Rapid Transcriptome Characterization of Green Ash (*Fraxinus pennsylvanica*) Using 454 Sequencing to Study Effects of Emerald Ash Borer (*Agrilus planipennis*) Infestation

Darla French and Rick Meilan
Department of Forestry and Natural Resources, Purdue University

Abstract

At present, North American ash species (genus *Fraxinus*) are under attack by Emerald Ash Borer (*Agrilus planipennis* Fairmaire; Coleoptera: Buprestidae; EAB), an aggressive, invasive insect native to southeast Asia, while native Asian ash species are comparatively resistant to this phloem-feeding insect. To date, little research has been done to determine the basis for susceptibility of North American ash species to EAB. It is possible that Asian species are resistant because of one or more metabolites they produce. One objective of my research is to compare transcriptome profiles of EAB-infested and uninfested control North American green ash trees (*F. pennsylvanica*) in order to better understand the mechanism(s) by which insect feeding acts on putative defense pathways in an EAB-susceptible tree. Following a field experiment, bark tissue samples were harvested and RNA extracted for transcriptomic profiling via high-throughput sequencing. Results have been summarized and are being analyzed to identify families of genes whose expression differs in the fed-upon and control trees. Future work will integrate these results with those from molecular and biochemical methods to predict candidate genes that may be involved in imparting resistance to the EAB.

Introduction

We know very little about the biochemical and molecular biological effects of EAB on *Fraxinus*, which presents challenges for defining experimental parameters for research meant to expand our knowledge in these areas. While we know physiologically how EAB kills ash trees (the wood-boring larvae effectively girdle a tree by feeding on its phloem), we know nothing about the effects of this herbivory on the defense pathways of *Fraxinus*. For example, how quickly does a tree respond? Also, are the effects local or systemic, induced or constitutive, direct or indirect? Experiments to answer these questions can best be designed by identifying a subset of genes whose expression levels are affected directly by insect feeding. This project is a first step toward identifying those genes through sequencing of expressed transcripts in a green ash tree that was fed upon by EAB larvae.

Methods

Two green ash trees (~1.5" DBH) were dug up in May 2009 from a plantation established in 2006 at the FNR Farm. These and two younger trees (~0.75" DBH) raised in the greenhouse were planted, enclosed, and irrigated regularly in separate cages at Huntington Nursery (Huntington, IN). In one cage, an Indiana Department of Natural Resources (IDNR) officer placed approximately 20 EAB adults trapped elsewhere, as well as an infested ash log from which EAB adults were allowed to emerge and infest the young trees. Control (uninfested) trees were planted in the other cage, which acted as an exclosure to prevent insects from reaching the trees. All trees were harvested in August 2009, well after infestation occurred and after larvae monitored in nearby ash trees were large enough to be seen easily. All stem material from each tree was immediately flash-frozen and stored at -80 ºC until RNA extraction. Transcriptomes of infested and control trees were then sequenced using 454 sequencing technology. Initial statistical analysis has focused on chi square and bivariate analyses in an effort to identify genes that are highly expressed in one tree but not the other.

Acknowledgements

Thanks to the following: Huntington Nursery for providing space to run the field experiment; Vince Burkle (IDNR) for performing the EAB infestations; Purdue Genomics Core Facility staff for making the cDNA libraries and performing the high-throughput sequencing; Han Wu (Purdue Statistics Consulting Service) for ongoing advice related to the statistical analyses for this project.

Results

Table 1. Selected run metrics for the 454 sequencing run of our two green ash trees. "Isotig" refers to multiple reads assembled into what is essentially a single transcript of a gene.

