Rosaceae CSREES NRI PGP Project Accomplishments

Summary of Impacts and Deliverables in Year 1 & 2
Compiled December 2006

EXECUTIVE SUMMARY

In 2005, the USDA-CSREES-NRI Plant Genome Program targeted the family Rosaceae. This report summarizes the accomplishments, broad impacts and deliverables resulting from the twelve projects as of December 2006.

One of the major genomic advantages of the Rosaceae is the small genome sizes (200 to 400 Mbp) of many of its diploid species. Therefore, the development of physical maps for forward genetics approaches and comparative mapping is a compelling research goal for the Rosaceae community. In peach, the physical map is nearing completion and the physical map data was and is being supplied to the community database, Genome Database Rosaceae (GDR), to facilitate gene cloning and comparative mapping. In addition, 400 genetically anchored SSRs were placed on the peach physical map, thereby substantially improving the integration of the genetic and physical map. In apple, a first draft framework of a genome-wide physical map was constructed. A total of 82,503 BAC clones, representing ~13× haploid genome equivalents, were assembled into contigs. Following automatic assembly, 74,281 BAC clones (92%) were assembled into 3,943 contigs, and then hand assembled into 2,702 contigs. In strawberry, the first publicly reported large-insert library has been characterized. Fifty fosmid inserts (average insert size of 35 kb) from a genomic library of the strawberry diploid model species *Fragaria vesca* were sequenced. About 1.75 Mb of genomic sequence was obtained (slightly under 1% of the *F. vesca* genome). Annotation of these fosmid sequences revealed nineteen flowering and fruit quality genes and many SSRs that will be key to integrating the physical and genetic maps.

Linkage map construction and gene/QTL trait associations in many rosaceous species also lagged behind that in many other crop plants. With these projects, major strides were made in the construction of genetic linkage maps and QTL - trait associations. Linkage maps were constructed for diploid and octoploid strawberry and sweet cherry and a map for chokecherry is being initiated. "Gene pair markers", a new marker type that takes advantage of the conserved location of adjacent genes, were demonstrated to be useful for comparative mapping in the Rosaceae as orthologous gene pairs were amplified in *Fragaria*, *Prunus*, *Rubus* and *Rosa*. In strawberry, the genetic basis of floral sex expression and repeat flowering are being investigated. Anomalies observed in segregating diploid strawberry populations indicated that the accepted theory that gender is controlled by one locus with three alleles may be too simple. Three genomic regions were shown to control repeat-flowering in strawberry. However, the genetic control was dependent upon the environment as in milder climates multiple genomic regions had moderate effects on repeat flowering but in hot summer environment only one region was responsible for the flowering phenotype. Understanding the genetic control of flower type and timing in strawberry and ultimately identifying markers useful for marker assisted selection will be critical for these yield-related traits. In sweet cherry, two highly significant fruit size QTL were identified for

fine mapping. As fruit size is the most important market driven trait for fresh market cherries and other *Prunus* fleshy fruits in the U.S. the ability to pre-select for large fruit size in breeding programs will be a major advance. In peach, fruit softening and freestone/clingstone phenotypic differences were demonstrated to be associated with allelic variants at the complex endopolygalacturonase locus. In addition, forty additional candidate genes for softening have been placed on the *Prunus* reference map and are being investigated for their potential roles in fruit quality.

Funded projects also contributed significant advances in the area of functional genomics. A multi-plasmid transformation system was developed for high-throughput generation of RNAi mutants in *Malus.* Universal PCR primers were developed for pHellsgate8-derived plasmids that can detect the presence of single or multiple EST silencing insertions in the RNAi transgenic, and provide sequencing template to determine the EST contained in the silencing insertion. Approximately 50% of the existing *Fragaria* ESTs in GenBank were generated, with more to come. The 9,600 high quality sequences contributed so far are from a cDNA library constructed from cold-stressed tissue. About one quarter of the cold-stress library sequences do not show similarity to a Rosaceae unigene in GDR, and therefore contribute totally new information about genes in this family. Lastly, statistical methods and computer programs are being developed to perform gene expression QTL analysis.

Finally, the genetic diversity of *Malus* accessions within the USDA National Plant Germplasm System was determined by genotyping ~950 *Malus sieversii*, >600 *Malus orientalis*, and >500 Chinese *Malus* accessions using 7 microsatellite primer sets. This information is being used to establish a core set of genotypes that will represent most of the genetic diversity available within *Malus* for use in breeding, screening and phenotypic data collection.

