Identification of adenovirus-encoded small RNAs by deep RNA sequencing

Hongxing Zhao a,*, Maoshan Chen a,b, Ulf Pettersson a

a Department of Immunology, Genetics and Immunology, Uppsala University, S-751 85 Uppsala, Sweden
b Bioinformatics center, Beijing Genomics Institute, Shenzhen, China

A R T I C L E   I N F O

Article history:
Received 5 February 2013
Returned to author for revisions 21 March 2013
Accepted 8 April 2013
Available online 6 May 2013

Keywords:
Adenovirus-encoded small RNA
RNA sequencing
Small RNA target prediction
Suppression of cellular gene expression
Viral defense mechanism

A B S T R A C T

Using deep RNA sequencing, we have studied the expression of adenovirus-encoded small RNAs at different times after infection. Nineteen small RNAs which comprised more than 1% of the total pool of small RNAs at least one time point were identified. These small RNAs were between 25 and 35 nucleotides long and mapped in the region of the VA RNA I and RNA II genes. However, the overlap was incomplete and some contained a few extra nucleotides at the 3’ end. This finding together with the observation that some of the small RNAs were detected before VA RNA expression had started might indicate that they are derived from other precursors than VA RNA I and II. Interestingly, the small RNAs displayed different expression profiles during the course of the infection suggesting that they have different functions. An effort was made to identify their mRNA targets by using computer prediction and deep cDNA sequencing. The most significant targets for the earliest small RNAs were genes involved in signaling pathways.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Human adenoviruses are non-enveloped icosahedral viruses with linear double stranded DNA genomes of 30–38 kb. The genome is transcribed from both strands and it is organized into several transcription units named mainly according to when they are expressed during the virus life cycle (Davison et al., 2003). Five early transcription units encode the E1A, E1B, E2, E3 and E4 proteins, two delayed early units encode the IVa2 and pIX proteins and there is one major late transcription unit (MLTU). The major functions of the early gene products are to force the host cell to enter the S phase (Bayley and Mymryk, 1994; Cobrinik, 1996) in order to provide optimal conditions for viral DNA replication and for suppression of the host antiviral response (Wold et al., 1994).

The major late gene products are the viral structural proteins which package the viral DNA into new virus particles. Group C adenoviruses also encode two small RNAs, called virus-associated (VA) RNA I and VA RNA II (Akusjarvi et al., 1980; Mathews, 1975; Pettersson and Philipson, 1975; Price and Penman, 1972; Reich et al., 1986). They are non-coding RNAs and transcribed by RNA polymerase III (Ma and Mathews, 1996; Weinmann et al., 1974).

Both VA RNAs are about 160 nucleotides long and GC rich. Expression of the VA RNAs begins during the early phase of infection and increases rapidly to a high level during the late phase (Reich et al., 1966; Soderlund et al., 1976). VA RNA I is the most abundant species and accounts for 95% of the total VA RNA pool. Inactivation of VA RNA I in adenovirus type 5 results in a 10–20 fold decrease in virus production, whereas deletion of VA RNA II alone has little impact on virus replication (Thimmappaya et al., 1982). Thus, the functional significance of VA RNA I is well documented whereas little is known about the function of VA RNA II. The primary function of VA RNA I appears to be to block the activity of RNA-dependent protein kinase (PKA), a double-stranded RNA activated inhibitor of translation (Akusjarvi et al., 1987; Ghadge et al., 1991; Maran and Mathews, 1988; O’Malley et al., 1986). PKA is an interferon-inducible serine-threonine protein kinase which phosphorylates the eukaryotic initiation factor 2 (eIF-2) resulting in the inhibition of cellular protein synthesis (Kitajewski et al., 1986; Levin et al., 1980; Petryslyn et al., 1983). It provides a potent anti-viral defense mechanism. VA RNA II also stabilizes ribosome-associated viral mRNAs resulting in enhanced levels of viral protein synthesis (O’Malley et al., 1989). In addition, VA RNA II binds efficiently to Exportin 5 (Exp5), thus interfering with the nuclear export of the cellular RNAi and miRNA precursors and Dicer processing (Anderson et al., 2005; Lu and Cullen, 2004). Finally, large amounts of VA RNA-derived small RNAs associate with RNA-induced silencing (RISC) complexes (Xu et al., 2009).

