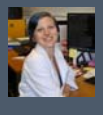


Identification of Differentially Expressed Genes for Mummy Berry (*Monilinia vaccinii-corymbosi*) Resistance in Blueberry

Ashley Yow[1], Kathleen Burchhardt[2] , Marc Cubeta[2], Hamid Ashrafi[1]

[1]North Carolina State University, Department of Horticulture

[2]North Carolina State University, Department of Entomology and Plant Pathology

BACKGROUND

- Molecular data for blueberries is severely lacking.
- Economic loss due to mummy berry has been repeatedly documented across N. America.
- Molecular data is needed to truly understand the genetic mechanisms behind pathogen resistance.
- Next-generation sequencing can generate large amounts of genomic data within a few weeks.
- RNA-seq is a powerful tool for measuring differential gene expression.
- In this study, we used a susceptible cultivar, 'Arlen', for RNA-seq analysis.

OBJECTIVES

- To develop transcriptome sequences of a blueberry cultivar (Arlen) from total RNA extracted from leaves, roots, and different developmental stages of flowers and fruit.
  - Mapping cv. 'Arlen' transcriptome reads to the 'W85-20' reference genome is a good indication of up- and down-regulated genes.
- To develop transcriptome sequences of a blueberry cultivar (Arlen) for mature fruit both infected and uninfected with mummy berry.
- To make a transcriptome assembly of cv. 'Arlen' from Illumina short reads.
- To make a genome assembly of *M. vaccinii-corymbosi* from Roche 454 reads.
- To identify candidate genes related to mummy berry resistance and genes related to the infection process.

METHODS

- Samples were freeze dried in the field and transferred to -80°C freezer on dry ice.
- RNA was extracted from leaves, flowers, and fruits.
- cDNA libraries were created using the KAPA mRNA-Seq kit for Illumina platforms.
  - 200-300bp desired insert size
- Libraries were run using the Illumina Miseq
  - 150bp read length, single-end
- 454 read sequences for *M. vaccinii-corymbosi* were obtained from Dr. Marc Cubeta.
- CLC Genomics v. 9.0.1 was used for trimming reads, de Novo assembly, RNA-seq mapping, and expression analysis.
  - Arlen sequences were mapped to both the 'W85-20' blueberry reference genome and an assembled genome for an NC isolate of *M. vaccinii-corymbosi*
  - For de novo assembly of cv. 'Arlen' we used a word size of 31, bubble size of 150, and minimum contig length of 300bp
  - For de novo assembly of *M. vaccinii-corymbosi* we used a word size of 19, bubble size of 334, and minimum contig length of 200bp
- BLASTx v. 2.4.0+ was used in Linux terminal for blasting contigs to the full NCBI non-redundant protein database.

ACKNOWLEDGMENTS

This research was partially supported by Southern Region Small Fruit Consortium Award No. 2016 R-04 and North Carolina State University.

The authors would like to thank Dr. Kathleen Burchhardt to collect and sequence one isolates of *M. vaccinium corymbosi*.

The authors would like to acknowledge NCSU's Genomic Sciences Lab for performing the sequencing using Illumina MiSeq.

The authors would like to thank Mr. Terry Bland for his help with tissue collection.

RESULTS

Figure 1: Tissues collected for RNA extraction

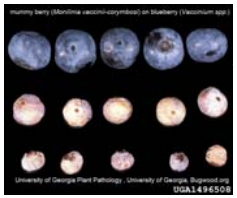


Figure 3: Visualization of mummy berry infection in blueberry fruits

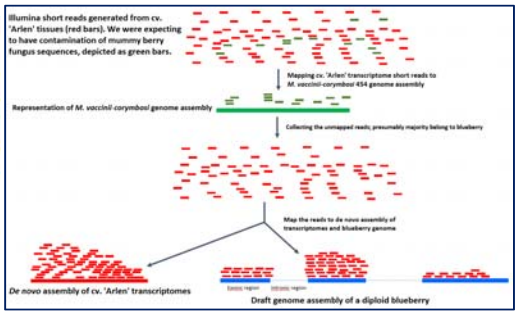


Figure 2: Workflow for performing RNA-Seq analysis

Assembly Stats

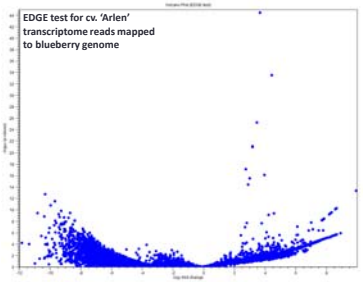
	Arlen Transcriptome Assembly Contig Measurements	*M. vaccinii-corymbosi* Assembly Contig Measurements
N50	1,084 bp	838 bp
Minimum	300 bp	200 bp
Maximum	15,731 bp	14,120 bp
Average	870 bp	754 bp
Count	49,803 contigs	18,988 contigs
Total	43,309,250 bp	14,320,408 bp

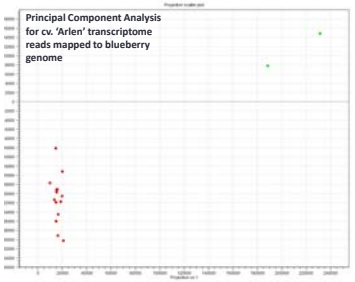
Table 1: Cv. 'Arlen' Transcriptome Assembly and *M. vaccinii-corymbosi* Genome Assembly Summary Statistics