Total # reads sequenced	1,121,026
Assembled reads	85.70%
Total # bases sequenced	406,284,567
Total # isotigs	31,089
Mean isotig size	973 bp
Median isotig size	1123 bp
Largest isotig	4080 bp
# annotated isotigs	21,122

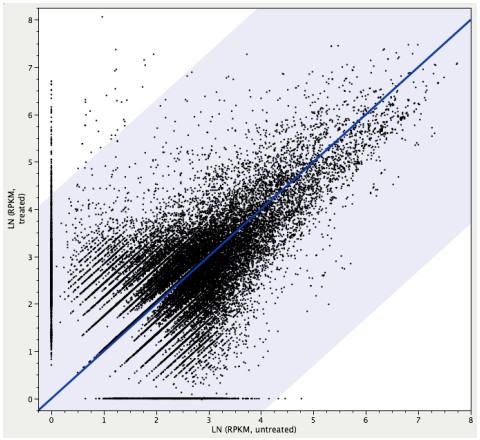

Figure 1. Bivariate fit of the natural log (LN) of the adjusted number of reads for each transcript in the uninfested control green ash (X axis) versus the EAB-infested green ash (Y axis). RPKM stands for reads per kilobase of exon model per million mapped reads, and is an adjustment made on the raw read counts for each transcript to account for differing overall numbers of reads per treatment. 31,008 transcripts are shown. For any one point, location along the X axis represents the relative expression for a single transcript in the control tree, while location along the Y axis represents the expression of that same transcript in the EAB-infested tree. The blue shading represents a 99.9% confidence interval (CI); any points lying outside the blue shading are potentially interesting genes.

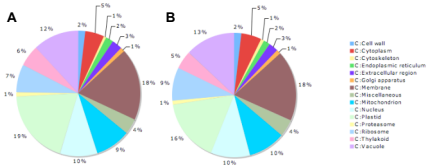

Figure 2. Percentage of total reads per category in uninfested (A) and infested (B) trees for transcripts tagged as **cellular components**. Fourteen categories contain 16,062 unique transcripts. Chi square analysis showed significant difference between the two treatments (X^2 = 6234.2 (N=1,716,862; df=14), *P* = 0.0000).

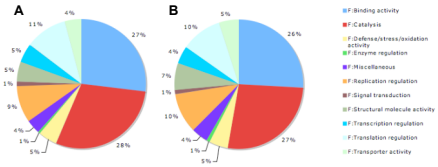

Figure 3. Percentage of total reads per category in uninfested (A) and infested (B) trees tagged as having **molecular functions**. Eleven categories contain 20,151 unique transcripts. Chi square analysis showed significant difference between the two treatments (X^2 = 3895.9 (N=1,456,639; df=9), *P* = 0.0000).

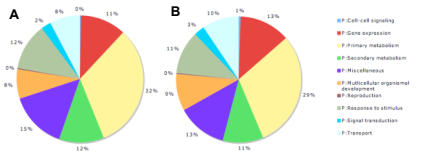

Figure 4. Percentage of total reads per category in uninfested (A) and infested (B) trees for transcripts tagged as involved in **biological processes**. Nine categories contain 17,232 unique transcripts. Chi square analysis showed significant difference between the two treatments (X^2 = 7232.0 (N=1,825,218; df=8), *P* = 0.0000).

Conclusions

The 159 points lying outside the 99.9% CI are potentially interesting genes showing very high expression levels (*i.e.*, a LN ratio of at least 4) under one treatment but not the other (Fig. 1). Any of the 156 genes that were up-regulated in green ash in response to insect feeding have the potential to serve as a biomarker. We have initially focused on a gene whose expression increased nearly 10-fold in the infested tree for possible development as a biomarker in *Fraxinus*. Studies are underway to determine whether this gene shows a strong response to MJ treatment as well as insect feeding; if so, it will be an ideal biomarker.

Further statistical analyses are underway to determine whether the subset of potentially interesting genes identified here (Fig. 1) are statistically different. The chi square tests presented above (Figs. 2-4) all showed significance, suggesting at least one transcript in at least one category across the board was expressed at a significantly different level in the infested tree versus the uninfested control tree. However, the chi square tests do not help identify which transcripts in particular are significant, which will be an important task toward determining what genes might be involved in tree response to insect feeding.