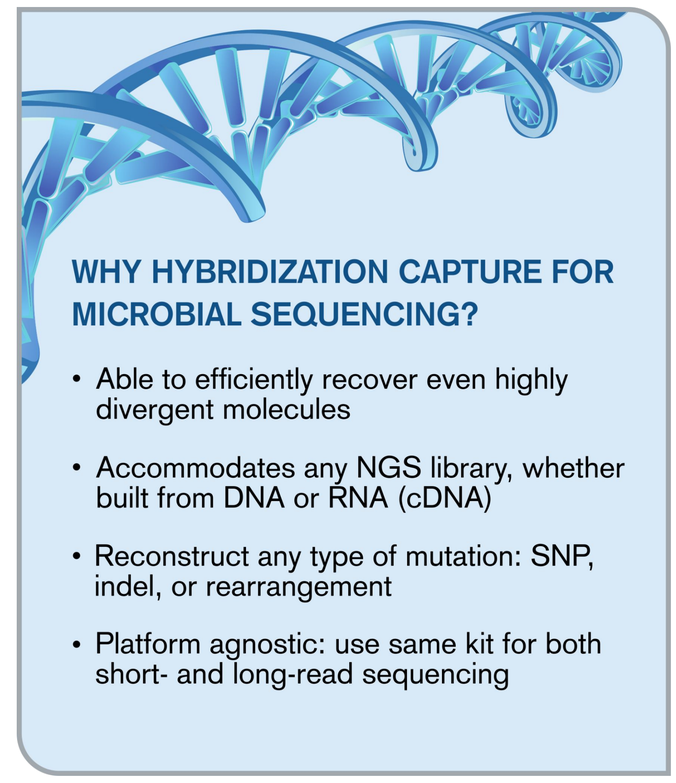
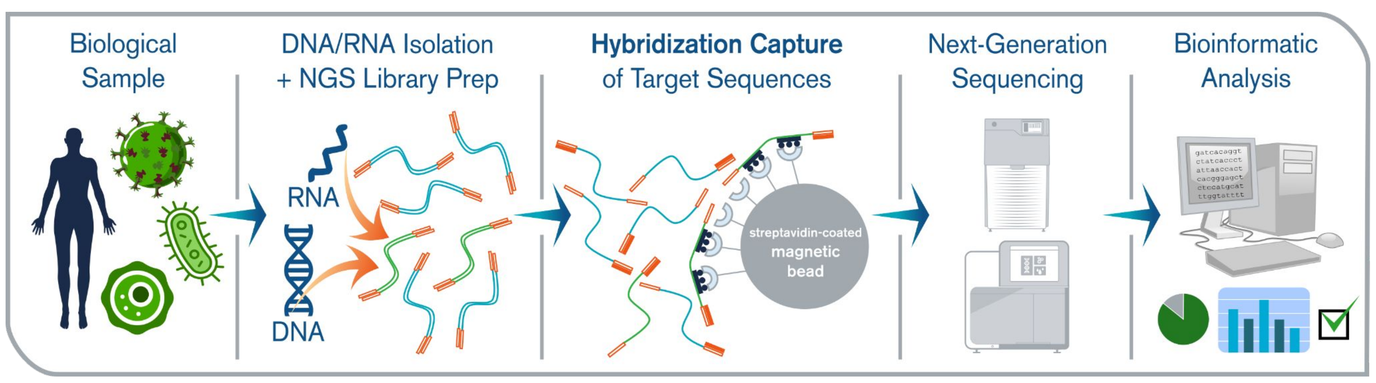
Simplified pathogen and commensal sequencing for plants and animals: Sensitive, versatile and cost-effective microbial genomics using NGS hybridization capture

Brian Brunelle, Alison Devault, Jacob Enk, Tianying Lan
Daicel Arbor Biosciences, Ann Arbor, MI, USA

The detection of pathogenic variants and pathobiome profiling in plant and animal samples using next-generation sequencing (NGS) is often impeded by the dominant host/background DNA and/or RNA, necessitating extremely deep total NGS in order to accurately resolve genomes/genes of interest or to fully characterize community members. Targeted NGS methods can dramatically reduce the overall costs of sequencing and data analysis per sample. Of the available targeted NGS methods, hybridization capture is the best technique for comprehensive, low-bias, and cost-effective genome or community sequencing of viruses and bacteria from complex samples. **Here we review the principles of hybridization capture for microbial sequencing applications and highlight several peer-reviewed studies relevant to the plant and animal research community, powered by efficient myBaits® target capture kits.**



Principle of Hybridization Capture for Efficient Microbial Sequencing



Application Highlights

Targeted NGS enables population genetic studies in plant pathogens

Pathogen virulence or avirulence is largely determined by pathogen effectors, which typically are encoded by a small portion of a given pathogen genome (e.g. < 1% of oomycete genomes) but exhibit high diversity among gene families. To establish a cost-effective alternative to whole genome sequencing, **Thilliez *et al.* (2019)** developed a hybridization capture-based Pathogen Enrichment Sequencing (PenSeq) workflow. By targeting the RXLR effectors, which is < c. 0.2% of the *Phytophthora infestans* genome and c. 0.62% of the *P. capsica* genome, the PenSeq workflow achieved an average c. 300-fold enrichment (56% reads on target) for *P. infestans* and c. 100-fold enrichment (63% reads on target) for *P. capsica* isolates, allowing characterization of allelic diversity in RXLR effectors and identification of presence/absence variations and polymorphisms in key pathogen genes.

