An Efficient Strategy of Screening for Pathogens in Wild-Caught Ticks and Mosquitoes by Reusing Small RNA Deep Sequencing Data

Lu Zhuang1*, Zhiyi Zhang1*, Xiaoping An1, Hang Fan1, Maijuan Ma1, Benjamin D. Anderson2, Jiafu Jiang1, Wei Liu1, Wuchun Cao1*, Yigang Tong1*

1 State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Fengtai District, Beijing, P. R. China, 2 College of Public Health and Health Professions, Emerging Pathogens Institute, University of Florida, Gainesville, Florida, United States of America

Abstract

This paper explored our hypothesis that sRNA (18–30 bp) deep sequencing technique can be used as an efficient strategy to identify microorganisms other than viruses, such as prokaryotic and eukaryotic pathogens. In the study, the clean reads derived from the sRNA deep sequencing data of wild-caught ticks and mosquitoes were compared against the NCBI nucleotide collection (non-redundant nt database) using Blastn. The blast results were then analyzed with in-house Python scripts. An empirical formula was proposed to identify the putative pathogens. Results showed that not only viruses but also prokaryotic and eukaryotic species of interest can be screened out and were subsequently confirmed with experiments. Specially, a novel Rickettsia spp. was indicated to exist in Haemaphysalis longicornis ticks collected in Beijing. Our study demonstrated the reuse of sRNA deep sequencing data would have the potential to trace the origin of pathogens or discover novel agents of emerging/re-emerging infectious diseases.

Citation: Zhuang L, Zhang Z, An X, Fan H, Ma M, et al. (2014) An Efficient Strategy of Screening for Pathogens in Wild-Caught Ticks and Mosquitoes by Reusing Small RNA Deep Sequencing Data. PLoS ONE 9(3): e90831. doi:10.1371/journal.pone.0090831

Editor: Amit Kapoor, Columbia University, United States of America

Received October 9, 2013; Accepted February 4, 2014; Published March 11, 2014

Copyright: © 2014 Zhuang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by the Natural Science Foundation of China (81222037, 81290344, 81130086, 81072250). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: caowc@nic.bmi.ac.cn (WC); tong.yigang@gmail.com (YT)

These authors contributed equally to this work.

Introduction

Recently, researchers have used deep sequencing of small RNA to identify microRNAs that function in the transcriptional and post-transcriptional regulation of gene expression in plants and animals [1] and also to be an effective methods of virus discovery in plants and invertebrates [2–4]. This approach utilizes the mechanism in which small interfering RNAs are generated during the viral immunity process [5]. RNA silencing, or interference, as a form of viral immunity, begins with the recognition of a viral double-stranded or structured RNA by the Dicer nuclease family [6–8], which results in short interfering RNA (21–26 nt).

Due to the genomic diversity of differing pathogens, the current metagenomic approaches for microbial analysis require specific protocols to detect DNA viruses, RNA viruses, and other cellular pathogens [9,10]. Because of this, sample processing is often labor intensive and costly. Since small RNA fractions could contain RNA metabolites derived from all RNA species, such as rRNAs, tRNAs, mRNA, snRNA, snoRNA [11], we hypothesize that it would be possible to use deep sequencing of sRNA as a universal strategy to identify multiple types of microorganisms other than viruses, including prokaryotic and eukaryotic pathogens. Therefore, in this study we demonstrate the use of sRNA deep sequencing method as a universal way to screen for multiple groups of pathogens, including viruses, bacteria, and eukaryotes, in wild-caught mosquitoes and ticks.

Methods

Collection of ticks and mosquitoes

Eight adult Haemaphysalis longicornis (H. longicornis) ticks were collected by dragging over the vegetation layer in the suburbs of Beijing, north of China (Fig. 1), pooled into one sample, and named XCP. Eighty-three larval H. longicornis ticks were hatched from eggs laid by adult ticks collected from Shanghai, east of China (Fig. 1), in 2011, pooled into one sample, and named CYP. Ticks were frozen at -80°C for 8 weeks until total RNA was extracted. Additionally, about 100 Anopheles sinensis (A. sinensis) mosquitoes were collected from Jinning, Yunnan province, southwest of China (Fig. 1), in 2009. Mosquitoes were stored in liquid nitrogen until total RNA was extracted. Morphologic features were observed under the anatomic microscope to identify the species and developmental stage of all tick and mosquito samples by entomologists (Y. Sun and R.M. Xu). The captured arthropods were not classified as endangered or protected species and were not privately owned. No specific permits were required for the described field studies.
Small RNA library preparation, sequencing, and data cleaning