Arlen Transcriptome Reads Mapping Stats

CV. Arlen	Location	Number of Reads	Reads Mapped to W85-20	% of Reads Mapped	Reads Mapped to *M. vaccinii-corymbosi*	% of Reads Mapped
Fruit Stage 1	Castle Hayne	1,257,005	767,816	61.08	7,283	0.58
Fruit Stage 2	Castle Hayne	1,271,969	789,008	62.03	3,084	0.24
Fruit Stage 3	Castle Hayne	1,487,476	855,195	57.49	6,573	0.44
Fruit Stage 4	Castle Hayne	1,206,255	731,511	60.64	3,164	0.26
Fruit Stage 4 Rep 2	Castle Hayne	987,653	584,972	59.23	2,492	0.25
Mummy Fruit Stage 4	Castle Hayne	1,150,852	169,709	14.75	448,469	38.97
Mummy Fruit Stage 4 Rep 2	Castle Hayne	982,291	156,014	15.88	403,572	41.08
Fruit Stage 3	Sandhills	1,121,100	617,924	55.12	7,897	0.70
Fruit Stage 3 Rep 2	Sandhills	1,139,034	681,415	59.82	4,895	0.43
Fruit Stage 4	Sandhills	1,104,765	680,120	61.56	3,267	0.30
Fruit Stage 4 Rep 2	Sandhills	1,031,052	636,119	61.70	3,371	0.33
FL1	Castle Hayne	820,782	505,015	61.53	6,170	0.75
FL3	Castle Hayne	1,021,020	583,335	57.13	8,058	0.79
FL5	Castle Hayne	1,031,618	567,937	55.05	10,578	1.03
L	Castle Hayne	893,901	502,060	56.17	5,898	0.66
R	Castle Hayne	1,193,456	718,407	60.20	6,357	0.53
All Tissues	--	**17,700,229**	**9,546,557**	**53.93**	**931,128**	**5.26**

Table 2: Statistics for cv. 'Arlen' transcriptome reads mapping to the diploid blueberry 'W85-20' reference genome as well as *M. vaccinii-corymbosi* genome assembly





- We were able to assemble 49,803 contigs from our cv. 'Arlen' transcriptome reads after filtering out reads associated with the *M. vaccinii-corymbosi* fungus.
  - The contig N50 for this assembly was 1,084 bp long and our largest contig was 15,731 bp long.
- As expected, we had a higher percentage of transcriptome reads from our infected mature fruits that mapped to the *M. vaccinii-corymbosi* genome assembly due to the presence of the pathogen in the fruit.
  - 38.97% for mummy fruit rep 1 and 41.08% for mummy fruit rep 2.
  - These two tissues also had the lowest percentage of reads map to the 'W85-20' diploid blueberry reference genome.
- The Principal Component Analysis showed that our mummy infected fruits were clustered together and uninfected tissues were clustered together based on RPKM of scaffolds in the 'W85-20' blueberry genome.
- Scaffolds in the 'W85-20' blueberry genome with a high fold-difference between our infected and non-infected tissue groups are of interest because they may contain genes that play a role in mummy berry infection.
  - Scaffold 66203; Scaffold 35130; Scaffold 88151; Scaffold 35999

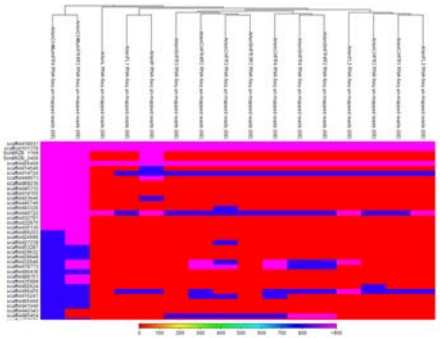


Figure 4: Hierarchical Clustering of Samples for cv. 'Arlen' transcriptome reads mapped to 'W85-20' blueberry reference genome

CONCLUSIONS

- Some of the cv. 'Arlen' transcriptome reads did not map to either the *M. vaccinii-corymbosi* genome or the 'W85-20' diploid blueberry reference genome.
  - The reason for this could be due to differences between cv. 'Arlen' and the 'W85-20' accession, or it could be due to an incomplete blueberry reference genome.
- Annotation on our 'W85-20' diploid blueberry reference genome will be needed in order to identify genes related to mummy berry infection. Currently, the genome is in the form of scaffolds, which could contain multiple genes.
- In the future, we will also include more cultivars on the extreme ends of the spectrum of resistance.
  - In particular we are interested in performing analyses on cv. 'Blue Gold', which is known to be highly susceptible to mummy berry.

REFERENCES

CLC Genomics Workbench 9.0.1 (https://www.qiagenbioinformatics.com/)

American Phytopathological Society & Oregon State Univeristy (http://www.apsnet.org/edcenter/intropp/lessons/fungi/ascomycetes/Pages/MummyBerry.aspx)

University of Georgia Plant Pathology , University of Georgia, Bugwood.org