The VA RNAs fold into a stable stem-loop structure which is similar to that of pre-miRNAs and are subjected to a similar processing (Ma and Mathews, 1993, 1996; Pe’ery et al., 1993). The structure can be divided into three regions: the panhandle apical stem, a more structured central domain, and the terminal stem. The termini of both VA RNA I and II are heterogeneous, due to variable transcription initiation and termination sites. VA RNA I and II are non-coding RNAs and transcribed by RNA polymerase III (Ma and Mathews, 1996; Weinmann et al., 1974).
has two initiations: one is at nt 10607 [VA RNAI(A)] and another is at nt 10610 VA RNAI(G)] (Xu et al., 2009). The latter is the major species and accounts for 75% of the total pool of VA RNAI. It has been demonstrated in vitro and in vivo that 2–5% of the VA RNAs are cleaved by Dicer in their terminal stem generating small viral RNAs named sRNA or mRNAs (Andersson et al., 2005; Aparicio et al., 2006). The most prominent mRNAs are made from the 3’ ends of VA RNAI and VA RNAII and named after their starting positions (Xu et al., 2007). Most abundant are mivaRI-137 (nt 10746) and -138 (nt 10747) processed from VA RNAI and mivaRII-138 (nt 11003) from VA RNAII. These small RNAs are efficiently incorporated into the RISC complex. It has been shown that 80% of RISC-bound miRNA at the late phase of infection is mivaRNA (Andersson et al., 2005; Mathews, 1995). Thus the RISC function seems to be under the control of Ad mivaRNA. The mivaRNAs are also produced from the 5’ end of VARNI albeit at a very low level and the incorporation into the RISC complex is inefficient (Xu et al., 2007).

During the last decade, increasing numbers of small RNAs have been identified and characterized and it has become evident that the small RNAs are critical regulators of gene function (Bartel, 2004; Seto et al., 2007; Zaratiegui et al., 2007). There are three main categories: short interfering RNAs (siRNAs which are ~21 nt in length), microRNAs (miRNAs, ~22 nt in length) and PIWI-interacting RNAs (piRNAs, ~24–32 nt in length) (Farazi et al., 2008). siRNAs and miRNAs are present in a broad range of eukaryotic species and are characterized by the double-stranded nature of their precursors (Ruby et al., 2006). In contrast, piRNAs are primarily found in mammals and exert their function in the germ line (Kim, 2006). They are processed from single-stranded RNA precursors and transcribed mainly from repetitive elements known as piRNA clusters by a Dicer-independent pathway. miRNAs act as key post-transcriptional regulators of gene expression, through modulating the translational efficiency and/or the stability of target mRNAs. In addition to eukaryotes, several virus families have been shown to encode small RNAs. Five Epstein-Barr virus–encoded premiRNAs were identified in infected cells (Pfeffer et al., 2004). Subsequent studies identified a total of 25 pre-miRNAs, which produce at least 44 mature miRNA species (Cai et al., 2006; Grundhoff et al., 2006; Zhu et al., 2009). Furthermore, many herpes viruses have been shown to encode miRNAs (Cai et al., 2005; Grundhoff et al., 2006; Pfeffer et al., 2005). These are derived from structured single stranded transcripts and thus defined as virus-encoded miRNAs. They are believed to regulate the expression of specific viral and cellular mRNAs. In addition, virus-encoded miRNAs or miRNA-like RNAs have been identified in many viruses including members of the polyomavirus family, Dengue virus, vesicular stomatitis virus, baculovirus, hepatitis C virus, West Nile virus, and human immunodeficiency virus (Aparicio et al., 2006; Cantalupo et al., 2005; Hussain et al., 2008; Parameswaran et al., 2010; Schopman et al., 2012; Seo et al., 2009; Singh et al., 2010; Sullivan et al., 2005).