As of December 2006 results from these projects have published in 22 journal articles. Results have also been disseminated via 38 presentations at scientific meetings, 10 seminars, 24 posters, and five web sites in addition to GDR. The projects created training opportunities for 16 graduate students, 24 undergraduate students, 16 post-doctorial scientists and 16 technical personnel. A broad international network within the Rosaceae community has resulted from the projects that include collaborations with over 20 scientists in thirteen countries.

Rosaceae CSREES NRI PGP Project Accomplishments
Summary of Impacts and Deliverables in Year 1 & 2
Compiled December 2006

Table of Contents

Project title: Completion of the Peach Genome Database: A Reference Genome for *Rosaceae*

PI: A. G. Abbott

Co-PIs: W.V. Baird, G. Reighard, B. Sosinski, P. Arus, D. Main, J. Tomkins

Accomplishments:

Aim 1: Complete construction of the physical/genetic map of peach using HICF (High information content fingerprinting) of the existing physical map framework contig clones and those of larger insert libraries of the haploid cultivar 'Lovell'.

- We have substantially completed this project aim. We have fingerprinted 20,000 BAC clones using HICF technologies. These fingerprints include the original physical map of peach and additional fingerprints from larger insert BAC libraries constructed from the haploid and dihaploid "Lovell". Current estimates of genome coverage are at approximately 80%. Currently, we are re-assembling the physical map of peach with these new data to obtain the more complete map. This map will be presented at PAG 2007. Once we have finished this assembly, if necessary, we will utilize BAC end sequencing and hybridization to close the remaining gaps.

Aim 2: Incorporate the physical map data into the Genome Database for Rosaceae (GDR) to provide a reference genome for identification and cloning of genes important to Rosaceous crop development and sustainability.

- All our data is submitted regularly to the database and the current peach genome database is being utilized by numerous laboratories in the search for candidate genes or markers.

Aim 3: Complete development of a high density genetic marker set anchored on the physical map to provide the tools necessary for marker assisted selection, comparative mapping and molecular map development in the less well characterized Rosaceae species.

- We have commenced physical mapping of 400 genetically anchored SSR's developed in the *Prunus* genomics community. These SSR sequences will strengthen the integration of the general *Prunus* genetics map and the physical map/EST database. We are already using the physical map in several comparative mapping projects to study genome evolution and to identify candidate genes for important characters in Rosaceae species.

Broad Impacts:

Although most of the outcomes we have anticipated, we are impressed with the overall utility of the data to work across species boundaries. We are using the data daily to

identify markers for candidate gene containing regions that carry characters for numerous traits of interest. In our initial sequence comparative work with other species, we find high preservation of genome structure in the *Prunus* species. We have also been surprised at the degree of colinearity of the peach genome with species outside the family such as poplar.

Deliverables:

Publications: These publications either directly result from or utilize the information being generated in this continuation project and the preceding project.

1. Zhebentyayeva, T. N., R. Horn, J. Mook, A. Lecouls, L. Georgi, A.G. Abbott, G. L. Reighard, G. Swire-Clark, and W.V. Baird. 2006. A physical framework for the peach genome. Acta Hort. 713: 83-88
2. Horn R., Lecouls A.-C., Callahan A., Dandekar A., Garay L., McCord P., Howad W., Chan H.,Verde I., Ramaswamy K., Main D., Jung S., Georgi L., Forrest S., Mook J., Zhebentyayeva T.N., Yu Y., Kim H.R., Jesudurai C., Sosinski B.A., Arus P., Baird V., Parfitt D., Reighard G., Scorza R., Tomkins J., Wing R., Abbott A.G. 2005. Candidate gene database and transcript map for peach, a model species for fruit trees. Theor.Appl. Genet. 110: 1419-1428
3. Sook Jung, Christopher Jesudurai, Margaret Staton, Zhidian Du, Stephen Ficklin, Ilhyung Cho, Albert Abbott, Jeffrey Tomkins and Dorrie Main, Sept. 9, 2004. GDR Genome Database for Rosaceae): integrated web resources for Rosaceae genomics and genetics research BMC Bioinformatics 2004, 5:130
4. Sook Jung, Albert Abbott, Christopher Jesudurai, Jeff Tomkins and Dorrie Main. 2005. Fequency, Type, Localization and Annotation of SSRs in Rosaceae ESTs, Functional and Integrated Genomics, 5(3):136-43.
5.. Jung,S., Main, D., Staton, M., Cho, I., Zhebentyayeva, T., Arús, P.and A. Abbott 2006. Synteny conservation between the *Prunus* genome and both the present and ancestral *Arabidopsis* genomes. *BMC Genomics* 2006, 7:81
6. Bielenberg, D.G., Y. Wang, G.L. Reighard and A.G. Abbott. 2006. Sequencing and analysis of the peach EVG locus. Acta Hort. 713: 73-82.
7. Lalli, D. A., Decroocq, V., Blenda, A.V. , Schurdi-levraud, V., Garay, L., Le Gall, O., Damsteegt, V., Reighard, G.L., and A. G. Abbott, 2005. Identification and mapping of resistance gene analogs (RGAs) in *Prunus:* A resistance map for *Prunus.* Theor. Appl.Genet. Theor. Appl. Genet. 111:1504-1513
8. Lalli, D.A.[1]*, Krška, B. [5],Abbott, A.G.[1], Badenes, M.L., Damsteegt, V.[3], Polák, J. [4] , Salava, J. A Genetic Linkage Map for an Apricot (*Prunus armeniaca* L.) BC1 Population Mapping *Plum Pox Virus* Resistance, submitted to Tree Genetics and Genomes

