Confined aquifers as viral reservoirs

Renee J. Smith,1* Thomas C. Jeffries,2 Ben Roudnew,1 Justin R. Seymour,2 Alison J. Fitch,1 Keryn L. Simons,1 Peter G. Speck,1 Kelly Newton,1 Melissa H. Brown1 and James G. Mitchell1
1School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia.
2Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, Sydney, NSW 2007, Australia.

Summary

Knowledge about viral diversity and abundance in deep groundwater reserves is limited. We found that the viral community inhabiting a deep confined aquifer in South Australia was more similar to reclaimed water communities than to the viral communities in the overlying unconfined aquifer community. This similarity was driven by high relative occurrence of the single-stranded DNA viral groups Circoviridae, Gemiviridae and Microviridae, which include many known plant and animal pathogens. These groups were present in a 1500-year-old water situated 80 m below the surface, which suggests the potential for long-term survival and spread of potentially pathogenic viruses in deep, confined groundwater. Obtaining a broader understanding of potentially pathogenic viral communities within aquifers is particularly important given the ability of viruses to spread within groundwater ecosystems.

Introduction

Confined aquifers typically lie deep below the surface and are permanently, or semi-permanently, separated from other groundwater by low permeability geological formations, which provide barriers to flow (Hamblin and Christiansen, 2004; Borchardt et al., 2007). These barriers are thought to protect the underlying groundwater from the overlying environment, and thus prevent the spread of contaminants into the freshwater reserves (Nolan et al., 1997). However, vertical fractures can lead to the formation of pathways for water movement, allowing for the introduction of surface contaminants, including microbial communities (Eaton et al., 2007). Among the microbial communities present, viruses have substantial potential for spread into deep aquifers due to their small, 27–75 nm, size (Borchardt et al., 2007).

Recent studies have shown that there are large reservoirs of viral cells present in deep aquifer ecosystems (Kyle et al., 2008; Roudnew et al., 2012; 2013). These studies have shown that viral abundances can be in the order of 10^4–10^7 cells ml^-1. However, the identity of these viral groups remains relatively unknown. Previous works in classifying viral groups present in deep aquifers have predominantly focused on viruses that are known to infect humans (Carducci et al., 2003; Borchardt et al., 2007), but a non-targeted approach allowing for a complete picture of the viral communities present in deep aquifers is currently lacking.

A recent metagenomic study of an aquifer system revealed a relatively high proportion (9%) of viral sequences (Smith et al., 2012) when compared with other aquatic environments, where the proportion is typically only 0.1–1% (Edwards and Rohwer, 2005; Williamson et al., 2008; Jeffries et al., 2011). Therefore, we sought to construct a viral community profile from the viral sequences within a microbial metagenome, those caught on a 0.22 μm filter, from unconfined and confined aquifer ecosystems. These data were compared to metagenomes from a number of other marine and freshwater environments.

Results and discussion

Groundwater samples were collected from the confined and unconfined Ashbourne aquifer system, in South Australia (35°18′S 138°46′E) in June 2010 (Appendix S1). The unconfined aquifer is exposed to overlying input, while the confined aquifer lies at 40–100 m, below a 15 m thick confining layer, and has been isolated from external input for approximately 1500 years (Banks et al., 2006). Separate recharge processes have led to distinct water sources that differ between the confined and unconfined aquifers (Banks et al., 2006; Smith et al., 2012). Metagenomes were sequenced using the GS-FLX pyrosequencing platform using Titanium reagents (Roche). The resulting 409 743 and 64 506 sequences from the confined and unconfined aquifers, respectively, were compared with the National Center for Biotechnology Information (NCBI) non-redundant database in the Community Cyberinfrastructure for Advanced Microbial
Ecology Research and Analysis (CAMERA) pipeline (Seshadri et al., 2007). BLASTX and an \( E \times 10^{-5} \) was used to identify hits.