The two groups of ticks were disrupted in liquid nitrogen by pestles, from which total RNA was extracted using the Animal Tissue RNA Purification Kit (LC Sciences, Houston, TX, USA), according to the manufacturer's instructions. Mosquitoes were pooled, cleaned in sterilized water, and dried with hygroscopic filter paper. Total RNA was extracted using the Total RNA Purification Kit (LC Sciences, Houston, TX, USA), according to the manufacturer’s instructions. The extracted total RNA was divided into two aliquots, one of which was sent for sRNA sequencing and the other was stored at −80°C for further testing.

Total RNA quality was then analyzed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Sequencing was conducted by LC Sciences (Houston, TX, USA), which included: 1) the brief purification of sRNA (18–30 bp length) from the total RNA; 2) reverse-transcription into cDNA; 3) sequencing using the Illumina GAIIx machine. Following sequencing, the sRNA library was generated according to Illumina’s sample preparation instructions.

The cleanup of the raw data was performed using a proprietary software package, AGGT101-miR v3.5 (LC Sciences, Houston, TX, USA). Briefly, the low-quality reads, simple artificial sequences, and adaptor sequences were removed. Unique sequences were then generated as clean data by collapsing the identical sequences, with the occurrence count of each unique sequence as a tag in the sequence name.

Bioinformatics prediction

Each clean read was compared against the NCBI nucleotide collection (non-redundant nt database) using Blastn with default parameters, and the top 10 hits be outputted in the blast results. After filtering the hits with identity names containing “uncultured organism” or “ribosomal RNA”, and a match length less than 20 bp, the hit with the highest “Max Score” for every query was picked up. The resulting hits were grouped by genus according to its GI number.

In order to identify the most likely pathogens, the number of reads (read-number, RN) and the total matched length (match-length, ML) of each genus were calculated. To minimize the influence of non-specific matches, the total number of base pairs of every genus included in the nt database were counted (Nt-total, NT), and an empirical formula was derived based on the principle that the host species be ranked at the top of the genus list since its sRNAs should be dominant in the dataset. The empirical formula used was: \[ \text{Ratio} = 1000 \times \frac{\text{ML} \times \text{RN}^2}{\text{NT}^3}. \] The results were ranked in descending order with the assumption that the higher a genus was ranked, the more likely it would exist in the sample. The putatively existing pathogenic species were supposed to be the species which are related with known pathogens and ranked highly (usually top-10) within every taxonomic domain. The sRNA sequences were then mapped with reference genomes of species of interest downloaded from GenBank using the program “Reference Assembly” of CLC Genomics Workbench (CLC bio, Aarhus, Denmark). All above-mentioned calculations were conducted with in-house Python scripts (available upon request).

Figure 1. Areas from which ticks and mosquitoes in the study were collected.
doi:10.1371/journal.pone.0090831.g001
Confirmation with polymerase chain reaction (PCR)

Two sets of primers were used to amplify the genomic sequences of target species. The first set of primers were designed using the reads from clean datasets mapped to reference genomes for detection of target species (Table 1). The second set of primers were adopted from previously reported amplification of 16s [12] or 18s rDNA [13] of targeted cellular species including *Rickettsia spp.*, *Coxiella spp.*, and *Aspergillus spp.* (Table 1). The target DNA stored at −80 °C was reverse-transcribed using SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA) and multiple clones were sequenced. The PCR products were cloned into the pGEM-T vector (Promega, Madison, WI, USA) and the resulting clones were then trimmed off and the resulting nucleotide sequences were compared against the NCBI nucleotide collection using Blastn with default parameters.