Presentations:

1. Zhebentyayea, T. N. 2006, The peach physical/genetic map database: a tool for Rosaceae genomics. Rosaceae Genome Conference 3, Napier, New Zealand

2. Lalli, D. A., Decroocq, V., Blenda, A.V., Schurdi-Levraud, V., . Garay ,L, , Le Gall, O., Damsteegt, V., Reighard, G. L., and Abbott, A. G. (2006) A Resistance map for *Prunus.* 3rd International Rosaceae Genomics Conference. Napier, New Zealand, March 19-22, 2006.

Meeting Abstracts:
Those cited in presentations above and PAG 2007:

Tatyana Zhebentyayeva[1], Ji Hae Jun[1], Bryon Sosinski,[4] Sriyani, Rajapakse[1], W. Vance Baird[2], Robert Ballard[1], , Ralph Scorza[3], and Albert Abbott[1] . A second generation linkage genetic map for f2 peach suncrest x bailey population

Albert Abbott[1], Tatyana Zhebentyayeva[1] , Renate Horn[1] , Ginger Swire-Clark[2], Jennifer Mook[1] , [1] Laura Georgi[1] , Samuel Forrest[1], Barbara Blackmon [3], Jeff Tomkins [3], Vance Baird[2] , Gregory Reighard[2] . An enchanced physical framework for peach

Web sites: Genome Database for Rosaceae (GDR)

Community resources generated (sequences, populations, plant materials)
All of our current work is directed at providing a community resource for Rosaceae researchers. We have generated in this project, a physical/genetic map resource for peach to serve as a model for Rosaceae genomics and gene discovery. We have provided a physical map for peach anchored on the general *Prunus* genetic map, provided extensive peach EST resources (~30,000 EST sequences from fruit, shoot and root) with several thousand unigenes mapped on the physical map/genetic map resource. We have provided a resistance gene map for *Prunus* that is already being used to identify potential candidate resistance genes for nematodes, Plum Pox Virus, powdery mildew and others. In addition we are providing targeted genomic sequences for comparative genome evolutionary studies among the species.

Patents: Evergrowing Gene Patent filed 2006

Training:
Under this current project and the former original project the following personnel have been associated.

Undergraduates:
2005-06 Jonathan Duncan: Molecular mapping of fruit quality characters in peach
2006-07 W. Thompson: comparative genomics of peach and apricot

Graduate Students:
D. Lalli, Ph.D. Genetics 2006 : Resistance gene map and comparative genomics of peach and apricot
S. Forrest, Ph,D. Genetic 2007: BAC library construction and EST mapping.
Fan Shenghua Ph.D. Genetics 2007: Mapping of chilling requirement genes.

Bode Olukulu Ph.D. Genetics 2009: Comparative genomics of peach and other important Rosaceae species

Postdoctoral Associates:
Anne-Claire Lecouls 2000-2002, Peach Physical mapping, and EST databases.
Tetyana Zhebentenyayeva 2003-present, Peach Physical mapping, EST databases, identification and characterization of important genes in Rosaceae biology.

Technicians:
Lilibeth Garay 2000-2004: EST mapping
Jennifer Mook 2001-2006: peach physical mapping, accounts manager

Visiting Scientists:
Dr. Ignazio Verde, Rome Italy, 2002, peach physical map
Dr. Jaroslav Salava Czech Republic, 2001-, Comparative mapping of peach and apricot disease resistance loci.
Dr. Marisa Badenes IVIA Valencia Spain 2003,2006 Comparative mapping of peach and apricot.
Claudia Ricciolini CNR Perugia 2004, RGA mapping in Prunus
Dr. Donato Giannino (Rome, Italy) May 05, Physical mapping in peach
Dr. Jahn Davik (Trondheim, Norway) Sept. 2006-Aug 2007, Comparative mapping in peach and strawberry.
Dr. Renate Horn (Rostock Germany) 2005- one month/year, identification of genes critical to tree architecture in peach.

