Enriching and mining soil and grain metagenomes for novel insecticidal proteins

Irina Shilova[1], Amy Jo Johnson[1], Alex Gulevich[1], Ryan Dowdy[1], Sunit Jain[1], Paul Loriaux[1], Ryan J Williams[2], Ian W Davis[2], Jeff A Haas[2], Shoko Iwai[1], Prasanna Ramachandran[1], Erica Rutherford[1], Kim M Wegener[2], Thomas Weinmaier[1], Yonggan Wu[1], James Baum[2], Todd Z DeSantis[1], Karim Dabbagh[1], Kristen Bennett[1]

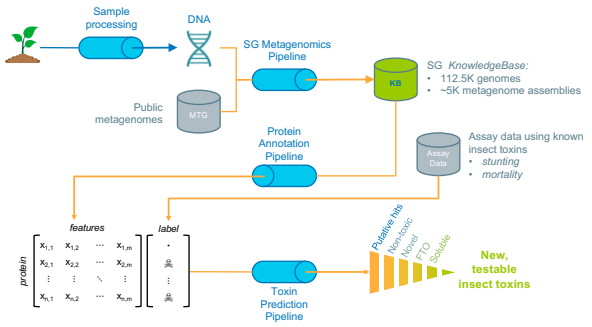
ABSTRACT

Insecticidal microbes and their products have been successfully used as effective and environmentally safe solutions against agricultural pests for more than half a century. However, identifying new insecticidal proteins that can overcome resistance is important to the long-term durability of this cropping system. Conventional approaches to discover novel insecticidal toxins are largely based on spore-selection and strain isolation techniques. Metagenomics allows tapping into the world of uncultured microbes. Soil is an extremely rich source of novel proteins, but high microbial diversity and relatively low representation of microbes that produce insecticidal proteins in soil are prohibitive for assembling genes for insecticidal proteins. To overcome this limitation, we have developed enrichment techniques and metagenomics pipelines coupled with metatranscriptomics approaches to identify novel insecticidal toxins from soil and other environmental samples such as grain dust samples. We used a mix of organic nitrogen and carbon nutrients to facilitate the growth of insecticidal bacteria. The resulting microbial community may have reduced diversity and increased representation of the desired taxa, which can more successfully be used in 'omics-based approaches. We will present the results of mining metatranscriptomes and metagenomes from soil and grain microbial communities in response to enrichments using the Second Genome discovery platform to specifically identify insect toxin genes.

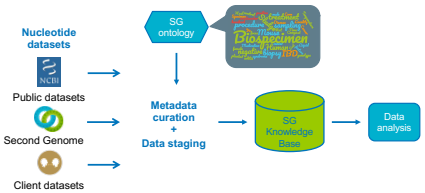
OBJECTIVES

- Determine if the insect-derived substance enriches soil and grain microbial communities for insecticidal bacteria and genes and affects transcript abundances for insecticidal genes
- Evaluate a metatranscriptomic approach for discovery of insecticidal genes

I. SG discovery platform from metagenomes is based on SG knowledge bases, 'omics pipelines and toxin-specific statistical modeling



II. The manually curated ontology system in the SG KnowledgeBase facilitates cross-study comparisons



- Data is obtained from public and private datasets and includes nucleotide and amino acid sequences, gene and transcript abundances and taxonomic annotations
- SG Ontology describes biospecimen and environmental metadata using curated ontology terms

III. APPROACH

- The soil and grain dust microbial communities were incubated with the insect-derived substance (IDS) enrichment. The control incubations included either LB or a base media without the insect-derived substance addition.
- The microbial community composition and diversity was accessed using the V4 16S rRNA Tag-sequencing
- Metagenomes (~10M paired-reads with Illumina NextSeq) were produced from soil enrichment and analyzed for the presence and diversity of insecticidal genes
- Metatranscriptomes (~20M paired-reads with Illumina NextSeq) were produced from grain dust enrichments. Transcript abundances were evaluated for genes homologous to insecticidal genes in comparison to enrichments in LB.

IV. RESULTS

1. Addition of IDS resulted in increased biomass and changes in soil microbial community

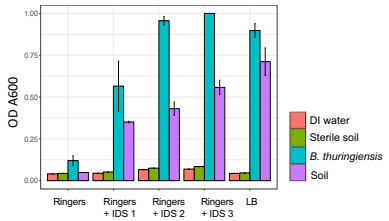


Figure 1. The growth of *Bacillus thuringiensis* HD-1 and soil bacteria was stimulated by the addition of IDS to Ringer's solution. OD was measured after 24 hours of incubation at 30°C, and each treatment had 8 replicates. IDS was added at three concentrations or modifications (IDS 1-3). The biomass yield in the IDS treatments 2 and 3 was similar to the yield in LB.