Deep sequencing technologies combined with bioinformatic strategies have revolutionized the identification of rare small RNAs. In this study, we have studied the expression of adenovirus-encoded small RNAs at different times after infection using the SOLid deep sequencing technology.

Materials and methods

Cell culture and adenovirus infection

Human primary lung fibroblast cells (IMR-90) (American Type Culture Collection) were cultured in Eagle’s minimum essential medium (MEM) with Earle’s salt and GlutaMAX™ supplemented with 10% fetal bovine serum, 100 U/ml penicillin, 100 μg/ml streptomycin, 0.1 mM nonessential amino acids, 1.5 g/l sodium bicarbonate and 1.0 mM sodium pyruvate. After reaching confluence, the cells were cultured for two more days in order to synchronize the cells. Over 95% of the cells were in the G0/G1 phase as indicated by FACS analysis (Zhao et al., 2007). Synchronized monolayer cells were mock-infected or infected with Ad2 at a multiplicity of 100 fluorescence-forming units (FFU) per cell (Philipson, 1961) in serum-free medium. After 1 h adsorption at 37 °C, the medium was replaced with MEM containing 10% FBS and incubated at 37 °C. Infected cells were collected at 6, 12, 24, and 36 h post infection (hpi). Mock-infected cells were collected at 6 hpi.

RNA extraction, cDNA library preparation, and sequencing

Total RNA from adenovirus or mock-infected IMR-90 cells was extracted using TRIZOL Reagent (Invitrogen). The quality of the input RNA in the 10–40 nt size range was controlled using the Agilent 6000 Pico chip on a Bioanalyzer (Agilent Technologies). Only RNA values above 7 were accepted. Enrichment of the small RNA was done by using the Invitrogen PureLink miRNA Isolation Kit according to the manufacturer’s protocol. Small RNA libraries were constructed using the SOLiD Total RNA-Seq Kit (Rev B, Life Technologies). Strand specific adapters were hybridized to the RNA before reverse transcription. The cDNA was then size-selected on a 10% TBE- Urea gel (Life Technologies) and the libraries were constructed after amplification (15 cycles). Emulsion PCR was performed using the EZ Bead System (Life Technologies) and the small RNA libraries were then sequenced on the SOLiD 5500xl system (35 bp read length, Life Technologies).

Data collection and mapping of reads

The reads were generally longer than the investigated RNAs because of possible remains of the adapter sequence. We used a strict algorithm to find the maximal overlap between the 3′-adapter and the end of the read. The trimmed reads of each library were then aligned to the whole genome sequence of human adenovirus type 2 (http://www.ncbi.nlm.nih.gov/nuccore/ NC_001405.1) using the Map reads module of LifeScope 2.5 software (http://www.lifetechologies.com/lifescope). Here we allowed two mismatches.

Identification of adenovirus-encoded small RNA

The read alignments were rigorously checked for quality. First we removed the reads mapping to the Coding DNA Sequence (CDS) regions. After sorting the reads by their expression number, small RNAs were collected if they have high predominance and change with time. We found that nearly all reads were exclusively located within the VA RNA region.

Prediction of adenovirus-encoded small RNA targets and function enrichment analysis

Potential cellular gene targets of adenovirus-encoded small RNAs were predicted by using RNAhybrid (http://bibiserv.techfak.uni-bielefeld.de/ahybrid/) using the reference human gene set (hg19). The target gene list was then cleaned by filtering through the Support Vector Machines (SVM) model.

The web based software DAVID (Huang da et al., 2009a, 2009b) was used to perform a gene ontology enrichment analysis of the differentially expressed genes. The annotated genes from the RefSeq data set were used as background. This analysis searched for biochemical pathways and gene ontologies that were...
overrepresented in our data, compared to a random sample of equal size from the background list.