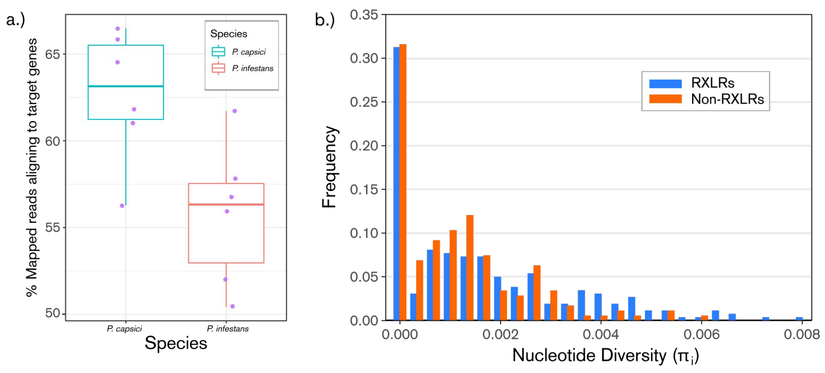


Fig. 1. (a) For 6 *P. capsici* and 6 *P. infestans* enriched isolates, percentages of mapped reads aligning to targeted genes. (b) For 6 *P. infestans* isolates, frequencies of 259 RXLR (blue) and 174 non-RXLR (orange) encoding genes with various nucleotide diversity (π_i) measurements. 7 RXLRs show significantly high π_i values. Data sourced and figures based on Thilliez *et al.* (2019), New Phytologist.

Comprehensive surveillance of antimicrobial resistance gene content

Profiling the genetic content of emerging resistant pathogens is essential to improve surveillance and monitor trends in antimicrobial resistance (AMR). While total DNA sequencing is a valuable discovery tool, it becomes much more resource-intensive and less sensitive for assessing highly complex metagenomic samples. **Guitor *et al.* (2019)** used myBaits to target >2,000 nucleotide sequences associated with antibiotic resistance. This approach not only produced higher on-target read counts and greater breadth of coverage than shotgun sequencing, but also identified additional markers that were inaccessible to shotgun sequencing alone.

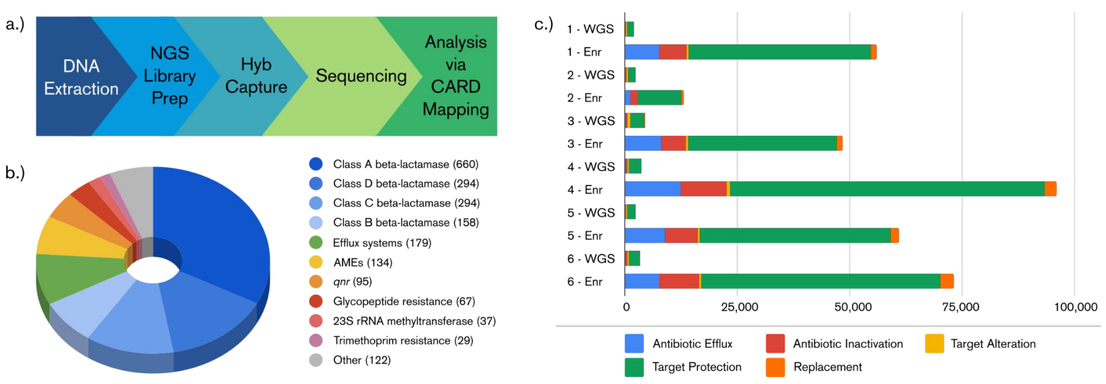


Fig. 2. (a) Overview of the myBaits workflow for identification of ARGs. (b) Breakdown of the CARD annotated resistance gene classes of 2,021 targeted AMR markers. (c) Comparison of read counts summarized per CARD resistance mechanism for 6 human gut metagenome samples between shotgun ('WGS') and enriched ('Enr') libraries. Data sourced and figures based on Guitor et al. (2019).

Sensitive viral diagnosis for epizootic and zoonotic pathogens

Due to large amounts of background information and small virus genome size, using non-targeted sequencing approaches to identify low-abundance pathogenic virus in clinical and environmental samples can be insensitive and costly. For the detection of epizootic and zoonotic viruses, **Wylezich *et al.* (2021)** developed a comprehensive hybridization capture-based panel "VirBaits" based on c. 18,800 complete viral genomes targeting 35 epizootic and zoonotic viruses, including African swine fever virus, Ebolavirus, Marburgvirus, Nipah henipavirus, Rift Valley fever virus, SARS-CoV-2, and more. Validation conducted in 23 complex diagnostic samples showed 10- to 10,000-fold enrichment, demonstrating a high diagnostic performance, including distantly related viruses.