A total of 14,266 and 1,003 assembled contigs from the confined and unconfined aquifers, respectively, had hits to viral sequences in the NCBI non-redundant database. Of these hits, a high percentage of viral sequences within the unconfined aquifer metagenome were unclassified, accounting for 53% of the total viral hits, while only 1% of sequences within the confined aquifer metagenome were unclassified. Of the classified sequences, 1% and 0% were double-stranded DNA (dsDNA) viruses, and 72% and 47% were single-stranded DNA (ssDNA) viruses (Table S1) in the confined and unconfined aquifers respectively. Previous viral metagenomic studies have shown that dsDNA viruses normally dominate over ssDNA viruses (Bench et al., 2007; Correa et al., 2013; Fancello et al., 2013). Their lack of dominance in this study suggests that the aquifer matrix may be excluding the larger dsDNA viruses.

The large proportion of ssDNA viruses were potentially related to the use of multiple displacement amplification (MDA) during sample preparation (Appendix S1), which can preferentially amplify ssDNA viruses (Kim and Bae, 2011). However, in our study, we have analysed contigs that can assist in reducing the effects of artefacts caused by MDA (Lasken and Stockwell, 2007; Rosario et al., 2009). Furthermore, the relative abundances of viral sequences were not used to infer relative abundances of the different DNA viruses. Instead, this study has focused on the overall community profile of viruses present within groundwater ecosystems.

To determine whether groundwater virus communities have intrinsic characteristics, the viral sequences from the confined and unconfined aquifer metagenomes were compared with metagenomes from a variety of other aquatic environments (Table 1) using a normalized Goodall’s Similarity Index (Goodall, 1964; 1966) in the MEtaGenome ANalyzer (MEGAN) (Huson et al., 2007). Despite geographical proximity, the confined aquifer viral consortia did not resemble those of the unconfined aquifer, and were instead most similar to the viral sequences in the metagenome from a reclaimed water sample, the reusable end-product of wastewater treatment, in Florida (Fig. 1) (Rosario et al., 2009; Roudnew et al., 2012; Smith et al., 2012). This similarity between the confined aquifer and the reclaimed water source is likely driven by exposure to similar extensive filtration within their environments, and suggests that filtering within deep groundwater ecosystems can alter microbial community composition.

The lack of similarity between the viruses within the confined and unconfined aquifers contradicts the patterns in bacterial taxonomy recently observed at the same site in South Australia, which showed that the confined aquifer total microbial metagenome, predominantly bacteria, was taxonomically more similar to that of the overlying unconfined aquifer than to any other environment (Smith et al., 2012). However, rationale for why the discrepancy between the patterns in bacterial and viral taxonomy is unknown and would require further investigation. Furthermore, the lack of similarity between the confined and unconfined aquifer viral communities suggests that viruses were not introduced into the confined aquifer from the overlying unconfined aquifer, indicating the long-term survival of viruses in groundwater. This, combined with estimates that the confined aquifer water is 1500 years old, suggests the long-term survival of viruses in groundwater.

To identify the taxa contributing to the similarity between the reclaimed water viruses and the confined aquifer viruses, community profiles were generated in MEGAN (Huson et al., 2007). The community profile indicated that the main taxa contributing to the similarity between the two metagenomes were ssDNA viruses (Fig. 2), accounting for 72% and 21% of the viruses in the confined aquifer and reclaimed water respectively (Fig. 2). Within the ssDNA viruses, members of the Microviridae dominated, accounting for 63% and 64% in the confined aquifer and reclaimed water source respectively. Within this group, DNA sequences were more similar to the members of the Gokushovirinae genus. The members of Gokushovirinae are known to infect obligate intracellular parasitic bacteria, including Bdellovirinae and Chlamydia (Cherwa and Fane, 2011), the latter of which has previously been identified in the

Table 1. Summary of publicly available metagenomes used in this study.