**Table 1.** Primers used in the study.

| Sample (Target) | Primer Name          | Primer                          | Annealing condition                  | Length |
|-----------------|----------------------|---------------------------------|--------------------------------------|--------|
| XCP (Rickettsia) | XCP-rick-2-1F        | GACGAAAACACCTCATCCAGG           | 50°C×4cycles; −1°C/step;             | 554 bp |
|                 | XCP-rick-2-1R        | TTAGAATAGAGTTGCGG              | 46°C×13cycles; 0.5°C/step;           |        |
|                 | XCP-rick-2-2F        | GTAGGGTTGCTTGTTG               | 46°C×13cycles; 0.5°C/step;           | 138 bp |
|                 | XCP-rick-2-2R        | CGAATAGACAGTGT               | 40°C×35cycles;                        |        |
| XCP (Coxiella)  | XCP-cox-1-1F         | KGATTGAATTTGCGA               | 48°C×13cycles; 0.5cycles/step;        | 959 bp |
|                 | XCP-cox-1-1R         | CCTACTCTGTTGACC              | 42°C×35cycles;                        |        |
|                 | XCP-Cox-1-2F         | ATAGGTTAGGACTGGAA             | 48°C×13cycles; 0.5cycles/step;        | 421 bp |
|                 | XCP-Cox-1-2R         | ATGACCTGTGGTTGCGC             | 42°C×35cycles;                        |        |
| CYP (Coxiella)  | CYP-Cox-2-1F         | AGGAAAGCAGTGATGTTGG           | 50°C×4cycles; −1°C/step;             | 777 bp |
|                 | CYP-Cox-2-1R         | CCTGACCCGAGCTCG               | 46°C×13cycles; 0.5°C/step;           |        |
|                 | CYP-Cox-2-2F         | CTGACCCGAGCTCG               | 40°C×35cycles;                        |        |
|                 | CYP-Cox-2-2R         | ATGAGCTGTGGTTGCGC             | 46°C×13cycles; 0.5°C/step;           | 476 bp |
| XCPSCYP (Coxiella) | XCP-Coxiella spp. 16s_F | ATTGAGAGGTGTGATGTTGG           | 53°C×40cycles;                       | 1457 bp|
|                 | XCP-Coxiella spp. 16s_R | CGCGTCTCCCGAGGTAGG           |                                      |        |
| XCPSCYP (universal 16s rDNA gene [12]) | 27F | GAGATTTGATGTTGCTGCAG         | 55°C×25cycles;                       | 1508 bp in *E.coli* |
| XCPSCYP (Aspergillus genus [13]) | Asp1 | GGCCCCTTTAATAGCAGGGTC     | 53°C×38cycles;                       | 362 bp in *Aspergillus niger* |
| mosquito (ESV) | ESVA-2-1F            | CGAATTGAGCGCAAGAAAGAAGATGA   | 57°C×13cycles; 0.5°C/step;           | 572 bp |
|                 | ESVA-2-1R            | AATAATCGCTGCGATTCAAA          | 51°C×35cycles;                        |        |
| mosquito (NDV) | NDV-RdRp-1-3F        | ATGGTGCCCTCAGAAGTA           | 53°C×13cycles; 0.5°C/step;           | 529 bp |
|                 | NDV-RdRp-1-3R        | GGTGAGAGGTATGTTGAGTT         | 46°C×35cycles;                        |        |

doi:10.1371/journal.pone.0090831.t001
coverage of reference viral genome was obtained for ESV (97.3%, reference GenBank Accession: NC016518) and NDiV (83.5%, reference GenBank Accession: NC015874). This finding also supported the increased likelihood of the sample containing these two viruses (Table 3).

Validation of microbes by polymerase chain reaction

Target sequences were amplified in the validation experiments performed using PCR assays (Fig. 2). The determined sequences of PCR products were subjected to BLAST analysis against the NCBI non-redundant nucleotide sequence database. Results confirmed the presence of Coxella burnetii in XCP and CYP, and Rickettsia spp. in XCP. Similarly, ESV and NDiV in A. sinensis were validated using the primers designed according to the sequences derived from sRNA sequencing. The blast results of the sequences derived from sRNA sequencing are shown in Table 3.

As no eukaryotic pathogens were found likely to be in tick or mosquito samples, we instead chose to validate this strategy using Aspergillus spp. (Table 3) and therefore confirmed its presence in both samples.

