Metagenomics sample preparation and sequencing

Oskar Erik Karlsson[1,2,3], Martin Norling[2], Erik Bongcam-Rudloff[2] ✉
[1]Department of Biomedical Sciences and Veterinary Public Health (BVF), Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden
[2]SLU Global Bioinformatics Centre, Department of Animal Breeding and Genetics (HGEN), Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden
[3]The Joint Research and Development Division of SLU and SVA, OIE Collaborating Centre for the Biotechnology-based Diagnosis of Infectious Diseases in Veterinary Medicine (OIE CC), Uppsala, Sweden

Motivation and Objectives

Metagenomic methods for detection of viruses provide new diagnostic tools to veterinary and public health laboratories, with powerful capacities to detect and to monitor the viromes in clinical samples. The Metagenomics methodology is divided into three main activities or steps: (1) wet-lab methodology; (2) sequencing; and (3) data analysis. Integrating all three parts is of critical importance to the results as well as their interpretation. Our groups at the OIE Collaborating Centre for the Biotechnology-based Diagnosis of Infectious Diseases in Veterinary Medicine, Uppsala, Sweden and at the SLU Global Bioinformatics Centre, Uppsala, Sweden are working with the development and evaluation of the methodological and technological platforms for viral metagenomics. Together with the National Veterinary Institute (SVA), we develop and test methods for extraction of viromes, feasibility of sequencing platforms to deliver metagenomic data-sets and evaluate bioinformatics tools as well as combine them into software packages for analysis and exploration of metagenomes, for separation, classification, assembly and visualization of genomic data in metagenomic samples. The aim of the work is to provide insight into using the metagenomics approach for detection of emerging viruses, monitoring wildlife for known pathogens as well as providing a tool for rapid characterization of viral pathogens in outbreak situations.

The main goal of the work is now to define Standard Operating Procedures for metagenomic investigations of clinical material, integrate current research within environmental ecology for experimental design in a clinical setting, continuously evaluate sequencing technologies (given the rapid turnover of sequencing platforms and chemistry) and integrating the methodology into the diagnostic labs by a risk based model for initiation of metagenomic investigations.

Methods

Samples are processed by whole genome nucleic acid extraction. Targeting the virome fraction of the sample by DNase/RNase treatment, degrading most of the host genome and microbiome genomes present in the original sample. Depending on sequencing technology, see next paragraph, sample might need amplification treatment or nucleic acid concentration. In the case of nucleic acid amplification the included bias must be considered in the final result (Belak *et al.*, 2013).

Sequencing strategies for metagenome retrieval range from clonal amplification combined with Sanger sequencing, to direct sequencing using the Illumina HiSeq system. Depending on chosen strategy, bias might be introduced into the final result. Besides the known bias introduced by the sequencing platforms, the platform can introduce indirect bias due to constraints on input material. By requiring high or low concentrations of input material while handling a mixed or genome depleted sample, bias is introduced by pre-sequencing processing of samples. The most obvious such bias is the whole genome amplification bias, inherent to both Multiple Displacement Amplification (MDA) and Sequence-independent, single-primer amplification (SISPA). Currently the Unknown Virus discovery platform at the OIE CC in Uppsala has successfully deployed our methodology into several sequencing systems; 454 Life Science/Roche, Illumina (HiSeq as well as MiSeq), IonTorrent. As NGS applications for metagenomics is now moving into its first decade the choice of sequencing technology seems to be fairly open, with good experimental design and biologically relevant questions taking precedence for good results.

Results and Discussion

The group has a fairly extensive history of both method adaption/development and applica-

tion in veterinary virology, providing an unbiased toolset for detection of emerging viruses and screening for known families of viruses (Granberg *et al.*, 2013; (Blomstrom *et al.*, 2009; Blomstrom *et al.*, 2010)

Currently large scale screening, prevalence studies as well as studies in viral hotspot zones are performed using the methodology. Integrating the methodology into the framework of bio-preparedness in outbreak situations is also a main goal within the scope of the AniBioThreat program.

With increasing availability of sequencing equipment the field of metagenomisc is only just entering its golden era. Given the possibility of an unbiased methodology for characterization of the microbiome several labs will provide great insight over the coming few years.

Acknowledgements

This work was supported by the Award of Excellence (Excellensbidrag) provided to SB by the Swedish University of Agricultural Sciences (SLU).

The authors would also like to acknowledge support of Uppsala Genome Centre and UPPMAX for providing assistance in massive parallel sequencing and computational infrastructure. Work performed at Uppsala Genome Centre has been funded by RFI/VR "SNISS" Swedish National Infrastructure for large Scale Sequencing and Science for Life Laboratory, Uppsala.

Writing of this publication has been supported by the framework of the EU-project AniBioThreat (Grant Agreement: Home/2009/ISEC/AG/191) with the financial support from the Prevention of and Fight against Crime Programme of the European Union, European Commission – Directorate General Home Affairs. This publication reflects views only of the authors, and the European Commission cannot be held responsible for any use which may be made of the information contained therein.

Support by BILS (Bioinformatics Infrastructure for Life Sciences) is gratefully acknowledged.

References

Belak, S., O. E. Karlsson, A. L. Blomstrom, M. Berg and F. Granberg (2013). "New viruses in veterinary medicine, detected by metagenomic approaches." *Vet Microbiol*.

Blomstrom, A. L., S. Belak, C. Fossum, J. McKillen, G. Allan, P. Wallgren and M. Berg (2009). "Detection of a novel porcine boca-like virus in the background of porcine circovirus type 2 induced postweaning multisystemic wasting syndrome." *Virus Res* **146**(1-2): 125-129.

Blomstrom, A. L., F. Widen, A. S. Hammer, S. Belak and M. Berg (2010). "Detection of a novel astrovirus in brain tissue of mink suffering from shaking mink syndrome by use of viral metagenomics." *J Clin Microbiol* **48**(12): 4392-4396.

Granberg G, Vicente-Rubiano M, Rubio-Guerri C, Karlsson OE, Kukielka D, Belák S, Sánchez-Vizcaíno JM. (2013) Metagenomic Detection of Viral Pathogens in Spanish Honeybees: Co-infection by Aphid Lethal Paralysis, Israel Acute Paralysis and Lake Sinai Viruses. PLOS One. Feb 27. doi: 10.1371/journal.pone.005759