<table>
<thead>
<tr>
<th>Database</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG-RAST</td>
<td>Unconfined aquifer</td>
<td>Smith and colleagues (2012)</td>
</tr>
<tr>
<td>MG-RAST</td>
<td>Confined aquifer</td>
<td>Smith and colleagues (2012)</td>
</tr>
<tr>
<td>MG-RAST</td>
<td>Botany Bay</td>
<td>Burke and colleagues (2011)</td>
</tr>
<tr>
<td>CAMERA</td>
<td>Metagenomic analysis of viruses in reclaimed water</td>
<td>Rosario and colleagues (2009)</td>
</tr>
<tr>
<td>CAMERA</td>
<td>Chesapeake Bay viroplankton metagenome</td>
<td>Bench and colleagues (2007)</td>
</tr>
<tr>
<td>CAMERA</td>
<td>Viral metagenome from the freshwater Lake Limnopolar</td>
<td>López-Bueno and colleagues (2009)</td>
</tr>
<tr>
<td>CAMERA</td>
<td>Viral metagenomes from terrestrial hot springs</td>
<td>Schoentelf and colleagues (2008)</td>
</tr>
<tr>
<td>CAMERA</td>
<td>Viral stromatolite metagenome</td>
<td>Desnues and colleagues (2008)</td>
</tr>
</tbody>
</table>

© 2013 Commonwealth of Australia, Environmental Microbiology Reports
These viruses may play a role in determining the composition of the bacterial community, which in turn may influence chemical cycling within the low-nutrient confined aquifer. Contigs matching individual Gokushovirinae sequences from the confined aquifer and reclaimed water source were aligned using CLUSTALW, with low sequence similarities observed. Due to incomplete sequence coverage, however, it is unclear as to whether there are novel Gokushovirinae members within the confined aquifer, or alternatively that different regions of the genomes were sequenced.

In the confined aquifer, members of the Circoviridae and Geminiviridae families accounted for 17% and 3%, respectively, while in the reclaimed water sample, these viral groups accounted for 6% and 7% respectively. Unclassified ssDNA viruses, predominantly Circovirus-like genomes, comprised 15% and 24% of the ssDNA viruses in the confined aquifer and reclaimed water, respectively. Nanoviridae were only found in the confined aquifer sample, accounting for 2% of ssDNA viruses overall (Figs 2 and 3). Circoviridae, Geminiviridae, Microviridae and Nanoviridae are all small viruses, with diameters of 17–30 nm (Storey et al., 1989; Gibbs and Weiller, 1999; Gutierrez et al., 2004). Thus, the dominance of these ssDNA viruses is consistent with the observations that small viruses have the greatest potential for transport through aquifers (Yates et al., 2014), and provides evidence for how filtration processes within deep aquifers alter viral communities.
Circoviridae, Geminiviridae and Nanoviridae all contain known plant or vertebrate pathogens (Gibbs and Weiller, 1999; Gutierrez et al., 2004). In particular, Circoviridae have been characterized from the tissues of birds, mammals, fish, insects, plants and algal cells (Delwart and Li, 2012), and have previously been found in other environmental metagenomic studies (Rosario et al., 2012). Unlike shallow unconfined aquifers, which are known to contain small invertebrates (Eberhard et al., 2005), confined aquifers are not known to contain plants or animals. This suggests that the ssDNA viruses identified in this study were introduced exogenously. Therefore, the identification of such viruses in this study from a 1500-year-old confined aquifer (Banks et al., 2006) suggests that the potential exists for long-term survival and spread of small viruses in groundwater. To our knowledge, this is the first metagenomic study to characterize viral groups present in a confined aquifer ecosystem, providing a broader understanding of viral communities within aquifers.

Acknowledgements

The authors gratefully acknowledge the funding provided by the Australian Research Council. R. J. Smith is the recipient of a Flinders University Research Scholarship (FURS).

Fig. 2. Community profile of unconfined aquifer, confined aquifer and reclaimed water metagenomes matching the NCBI non-redundant database in CAMERA. Non-redundant metagenomic sequences were assembled and identified using the BLASTX algorithm and E < 1 × 10^{-5} against the NCBI non-redundant database using CAMERA (Seshadri et al., 2007). Normalized abundances were then used to generate a community profile in MEGAN (Huson et al., 2007). Phyla expanded to family level where available.

Fig. 3. ssDNA viruses % relative abundance in the confined aquifer and reclaimed water samples identified by BLASTX to the viral proteins database in CAMERA (Seshadri et al., 2007).
References


Seshadri, R., Kravitz, S.A., Smarr, L., Gilna, P., and Frazier,


Supporting information
Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

Table S1. Relative proportion of matches to the viral proteins database taxonomical hierarchy.

Appendix S1. Supplementary methods.