023 Metagenomics and Targeted Capture Next Generation Sequencing to Detect and Identify Insects in Foods

Monica Pava-Ripoll[1], Mark K. Mammel[2], Elizabeth Reed[3], Martine Ferguson[4], and Padmini Ramachandran[3]

[1]CFSAN, Office of Food Safety (OFS), College Park, MD; [2]CFSAN, Office of Applied Research and Safety Assessment (OARSA), Laurel, MD; [3]CFSAN, Office of Regulatory Science (ORS), College Park, MD; [4]CFSAN, Office of Analytics and Outreach (OAO), College Park, MD.

Abstract

Advances in next generation sequencing (NGS) metagenomics, as well as bioinformatic analyses based on reference genome comparisons, provide the basis for developing novel methods to evaluate for the presence of arthropods in food. Following the development of an internal reference collection (MitochonTrakr) of approximately 20,000 mitochondrial genomes of eukaryotes, 35% of them being arthropods (e.g., insects and mites), we constructed insect-specific probes to enhance the detection of low target DNA levels. The purpose of this study was to evaluate the effectiveness of targeted enrichment via hybridization-based capture using custom insect baits in increasing the detection and identification of insect fragments in a food matrix using NGS metagenomics.

NGS libraries were prepared, with or without targeted enrichment, from genomic DNA extracted from whole wheat flour samples spiked with fragments of a stored-product insect pest, the Indian meal moth, *Plodia interpunctella*. Samples were spiked at 8 levels: 0, 1, 2.5, 5, 7.5, 10, 100, and 1000 ppm, five replicates per spike level (n=40). Insect identification was performed by assigning and matching sequences to the MitoK-mer database. The four-parameter logistic analysis was used to calculate limit of detection (LOD) in pre- and post-target capture data. Targeted capture decreased sequence reads from the food background and increased the LOD by 3.2-fold, from 30.7 to 9.7. This metagenomic approach may be a useful tool for detecting and identifying the presence of low levels of insect fragments in food.

Introduction

Insects are capable of adulterating food products. Current methods for the detection of insect fragments involve their extraction from the food matrix, followed by microscopic analysis to quantify them and determine their taxonomic identification. Both quantification and identification are relevant for regulatory purposes as they provide information on the level and origin of the insect adulteration (for example, field pests versus storage pests). However, microscopy-based taxonomical identification of the insect fragments is time-consuming, requires a lot of expertise from the analysts, and does not always provide resolution to the species level.

Metagenomics allows for the development of new techniques for better detection of eukaryotes such as insects in food samples. When combined with mitochondrial sequencing, metagenomics becomes highly sensitive for the detection of eukaryotic species due to the multiple copies of mitochondrial DNA. After developing an in-house reference collection (MitochonTrakr[1]) of about 20,000 eukaryotic mitochondrial genomes (35% of which are arthropods, including insects and mites), we proceeded to create the MitoK-mer database, which is a combination of short mitochondrial sequence fragments (K=30) of all eukaryote species in our reference collection. Targeted hybridization-based capture has also been shown to be an efficient method to further enhance sensitivity for low-level detection of target DNA in a complex food background. Combining metagenomic sequencing with target enrichment and robust databases, it allows for precise taxonomical identification of low abundance target DNA. The objective of this study was to evaluate next generation sequencing (NGS) metagenomics and target enrichment using hybridization-based capture with insect-specific baits to increase the sensitivity of insect detection and identification in food samples.

Materials and Methods

- Developed an in-house reference collection (MitochonTrakr) of about 20,000 eukaryote mitochondrial genomes (mitogenomes) including arthropods such as insects and mites (Figure 1a).
- Created a MitoK-mer database containing millions of K-mers (K=30) from all mitogenomes in the MitochonTrakr reference collection.
- Used mitochondrial K-mers from approximately 5,600 arthropods in the MitoK-mer database to design a target capture panel containing about 260,000 insect probes for hybridizing insect DNA (Figure 1b).
- Crushed adult insects from the Indian meal moth, *Plodia interpunctella*, in liquid nitrogen and spiked insect fragments in 10 g of whole wheat flour at 8 levels: 0, 1, 2.5, 5, 7.5, 10, 100, and 1000 ppm (5 replicates per spike level; n=40) (Figure 2a).
- Extracted genomic DNA from all 10 g of samples (Figure 2b).
- Prepared pre- and post-target capture libraries using the KAPA HyperPlus kit and sequenced amplified libraries on the Illumina MiSeq system.
- Analyzed sequencing data using the MitoK-mer database to estimate the relative abundance of sequence reads.
- Used the four-parameter logistic analysis to calculate limit of detection (LOD) in pre- and post-target capture data.