Results

1. Identification of adenovirus-encoded RNAs by cDNA sequencing

Previously, we have studied the transcriptional profiles in adenovirus infected IMR-90 cells. Three cDNA libraries, prepared from mock, Ad2 infected cells at 12 (Ad2-12hpi) and 24 (Ad2-24hpi) hours after infection were sequenced and generated 50–54 million 76 bp long sequence reads per sample. Deregulation of cellular mRNA expression was the main focus in our previous publication (Zhao et al., 2012). This study also generated a comprehensive data set of adenovirus gene expression. Among detected sequence reads, 4.4% at 12 hpi and 45% at 24 hpi could be mapped to the adenovirus genome. At 12 hpi, only adenovirus early genes E1A, E1B, E3, E4 and E2A were expressed. Following the progress of the infection, adenovirus gene expression was dramatically changed and transcripts from the major late transcription unit (MLTU) became very abundant. An interesting finding was that RNA transcripts spanning from 10700 to 10800 were the most abundant and reached more than 180,000 reads. This region covers the VA RNA genes and we suspected that the peak could contain a pool of small RNAs. To test the hypothesis, we performed small RNA sequencing.

2. Detection of adenovirus-encoded RNAs in the 10–35 nt range by RNA sequencing

The small RNA pool in adenovirus infected IMR-90 cells was sequenced by SOLiD sequencing technology. Hereby, the expression of miRNAs and other non-coding RNAs of both cellular and viral origin could be discovered and profiled in an unbiased way. In the present study we focused on the expression of adenovirus-encoded small RNAs. Sequencing was performed on five RNA samples extracted from mock and adenovirus infected cells at 6, 12, 24 and 36 hpi. As shown in our previous study, these time points represent different stages of the infectious cycle, i.e. before any adenoviral gene expression, immediate early gene (E1a) expression, adenoviral DNA replication and viral late gene expression, respectively (Zhao et al., 2007). Thus, we could correlate the expression of small RNAs with the progression of the infection.

The expression of adenovirus-encoded small RNAs increased rapidly after infection. There were 97,408, 844,458, 1,866,214 and 4,566,034 sequence reads that could be mapped to adenovirus genome at 6, 12, 24 and 36 hpi, respectively (Table 1). It is noteworthy that the most dramatic increase, approximately 9-fold, occurred between 6 and 12 hpi. Then, there was a relative steady increase of about 2-fold from 12 to 24 and from 24 to 36 hpi. More than 98% of the adenovirus-encoded small RNAs were mapped within the VA RNAI and VA RNAII region at 6 hpi, which was then slightly decreased to 89.6% at 36 hpi. The majority of small RNAs was mapped within the VA RNAI region. Using 1000 sequence reads as a cut-off level, 207 small RNAs were identified (see Table in supplementary material). Nearly all of them (205 out of 207) were mapped to the VA RNAI and VA RNAII regions as shown in Fig. 1. Included in this figure is the expression of unfractionated adenovirus RNA at 24 hpi (blue curve) detected by cDNA sequencing of total RNA. The locations of the previously identified adenovirus specific mivaRNAs are also shown (Andersson et al., 2005; Sano et al., 2006; Xu et al., 2007). The majority of the adenovirus-encoded small RNAs co-localized with the highly expressed VA RNAI. There were four major groups of small RNAs. All small RNAs in each group overlap but vary in length at the 5’ or 3’ end or both. The small RNAs in the first groups comprise two major clusters which differ from each other by three nt at the 5’ end, corresponding to the transcription initiation sites for VA RNAI (G) and VA RNAI (A). Surprisingly, more than 60% of detected small RNAs were mapped to the region spanning from the central domain to the 3’ end of VA RNAI (nt 10702 to 10772). These RNAs were very heterogeneous at both the 5’ and the 3’ ends. Less than 30% of the small RNAs mapped to the VA RNAII region. One cluster contained RNAs with identical 5’ ends matching the initiation site of VA RNAII. Another cluster which mapped to the 3’ end of VA RNAII was heterogeneous at both the 5’ and the 3’ ends spanning a region from nt 10968 to 11030.