Phylogenetic Analysis of target species

To explore the evolutionary status of identified pathogens of interest, we inferred dendrograms using the maximum parsimony method. The Rickettsia spp. identified in XCP was shown to be clustered within Rickettsia spotted fever group and most related with uncultured bacterium clone HLX-1 (JN866573) which was detected from H. longicornis collected in China (Fig. 3A). Our sequence and uncultured bacterium clone HLX-1 constituted a separate branch, indicating the presence of a uncharacterised Rickettsia species in the XCP sample.

In the phylogenetic tree (Fig. 3B), the 16S rRNA gene amplified from the sample CYP was clustered with that of H. longicornis symbiont A, while differing from those of Coxella burnetii strains (RSA 331, Dugway 3108-111, Chug Q212). The 16S rRNA gene of the XCP sample was clustered with that of uncultured bacterium clone DX-68 (JN866592) from Dermacentor silvarum in China, uncultured bacterium clone HLX-3 (JN866574), and uncultured bacterium clone HLC26 (JN866567) from H. longicornis in China. H. longicornis symbiont A belongs to the genus Coxella (Taxonomy ID: 776). The uncultured bacterium clones DX-68, HLX-3, HLC26 were in a cluster with that of Coxella spp. (Rhipicephalus sanguineus sibimont, D84559), which was clustered with C. burnetii (CP000890, CP000733, CP001019). The result showed that the predicted Coxella spp. in the sample CYP and XCP were in different clusters representing different species.

The phylogenetic analysis of Aspergillus spp. (Fig. 3C) demonstrated that our sequences were clustered with the Aspergillus

### Table 2. Coverage of respective microbe genomes by mapped sRNA reads.

| Sample       | Species of pathogens          | Read number | Base number | Genome coverage (%) |
|--------------|-------------------------------|-------------|-------------|---------------------|
| CYP          | Coxella burnetii (NC011528)   | 41316       | 969621      | 3.50                |
|              | Aspergillus nidulans FGSC A4  | 64269       | 1159702     | 2.1                 |
| XCP          | Coxella burnetii (NC011528)   | 22702       | 533741      | 2.50                |
|              | Rickettsia peacockii (NC012730)| 15576       | 369254      | 3.30                |
|              | Aspergillus nidulans FGSC A4  | 47753       | 864327      | 1.5                 |
| A. sinensis  | Espirito Santo virus (NC016518)| 8891        | 202014      | 97.3                |
|              | Nam Dinh virus (NC015874)     | 7346        | 166439      | 83.5                |

### Table 3. BLAST Results of confirmation experiment.

| Sample                  | Primer     | Target Genus/Species | Length | Top Hit Descriptions                        | Query Coverage | Max ID  |
|-------------------------|------------|----------------------|--------|---------------------------------------------|----------------|--------|
| Adult H. longicornis    | XCP-Cox-1-2F | Coxella             | 474 bp | Coxella burnetii CbuG_Q212 (CP001019)        | 67%            | 98%    |
| ticks from Beijing      | XCP-Cox-1-2R |                    |        |                                              |                |        |
|                         | XCP-nick-2-2F | Rickettsia         | 191 bp | Rickettsia montanensis str. OSU 85–930 (CP001019) | 99%            | 98%    |
|                         | XCP-nick-2-2R |                    |        |                                              |                |        |
| Nymphal H. longicornis  | CYP-Cox-2-2F | Coxella             | 480 bp | Coxella burnetii CbuG_Q212 (CP001019)        | 99%            | 98%    |
| ticks from Shanghai     | CYP-Cox-2-2R |                    |        |                                              |                |        |
|                         | Asp1        | Aspergillus         | 362 bp | Uncultured Aspergillus clone 1186 (DQ451600) | 100%           | 100%   |
|                         | Asp2        |                      |        |                                              |                |        |
| Nymphal H. longicornis  | Asp1        | Aspergillus         | 362 bp | Aspergillus candidus strain CBS 567.65 (GU733348) | 100%           | 100%   |
| ticks from Beijing      | Asp2        |                      |        |                                              |                |        |
| A. sinensis             | NDV-RdRp-1-3F | Nam Dinh virus     | 529 bp | Nam Dinh virus isolate SZ11714Z (CP001019)    | 100%           | 99%    |
|                         | NDV-RdRp-1-3R |                    |        |                                              |                |        |
| A. sinensis             | ESVA -2-1F  | Espirito Santo virus segment A | 572 bp | Espirito Santo virus segment A | 98%            | 98%    |
|                         | ESVA -2-1R  |                      |        |                                              |                |        |
species. The sequence amplified from sample XCP was clustered with *Aspergillus restrictus*, while the sequence amplified from CYP was closely related with that of *Aspergillus fumigatus*, *Aspergillus niger*.

