strains of hepatic A virus and other picornaviruses. *J Virol* 1987; 61: 50-9.
- [26] Oberste MS, Maher K, Nix WA, *et al.* Molecular identification of 13 new enterovirus types, EV79-88, EV97, and EV100-101, members of the species human enterovirus B. *Virus Res* 2007; 128: 34-42.
- [27] Robertson BH, Jansen RW, Khanna B, *et al.* Genetic relatedness of hepatitis A virus strains recovered from different geographical regions. *J Gen Virol* 1992; 73: 1365-77.
- [28] Nainan OV, Xia G, Vaughan G, Margolis HS. Diagnosis of hepatitis A infection: a molecular approach. *Clin Microbiol Rev* 2006; 19: 63-79.
- [29] Jokel P, Joki-Korpela P, Maaronen M, Glumoff V, Hyypia T. Detection of human picornaviruses by multiplex reverse transcription-PCR and liquid hybridization. *J Clin Microbiol* 2005; 43: 1239-45.

---

Received: February 2, 2009

Revised: February 20, 2009

Accepted: February 27, 2009

© Ayodeji *et al.*; Licensee Bentham Open

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.