The length of the small RNAs ranged from 25 to 35 nt (as shown in Fig. 2 and Table in supplementary data). The most significant group was 35 nt long. However, it is possible that this group included longer RNAs, since the sequence reads were only 35 nt long. Therefore, we were unable to determine the 3’ end of these RNAs unambiguously. A surprising finding was that several of small RNAs have 3 ends which extend beyond the 3’ ends of VARNAI and II.

3. Expression of distinct sets of small RNA at different stages of the infectious cycle

Fig. 3 shows the alignment of small RNAs expressed at different times after infection to the VA RNA gene region. At 6 hpi, when E1A expression could barely be detected, a significant amount of small RNAs was identified mapping to the 5’ end of VA RNAI (with more than 53,000 reads). To reveal the expression difference, the relative abundances of the adenovirus-encoded small RNAs were calculated for each small RNA species in relation to the total pool of all detected adenovirus-encoded small RNAs at each time point. Small RNAs with more than 1% value were considered as significantly expressed and included in Table 2. At 6 hpi, the most abundant small RNA (54.6%) maps to 5’ end of VA RNAI, from nt 10607 to 10631 (VA sRNA10607–10631). Two small RNAs (VA sRNA10866–10897 and VA sRNA10866–10897) which mapped to the 5’ end of VA RNAI represented the second most abundant small RNAs (7.9%). Remaining RNAs mapped mostly to the 3’ end of VA RNAI and VARNAl. At 12 hpi, the overall expression of small RNAs increased 9-fold. Expression of VA sRNA10607–10631 was increased 10-fold and accounted for 63.5% of adenovirus-encoded small RNAs at this time. Although the small RNAs that mapped to the 3’ end of VA RNAI was increased 3-fold, they represented less than 1% of the small RNAs at this time point. Small RNAs mapping to VA RNAII,
especially to the 3’ end, were also increased (about 10% in total). About 5.7% of the small RNAs mapped to the 5’ end of VA RNAII. The change of the expression profile at 24 hpi was dramatic. The fraction of VA sRNA \[10607-10631\] decreased to 28.4%. Instead, small RNAs that mapped to the 3’ side of the central domain of VA RNAI from 10702 to 10736 (or longer) increased at 24 hpi (12.7% in total) and became the most abundant class at 36 hpi (49.2% in total). In addition, a group of small RNAs that mapped to the 3’ end of VA RNAII reached its highest level (> 16.6%) at 24 hpi, but was decreased at 36 hpi.

Prediction of small RNA targets

The possible functions of the small RNAs were studied by prediction of their potential targets using RNAhybrid, which is a tool for finding the minimum free energy of hybrids between long mRNAs and small RNAs. The identified potential target mRNAs were then subjected to functional enrichment analysis by using gene functional annotation resources (DAVID). Three highly expressed small RNAs, Ad VA sRNA \[10607-10731\], Ad VA sRNA \[10702-10736\] and Ad VA sRNA \[11001-11025\], were chosen for this analysis, because of their abundance and their expression kinetics. The results showed that the predicted targets of Ad VA sRNA \[10607-10731\] were involved in various signaling pathways, cell surface structure, and ligand-receptors interaction (Table 3). In contrast, the predicted target genes of Ad VA sRNA \[11001-11025\] were mainly involved in cellular metabolic pathways such as nitrogen metabolism, glycan biosynthesis, and fatty acid biosynthesis. No significant cellular targets were predicted for Ad VA sRNA \[10702-10736\] perhaps because of its length. As a step further, we compared the predicted target genes with genes that were found to be deregulated in our previous study (Zhao et al., 2012). Interestingly our previous study showed that cellular signaling pathways were the most significantly suppressed gene group at 12 hpi. The most significant class of predicted targets of Ad VA sRNA \[10607-10731\] was various protein kinases (Table 4). These kinases participate in several signaling pathways that are involved in cellular stress/immune response and apoptosis, (such as MAP kinases: MAPKAPK2 MAP3K14 and MAP2K4, serine/
threonine kinase: STK17A, and STK40, as well as NUAK2 and LAST2), growth/cell cycle arrest (such as AURKA, PLK3, RPS6KA2, STK10 and SPEG) and cytoskeletal organization, vesicle trafficking and nuclear transport (DCLK2, DMPK, SGK223, MAST4, PAK4 and WNK4). Furthermore, numerous small GTPases (RAB4B, RAB22A, RAB2B, RAB40C, RAB5B, ARF4 and RND3) and their regulators, GTPase-activating proteins (EVI5L, TBC1D2, TBC1D10B, TBC1D23 and TBC1D10A) and guanine nucleotide exchange factors (ARHGEF39, ARHGEF40, PLEKHG4, PLEKHG2 and ARHGEF17) were found among the target genes. It is more noteworthy that most of these genes were down-regulated already at 12 hpi. Taken together the results indicated that the immediately expressed small RNAs target cellular signaling pathways involved in various cellular processes and that the expression profile of these RNAs is dynamic and changes over time.