For the phylogenetic analysis of NDiV (Fig. 3D), the RdRp gene was amplified and sequenced, as it is conserved in nidoviruses and suitable for use in evolutionary study [15]. Since the RdRp gene nucleotide sequences showed remarkably high diversity, making it difficult to find homology between them, we adopted the amino acid sequences of the RdRp genes to construct the phylogenetic tree. The result showed that the identified NDiV (KC776320, Nam Dinh virus isolate Yunnan) was clustered with different isolates of NDiV (JQ996713, JQ996715, JQ996712) and Cavally virus isolate C79 (HM746600), and was in different branches from Coronaviridae and Arteriviridae.

**Discussion**

In the present study, we demonstrated that not only viruses, but also prokaryotic and eukaryotic pathogens could be screened out of samples using sRNA deep sequencing data.

For the discovery of viruses, our strategy described in this study showed a higher sensitivity of discovering viruses compared with our previous study. Huang *et al* [16] only reported the discovery of Mosquito X Virus (MXV) which is 97% identical to ESV in the same sRNA dataset *A. sinensis* as we use in this study. In that study, he used the sRNA data to BLAST only with the virus database, and no statistical analysis was applied. In our study, the non-redundant nt database (NCBI) were used for BLAST procedure. Statistical analysis of blast results revealed the existence of another virus, NDiV in the sample and subsequent experiments confirmed the prediction.

More over, our strategy also showed the feasibility of screening pathogens of prokaryotes and eukaryotes. Specially, the *Rickettsia* spp., identified in *H. longicornis* captured in Beijing was shown to be most similar with uncultured bacterium clone HLX-1. Phylogenetic analysis of the 16S rRNA gene indicated it likely be a novel species of *Rickettsia* spotted fever group (SFG) which unites a phylogenetically well-defined clade of Rickettsiae that are distinct from other species and that have a life cycle involving arthropods, mainly ticks [17]. SFG includes a number of pathogenic organisms that cause so-called tick-borne (TB) rickettsioses, which can cause

---

**Figure 2. Experimental confirmation of predicted pathogens of interest predicted by bioinformatics.** A), The second run of nested PCR amplification of the sample XCP (*Heamaphysalis longicornis* ticks collected from Beijing) to confirm the predicted *Coxiella* spp. B), The second run of nested PCR amplification of the sample XCP (*Heamaphysalis longicornis* ticks collected from Beijing) to confirm the predicted *Rickettsia* spp. C), The second run of nested PCR amplification of the sample CYP (*Heamaphysalis longicornis* ticks collected from Shanghai) to confirm the predicted *Coxiella* spp. D), The PCR amplification of the sample XCP (*Heamaphysalis longicornis* ticks collected from Beijing) to confirm the predicted *Aspergillus* spp. E), The PCR amplification of the sample CYP (*Heamaphysalis longicornis* ticks collected from Shanghai) to confirm the predicted *Aspergillus* spp. F), The PCR amplification of the sample *A. sinensis* collected from Yunnan to confirm the predicted ESV. G), The PCR amplification of the sample *A. sinensis* collected from Yunnan to confirm the predicted NDV. M, DNA marker; S, Sample; N, negative control. doi:10.1371/journal.pone.0090831.g002
diseases such as Rocky Mountain spotted fever in humans [18]. Additionally, many types of SFG rickettsia have been reported in Asia, Africa, North America, South America, Europe, and Australia [19], reinforcing the plausibility of this study’s predicted Rickettsia spp. strain. In China, there are five species of tick-transmitted SFG rickettsiae that have all been isolated, named R. sibirica [20], R. mongolotimona [21], R. heilongjiangiensis, R. hulinii [22], and BJ-90 strain [23]. Furthermore, molecular evidence of R. raoultii and R. slovaca has been reported in the northeast and northwest of China [24,25]. Serological evidence of Rickettsia japonica [26], Rickettsia conorii, Rickettsia akari [27] has also been reported. The prediction of the genetic sequence of Rickettsia spp. in this study appears to be unique, suggesting a possible novel strain of the bacterium. This finding highlights that potential value this technique could have for species discovery.