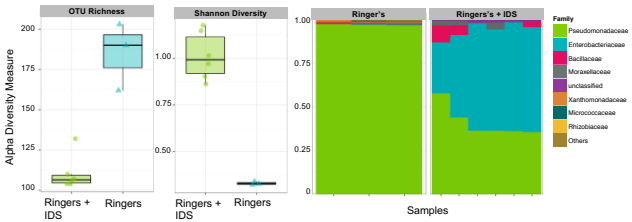
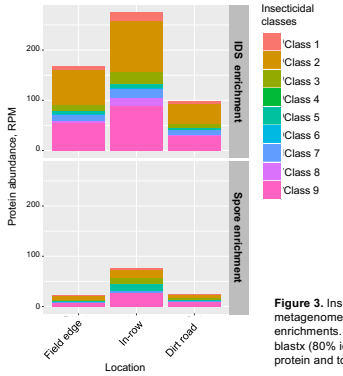


Figure 2. Alpha diversity (left) and relative OTU abundances at the family level (right) in soil samples incubated in Ringers's with the addition of IDS or in Ringer's solution alone.

- Incubation of soil microbial community with the addition of IDS resulted in decreased number of OTUs and increased Shannon diversity compared to the community in Ringer's alone
- Relative abundances of OTUs from family *Enterobacteriaceae* (unclassified at the genus level) increased significantly in the IDS enrichment: from 0.1 % to ~50%
- Other significant increases in the IDS enrichment were seen for OTUs from families *Xanthomonadaceae, Micrococcaceae, and Rhizobiaceae*
- OTUs that identified as *B. thuringiensis* increased in relative abundance in the IDS enrichment (3%) relative to Ringers (0.3%)

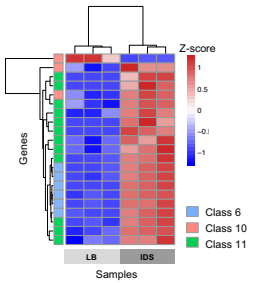
2. Enrichment of soil microbial community with the insect-derived substance stimulated growth of bacteria containing insecticidal genes



- 23 soil samples were collected from three locations in an agricultural field: in-row, field-edge and dirt road
- Soil samples were enriched in 2 ways: 1) addition of the IDS and 2) selecting for spore-forming members by heat-treatment
- Metagenomes obtained from the IDS enrichment contained significantly higher number of insecticidal genes than metagenomes obtained from the same soil samples enriched for spore-forming bacteria.

Figure 3. Insecticidal protein abundances summarized by group found in metagenomes from IDS (top panel) and spore-selected (bottom panel) enrichments. Read hits to known insecticidal proteins was determined using blastx (80% identity over 20 AA length) and normalized to the length of protein and to million of reads (RPM, read per protein per million).

3. The addition of IDS resulted in the increase of transcript abundances for specific insecticidal genes



- A grain dust sample was incubated with the IDS addition and in LB
- RNA samples were collected 20 hours after the start of incubation
- The transcript abundances for insecticidal genes were determined by aligning assembled transcripts to the known insecticidal genes (>78% identity global alignment and insecticidal HMM confirmation)
- Metatranscriptomes obtained from the IDS enrichment contained significantly higher number of transcripts for specific insecticidal genes than metatranscriptomes obtained from the enrichment in LB
- However, it is not clear if the transcript abundance changes observed here are not due to changes in community composition (under investigation)

Figure 4. Insecticidal genes had significantly higher transcript abundances in the IDS enrichment relative to the LB enrichment (ANOVA $p<0.05$). Only genes that were fully assembled are shown here. The transcript abundances in each library were normalized to gene length and to library sequencing depth.

CONCLUSIONS

- The IDS has the potential to promote the growth of insecticidal taxa from soil samples, resulting in the increased number of insecticidal genes in metagenomes. Moreover, the IDS may up-regulate transcription of insecticidal genes in the grain dust microbial community.
- These enrichments in insecticidal gene sequences in metagenomes and also in metatranscriptomes will help discover novel insecticidal proteins.
- The Second Genome Discovery platform is applicable to other fields for the discovery of microbial proteins beneficial for agriculture and human health.

ACKNOWLDGEMENTS

Kayleen Sabino (SG) and Edward Raack (SG) for sequencing, Andrew Han (SG) for discussions.

[1]Second Genome, Inc, 341 Allerton Ave, South San Francisco, CA 94080, USA, [2]Monsanto, 700 Chesterfield Parkway West, Chesterfield, MO 63017, USA

SECOND GENOME
THE MICROBIOME COMPANY
MONSANTO

Contact Information:
kristen@secondgenome.com
irina@secondgenome.com