![Graph](image)

**Fig. 3.** A compilation of adenovirus-encoded small RNAs within the VA RNA region at 4 different time points after infection. Each small RNA was aligned to the VA RNA region (from 10600 to 11000). The expression levels are illustrated by the heights of the peaks. The positions of VA RNAI and II are indicated below each diagram.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Expression of adenovirus-encoded small RNAs with more than 1% of the total pool at different stages of infection.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genome location</strong></td>
<td><strong>Sequence reads</strong></td>
</tr>
<tr>
<td>VA RNAI</td>
<td><strong>Mock</strong></td>
</tr>
<tr>
<td>10607-10631</td>
<td>92</td>
</tr>
<tr>
<td>10702-10731</td>
<td>9</td>
</tr>
<tr>
<td>10702-10733</td>
<td>14</td>
</tr>
<tr>
<td>10702-10734</td>
<td>10</td>
</tr>
<tr>
<td>10702-10735</td>
<td>18</td>
</tr>
<tr>
<td>10702-10736</td>
<td>149</td>
</tr>
<tr>
<td>10723-10757</td>
<td>5</td>
</tr>
<tr>
<td>10730-10754</td>
<td>2</td>
</tr>
<tr>
<td>10747-10771</td>
<td>10</td>
</tr>
<tr>
<td>10747-10772</td>
<td>4</td>
</tr>
<tr>
<td>VA RNAI</td>
<td><strong>10866-10890</strong></td>
</tr>
<tr>
<td>10866-10896</td>
<td>3</td>
</tr>
<tr>
<td>10866-10897</td>
<td>12</td>
</tr>
<tr>
<td>11001-11025</td>
<td>2</td>
</tr>
<tr>
<td>11001-11026</td>
<td>4</td>
</tr>
<tr>
<td>11002-11026</td>
<td>0</td>
</tr>
<tr>
<td>11002-11030</td>
<td>0</td>
</tr>
<tr>
<td>11003-11027</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Functional enrichment of small RNA targets by DAVID analysis.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction target of VA sRNA10607-10631</strong></td>
<td><strong>Predicted target of VA sRNA11001-11025</strong></td>
</tr>
<tr>
<td>Calcium signaling pathway</td>
<td>Nitrogen metabolism</td>
</tr>
<tr>
<td>MAPK signaling</td>
<td>Glycan synthesis</td>
</tr>
<tr>
<td>Focal adhesion</td>
<td>Glycosphingolipid biosynthesis</td>
</tr>
<tr>
<td>Gap junction</td>
<td>Adipocyte signaling pathway</td>
</tr>
<tr>
<td>ECM-receptor interaction</td>
<td>Fatty acid biosynthesis/metabolism</td>
</tr>
<tr>
<td>ErbB signaling pathway</td>
<td>Metabolic pathway</td>
</tr>
<tr>
<td>Insulin signaling pathway</td>
<td>Lysosome</td>
</tr>
<tr>
<td>Long-term depression</td>
<td>ECM-receptor interaction</td>
</tr>
<tr>
<td>GnRH signaling pathway</td>
<td>TGF-beta signaling pathway</td>
</tr>
</tbody>
</table>