Coxiella spp. and Aspergillus spp., from samples XCP and CYP, were confirmed with additional experimentation. Interestingly, the discovery of Coxiiella spp. in larval H. longicornis ticks, hatched from the eggs of adult ticks collected from Shanghai, indicated that transovarian and transtadial transmission of the Coxiiella species had occurred consistent with that of a novel Coxiiella-like agent [28].

However, non-specific match would occur in the prediction results. For example, from the H. longicornis ticks sample, Heamaphysalis was demonstrated as the top eukaryote among tested samples (Table S1). Amblyomma, Ixodes, Dermacentor, Aponomma were also in the top 10 genera of Eukaryota. Some irrelevant genera such as Spirogyra and Psathyropus also appeared in the top 10 genus of Eukaryota, but with short total matched lengths and small numbers of matched reads. These non-specific matches could have been caused by the following reasons: 1) the sequencing platform Illumina GAIIx had an inherent error rate roughly 0.1%–0.5%; 2) sequence homology exists between evolutionarily related species; 3) the BLAST program allows mismatches, which could result in false species assignments; 4) the sequence richness, diversity, and evenness of different species deposited in the nt database were biased. Due to the short length of the reads, these reasons described above would affect the predicted results.

Our study showed for the first time the capacity to screen for a variety of microbial pathogens, not just RNA [2] and DNA [29] viruses, but also prokaryotic, and eukaryotic pathogens, using the sRNA deep sequencing data obtained from wild-caught ticks and mosquitoes. Since microRNAs were originally discovered as critical regulators of developmental timing events in Caenorhabditis elegans [30], interest in understanding microRNAs has increased. Under this context, the sRNA deep sequencing technique was further developed and large amounts of sRNA deep sequencing data were derived from a variety of species. This readily available sRNA deep sequencing data can be used or re-used to find...
putative pathogens, which could be associated with a variety of known and unknown diseases. Our study showed the possibility of discovering pathogens by advanced mining of the sRNA deep sequencing data. This strategy of reuse of sRNA deep sequencing data with the ability of discovering all spectrums of microbial pathogens could have important application in pathogen screening, early warning and tracing the origins of emerging/re-emerging infectious diseases.

Supporting Information

Table S1 Top 10 genus of Eukaryota predicted from deep sequencing data of small RNAs.

Table S2 Top 10 genus of Bacteria predicted from deep sequencing data of small RNAs.