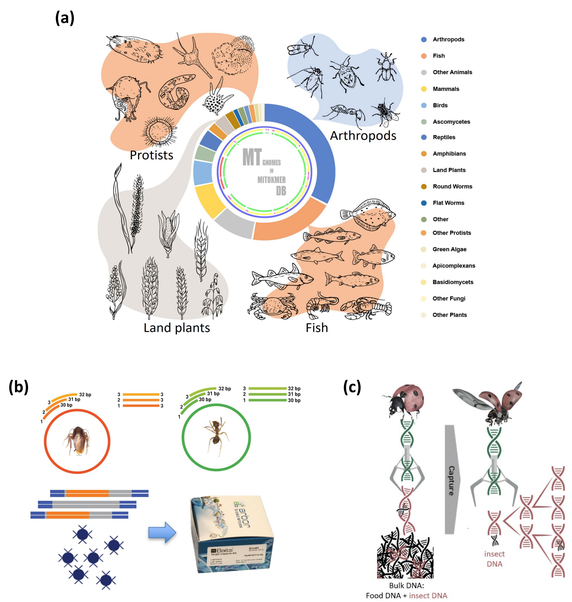

Figure 1. Schematic representations: (a) the MitochonTrakr reference collection and the MitoK-mer database containing eukaryote mitogenomes, (b) insect custom baits design for target enrichment, and (c) targeted capture-based hybridization of insect DNA from a food background.

Results and Discussion

- Identification of eukaryotic sequences from metagenome assemblies using the K-mer comparison approach and our MitoK-mer database resulted in the recovery of most insect sequences from the food matrix.
- Sequence reads from the Indian meal moth, *P. interpunctella*, increased when using targeted capture with custom-designed insect baits.
- While targeted capture decreased sequence reads from the food background, it also increased sequence reads from the target insect in non-spiked samples (Figure 3).
- The relative abundance of spiked insect fragments ranged from 0.1 ± 0.07% to 62.6 ± 2.5% in spiked samples with no baits and from 16.3 ± 1.4% to 99.1 ± 0.05% in post-targeted capture samples (Table 1).
- Targeted capture NGS increased the LOD by 3.2-fold, from 30.7 to 9.7 ppm (Figure 4).
- One of the advantages of the NGS target-based capture approach is that it preferentially amplifies low amounts of target DNA, reducing the breadth and achieving deeper sequencing coverage. This potentially allows for an increased number of samples in a single run.

Figure 2. Preparation of spiking experiment. (a) The Indian meal moth, *Plodia interpunctella*, is crushed in liquid nitrogen and fragments weighted to add them to whole wheat flour. (b) Genomic DNA (gDNA) is extracted from spiked and non-spiked samples.

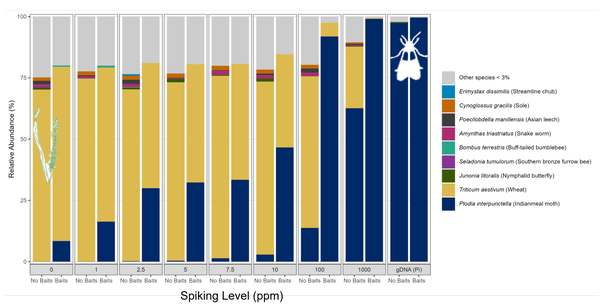

Figure 3. Percent relative abundance of sequence reads from spiked samples in libraries prepared with and without insect baits.

Table 1. Relative abundance (average ± SE) at each spike level (in ppm) from spiked and non-spiked samples prepared with and without insect baits.

	0	1	2.5	5	7.5	10	100	1000
No Baits	0.1 ± 0.07%	0.1 ± 0.07%	0.2 ± 0.2%	0.4 ± 0.2%	1.4 ± 0.3%	2.9 ± 0.8%	13.8 ± 1.7%	62.6 ± 2.6%
Baits	8.4 ± 1.2%	16.3 ± 1.4%	30.0 ± 1.6%	32.3 ± 2.3%	33.4 ± 1.7%	46.6 ± 2.7%	91.9 ± 1.2%	99.1 ± 0.06%

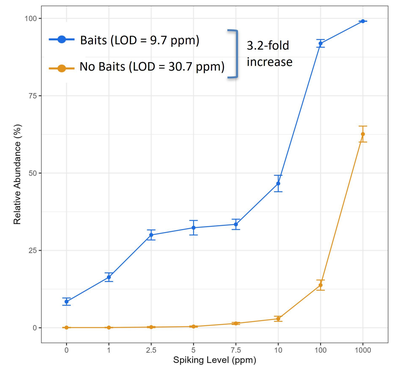

Figure 4. Limit of detection (LOD) of spiked and non-spiked samples prepared with and without insect baits.

Conclusion

- Metagenomics and targeted capture NGS allow for detection of low quantities of insect fragments in food samples that unbiased sequencing would not be able to detect.
- Target capture using our panel of insect baits:
  - Successfully increased the sensitivity of detection of low levels of insect fragments in food samples.
  - Provided accurate level of insect species identification.
- This metagenomics targeted capture NGS approach has the potential of being a useful tool for detecting and identifying insect fragments in food samples.

Acknowledgements: The authors are grateful to Amy K. Miller from CFSAN/OFS for managing the insectarium at CFSAN and for providing insect specimens to perform spiking experiments.

Reference: [1]Ramachandran Et al. (2019). Mitochondrial DNA Part B, 4:1, 292-293.

Disclaimer: The findings and conclusions in this research are those of the authors and do not necessarily represent the views or the official position of the U.S. Food and Drug Administration (U.S. FDA). The names of vendors or manufacturers are provided as examples of available product sources; inclusion does not imply endorsement of the vendors, manufacturers, or products by the U.S. FDA, or by the U.S. Department of Health and Human Services.