Furthermore, the analysis of enrichment using DAVID software revealed several pathways that were significantly associated with the sRNA targets. These pathways include calcium signaling pathway, nitrogen metabolism, MAPK signaling, glycan synthesis, focal adhesion, glycophosphoinositide biosynthesis, gap junction, ECM-receptor interaction, ErbB signaling pathway, adipocytokine signaling pathway, insulin signaling pathway, long-term depression, and GnRH signaling pathway. These findings suggest that the sRNAs may be involved in the regulation of various cellular processes and signaling pathways.
species accounted for a minor fraction of all reads. Surprisingly, remaining reads were spread over the genome and each comprised more than 1% of the total pool of small RNAs at any time point. 200 small RNAs were identified using 1000 reads as the cut off end. The small RNAs identified here were identical to the none of the small RNAs identified here were identical to the previously detected adenovirus-encoded small RNAs (miRNA137, -138 and mivRNA11-138). We did detect small RNAs that overlapped with them, but they were 2-15 nt longer at the 3’ or 5’ end. The small RNAs identified here were longer than most well defined cellular small RNAs, such as miRNA and siRNA. It could be argued that our sequencing procedure failed to detect miRNAs less than 25 nt long. This seems however, unlikely since we did detect expression of more than 300 annotated cellular miRNAs.

A striking finding was that the expression of different adenovirus-encoded small RNAs increased and declined during pathways that are involved in stress/immune response, cell growth/cycle and intracellular transport.

Discussion

The aim of the present study was to perform a comprehensive and unbiased study of small adenovirus-encoded RNAs. More than 200 small RNAs were identified using 1000 reads as the cut off level. Many of them were overlapping and differed because of heterogeneity at both the 3’ and 5’ ends and only 19 of them comprised more than 1% of the total pool of small RNAs at any time point. Between 90 and 98% of the reads mapped in the VA RNA region. Remaining reads were spread over the genome and each species accounted for a minor fraction of all reads. Surprisingly, none of the small RNAs identified here were identical to the previously detected adenovirus-encoded small RNAs (miRNA137, -138 and mivRNA11-138). We did detect small RNAs that overlapped with them, but they were 2-15 nt longer at either the 3’ or 5’ end. The small RNAs identified here were longer than most well defined cellular small RNAs, such as miRNA and siRNA. It could be argued that our sequencing procedure failed to detect miRNAs less than 25 nt long. This seems however, unlikely since we did detect expression of more than 300 annotated cellular miRNAs. For instance, microRNA 21 was covered by more than 10,000 reads. One possibility is that adenovirus utilizes a unique mechanism for generating small RNAs. A noteworthy difference is that adenovirus VA RNAI and II are synthesized by pol III.

A striking finding was that the expression of different adenovirus-encoded small RNAs increased and declined during
different phases of the infection. At 6 hpi, VA sRNA10607–10631 was the most highly expressed species. It may represent the earliest adenovirus gene expression, since no expression of other adenovirus genes has been detected at this time point in IMR-90 cells (Zhao et al., 2007). It is tempting to speculate that this RNA plays an important role during the very early phase of infection. To gain support for this hypothesis, we performed a target predication and the results suggested that mRNAs encoding proteins in signaling pathways are the targets. In line with this we have shown before by deep cDNA sequencing (Zhao et al., 2012) that the most significantly suppressed gene group at 12 hpi was implicated in cell signaling pathways. The predicted targets included various kinases, small GTPases and their regulators, guanine nucleotide exchange factors and GTPase-activating proteins. Among them p38 MAPK is noteworthy since MAPKs mediate responses to a diverse array of stimuli, including virus infection. The main downstream target genes regulate proliferation, cell survival and apoptosis, cell growth/cell cycle arrest and cytoskeleton. In fact, immediate activation of several cellular signaling pathways, such as PKA signaling pathways were down-regulated as revealed by cDNA sequencing. The adenovirus-encoded small RNAs may thus contribute to the virus. On the other hand, TGF signaling pathway was one of the targets, which agrees with our previous results.