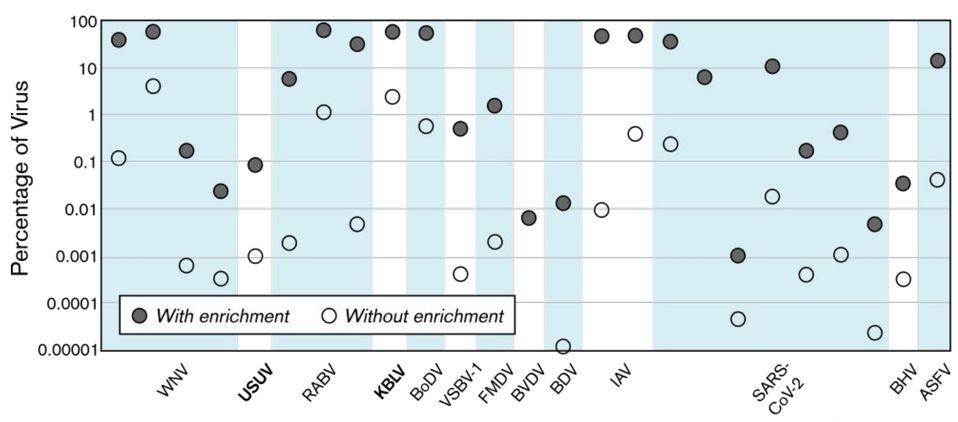


Fig. 3. 25 virus datasets displayed for 23 samples from various animal tissue and mock samples: Percentage of sequenced reads mapping to various virus genomes (labels on x-axis) both with (grey circles) and without (white circles) target enrichment with the VirBaits panel. Data sourced and figure based on Wylezich *et al.* (2021), Microbiome, Fig. 3, p. 8.

High taxonomic resolution analysis of Pacific oyster pathobiota

Accurate profiling and phylogenetic analysis of the pathobiota (the community of pathogens) has been impeded by methodological constraints, such as the lack of phylogenetic value of the 16S rRNA gene and low detection ability of ultra-low-abundance species by shotgun metagenomic sequencing. To address these limitations, **Lasa *et al.* (2019)** for the first time applied target enrichment sequencing for selective capturing of 884 phylogenetic and virulence markers of the pathogenic community in oyster tissue. The highly efficient myBaits enrichment allowed the detection and relative quantification of members of the oyster pathobiota with high taxonomic resolution up to the species level.

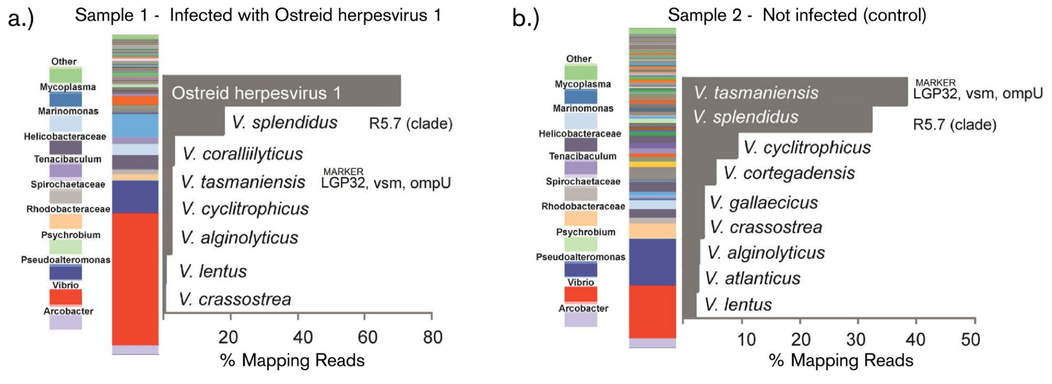


Fig. 4. Results of pathobiota target capture on 2 samples infected (a) or not infected (b, control) with Ostreid herpesvirus 1. *Vibrio* pathogens in control sample were detected. Figure modified from Lasa *et al.* (2019) Enviro. Microbiol., Fig. 4, pg 4555, under the terms of CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

References

Thilliez, G. J. A., *et al.* (2019) *New Phytologist*
Guitor, A.M. *et al.* (2019) *Antimicrobial Agents & Chemotherapy*
Wylezich, C., *et al.* (2021) *Microbiome*
Lasa, A., *et al.* (2019) *Environmental Microbiology*

web: www.arborbiosci.com
email: info@arbor.daicel.com
phone: 1-734-998-0751
twitter: @ArborBio

Custom and Predesigned Targeted Sequencing Kits For Any Application
www.arborbiosci.com/myBaits
01/2022