References

1. Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet 8: 93–103.
2. Wu Q, Lao Y, Lai R, Lai N, Dai E, et al. (2010) Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. Proc Natl Acad Sci U S A 107: 1606–1611.
3. Giampetruzzi A, Rousi V, Roberto R, Mahssini U, Yoshikawa N, et al. (2012) A new grapevine virus discovered by deep sequencing of virus- and viroid-derived small RNAs in Cv Pinot gris. Virus Res 163: 262–268.
4. Kreuze JF, Perez A, Untervos M, Quispe D, Fuentes S, et al. (2009) Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. Virology 388: 1–7.
5. van Mierlo JT, van Cleef KW, van Rij RP (2010) Small Silentings RNAs: Picing Together a Viral Genome. Cell Host Microbe 7: 87–89.
6. Aliyari R, Ding SW (2009) RNA-based viral immunity initiated by the Dicer family of host immune receptors. Immuno Rev 227: 176–186.
7. Ding SW, Voinnet O (2007) Antiviral immunity directed by small RNAs. Cell 135: 1567–1568.
8. Mlotshwa S, Pruss GJ, Vance V (2008) Small RNAs in viral infection and host defense. Trends Plant Sci 13: 375–382.
9. Caillé Al, Lang AS, Suttle CA (2006) Metagenomic analysis of coastal RNA virus communities. Science 312: 1795–1798.
10. Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, et al. (2006) Metagenomic analysis of the human distal gut microbiome. Science 312: 1355–1359.
11. Cole C, Sohala A, Lu C, Thatchter SR, Bowman A, et al. (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. RNA 15: 2147–2160.
12. Galkiewicz JP, Kellogg CA (2008) Cross-kingdom amplification using bacteria-specific primers: complications for studies of coral microbial ecology. Appl Environ Microbiol 74: 7829–7831.
13. Garcia ME, Blanco JL, Caballero J, Gargallo-Viola D (2002) Anticoagulants interfere with PCR used to diagnose invasive aspergillosis. J Clin Microbiol 40: 1367–1368.
14. Huang Y, Mi Z, Zhuang L, Ma M, An X, et al. (2013) Presence of entomobirnaviruses in Chinese mosquitoes in the absence of Dengue virus co-infection. J Gen Virol 94: 663–667.
15. Nga PT, Parquet Meol C, Lauber C, Partida M, Nakashima T, et al. (2011) Discovery of the first insect nidorovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. PLoS Pathog 7: e1002215.
16. Huang Y, Mi Z, Zhuang L, Ma M, An X, et al. (2012) Presence of entomobirnaviruses in Chinese mosquitoes in the absence of dengue virus co-infection. J Gen Virol 94: 663–667.
17. Parola P, Paddock CD, Raoult D (2005) Tick-borne rickettsioses around the world: emerging diseases challenging old concepts. Clin Microbiol Rev 18: 719–756.
18. Dantas-Torres F (2007) Rocky Mountain spotted fever. Lancet Infect Dis 7: 724–732.
19. Walker DH, Paddock CD, Dummer JS (2008) Emerging and re-emerging tick-transmitted rickettsial and chafichial infections. Med Clin North Am 92: 1345–1361.
20. Zhang L, Jin J, Fu X, Raoult D, Fouquier PE (2006) Genetic differentiation of Chinese isolates of Rickettsia sibirica by partial ompA gene sequencing and multipacer typing. J Clin Microbiol 44: 2465–2467.
21. Raoult D, Brouqui P, Roux V (1996) A new spotted-fever-group rickettsiosis. Lancet 348: 412.
22. Zhang JZ, Fan MY, Wu YM, Fouquier PE, Roux V, et al. (2000). Genetic classification of "Rickettsia helongiangii" and "Rickettsia hulini," two Chinese spotted fever group rickettsiae. J Clin Microbiol 38: 5498–5501.
23. Zhang JZ, Fan MY, Xu JX, Raoult D (2000) Phylogenetic analysis of the Chinese Rickettsia isolate BJ-90. Emerg Infect Dis 6: 432–433.
24. Cao WC, Zhan L, De Vlas SJ, Wen BH, Yang H, et al. (2008) Molecular detection of spotted fever group Rickettsia in Dermacentor silvarum from a forest area of northeastern China. J Med Entomol 45: 741–744.
25. Tan XC, Lu GY, Shen H, Xie JR, Luo J, et al. (2012) First report on the occurrence of Rickettsia slovaca and Rickettsia raoulti in Dermacentor silvarum in China. Parasit Vectors 5: 19.
26. Lin Y, Sun X, Zhan Z (1999) Serology and pathology study of SFG Rickettsia in Hainan Island. Chinese journal of zoonoses 02: 3.
27. Chen Z, Zhang M, He J (1999) Study of SFG Rickettsia in animals from Ningxia, Fujian province. Chinese Journal of Vector Biology and Control 5: 343–345.
28. Reeves WK, Lofts AD, Priestley RA, Wills W, Sanders F, et al. (2005) Molecular and biological characterization of a novel Coxiella-like agent from Caritis capensis. Ann N Y Acad Sci 1063: 343–345.
29. Ma M, Huang Y, Gong Z, Zhuang L, Li C, et al. (2011) Discovery of DNA viruses in wild-caught mosquitoes using small RNA high throughput sequencing. PLoS One 6: e24736.
30. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-1 encodes small RNAs with antisense complementarity to lin-14. Cell 70: 843–854.