Another important question concerns the substrates from which the viral small RNAs are excised. Are they cleaved from VA RNAI and II? Although we lack direct evidence some observations suggest that they are excised from other precursors than the VARNAs. First of all, expression of some small RNAs starts very early, before the VA RNAs are detected. There is, moreover, no correlation between expression levels of the small RNAs and expression levels of VA RNAI and II. Secondly, the sequences of the most highly expressed small RNAs did not completely match that of VA RNAI. It has been shown that VA RNAI has two initiations sites, one at nt 10607 [VA RNA(A)] and another at nt 10610 [VA RNA(G)] (Xu et al., 2009). The latter is the major species and accounts for 75% of total VA RNAI. Our results showed that VA sRNA10607–10631 was the major species and accounted for 99% of small RNAs that mapped to the 5’ end of VA RNAI during the early phase of infection. Furthermore, several small RNAs that mapped to the 3’ end of VA RNA extend beyond the 3’ ends of the VA RNAs. Lastly, from a structural point of view it would not be expected that a small RNA is generated from the central domain of a precursor. Therefore it seems unlikely that sRNA10702–10736 is excised from VA RNAI. Thus, we propose that adenovirus-encoded small RNAs are produced from a different set of VA RNAs which are subjected to very efficient cleavage soon after they are produced.

Cellular gene expression is regulated by complex networks. The adenovirus-encoded small RNAs are likely to be a part of these networks. The rapid expression of VA sRNA10607–10631, suggested an important function in the early phase of the infection. Our previous microarray studies showed that the regulation of host gene expression in response to the incoming virus is very rapid. Eighteen cellular genes were up-regulated already at 1 hpi (Granberg et al., 2006). Eight out of twelve annotated genes encoded transcription factors that are implicated in cell growth arrest (ATF3, ATF4, KLF4, KLF6, ELK3 and GABPB2), as well as antiviral immune response and apoptosis (CEBPB and NRA1). In addition, the signal regulators RGS4 and ARHE were also up-regulated. RGS4 is a small GTPase activating protein and ARHE is a member of the small GTPase protein superfamily. Regulation of cellular gene expression during this period must be independent of adenoviral gene expression and is likely to be triggered by the attachment of the virus to cell surface receptors, the entry process of the virus, its intracellular transport along the microtubules or the possible intracellular recognition of viral nucleic acids by Toll-like receptors. MAPK cascades are key components of the signaling networks that sense cellular exposure to environmental stress. Most of the up-regulated genes are either involved in this pathway or are downstream targets. However, activation of these immediate response genes is transient and their expression had decreased at 6 hpi. An important question is how the earliest host cell anti-viral response is repressed? Although it is well established that E1A is the first viral gene to be transcribed, they are first detected at 12 hpi in IMR-90 cell used in our studies. E1A cannot thus be responsible for the suppression of host cell immediate anti-viral response. We show here that some small RNAs, particularly VA sRNA10607–10631 are expressed already at 6 hpi. Furthermore, we show here that Ad VA sRNA10607–10631 are likely to target genes involved in diverse signaling pathways, including the MAPK pathway. Further support for this notion, was the finding that many predicted targets, especially those involved in signaling pathways were down-regulated as revealed by cDNA sequencing. The adenovirus-encoded small RNAs may thus constitute the front-line defense and be crucial for the survival of the virus.

Acknowledgments

Sequencing was performed by the Uppsala genome center, Science For Life Laboratory in Uppsala. We thank Linnea Nyberg for excellent sequencing, Christian Tellgren-Rothfor for help with the data analysis. This work was supported by the Kjell and Märta Beijer Foundation.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2013.04.006.

References