Collaborations:

All or many of the above listed visiting scientists have worked in collaborations that this project has made possible. In addition, we are beginning to collaborate with the strawberry community to develop a physical map of strawberry, we are working collaborative with the Scottish Crops Institute to comparatively sequence regions of the genome of peach and raspberry carrying genes that control winter dormancy, we are initiating a project to understand nitrogen fixation in Rosaceae with several investigators at Clemson and in the US, we are collaborating with several groups on biomass production in peaches and other trees and finally, we are pushing very hard to get the genome of peach completely sequenced as a reference genome for Rosaceae with other members of the Rosaceae community.

Project Title: Gender Determination, the Key to Germplasm Utilization in Strawberry: Genetic Mapping and Colinearity with Peach.

PD: Tia-Lynn Ashman

Co-PI: Kim S. Lewers

Accomplishments:

Objective 1: Identify and characterize gender-related genes in octoploid wild *Fragaria* via QTL mapping.

- In April 2005, we performed crosses for mapping population 1 and testcrosses to verify candidate genotypes of parent material.
- Before we identified the parents for Mapping Population 1, we screened all eight potential parental candidates with 94 of the available SSR markers and three relevant SCAR markers.
- In April 2005, interspecies hybrid crosses were performed to generate paternal material (half *F. virginiana*, and half *F. chiloensis*) to cross with the "female" parent from Mapping Population 1 in order to generate Mapping Population 2 which will be used to confirm the utility of the markers identified using Mapping Population 1.
- In June 2005, seeds from 28 test crosses were germinated at the University of Pittsburgh. In December 2005, the testcross progeny received a cold treatment, and were returned to the greenhouse in January 2006 where the flower and fruit set phenotypes were scored.
- In January 2006, we planted seed from all nine potential mapping populations and grew them until we identified the desired parents and, therefore, the desired mapping population.
- Mapping Population 1 was clonally propagated to generate six copies. Three copies were evaluated for flowering data in a greenhouse at the University of Pittsburgh.
- In July 2006, DNA was extracted from parents and progeny of Mapping Population 1.
- In August 2006, parents and progeny were planted in a randomized complete block design in the plasticulture production system in a field at the USDA-ARS Beltsville Area Research Center.
- In September through October 2006, flowering and fruiting data were collected from the field planting.
- In April 2006, we performed crosses for Mapping Population 2, and in June 2006, seeds from the cross were planted out and cultivated in the greenhouse at the University of Pittsburgh.
- Meanwhile, a total now of 184 primer pairs have been tested with the parents of Mapping Population 1. Of these, 121 detected polymorphisms, 33 have been used with Mapping Population 1, and 7 linkage groups have been identified from the expected 28 linkage groups.

Objective 2: Determine the level of similarity between Peach (or *Prunus*) and *Fragaria* genomes. Specifically, will markers linked to gender in strawberry amplify a product in peach or have high sequence similarity to a peach sequence? If so, and

the sequence is already mapped in peach, is it near any sterility–related, flowering, or yield loci?

- All SSR primer pairs we are testing in strawberry are also being tested with a pool of 'Bailey' and 'Suncrest' peach DNA acquired from Ann Callahan of the USDA-ARS lab in Kearneysville, West Virginia.
- We have identified two primer pairs from strawberry sequences that amplify a lot of product from the peach DNA we are testing. We test a pool of DNA from both parents, so we do not know if they detect a polymorphism.
- We also test each primer pair with a pool of blackberry DNA and a pool or red and black raspberry DNA. We have identified 21 new strawberry EST-derived primer pairs that amplify a product from blackberry DNA, and 16 that amplify a product from raspberry.
- We were unable to test any "bridging" markers on the candidate parents, because our collaborator, Dorrie Main, has been unable to provide these primer sequences to us. To create "bridging" markers, mapped peach genes were to be compared with all available strawberry sequences to identify those with very similar sequences. Primers were to be developed from the strawberry sequences, preferably from exons spanning introns.

Broad Impacts:

- Two SSR markers, ARSFL 14 and ARSFL 9, are linked in our octoploid Mapping Population 1 as they are in *F. vesca*, diploid progenitor. This helps substantiate the role of *F. vesca* as a diploid genomic reference for strawberry research that could be complicated by the higher ploidy level of cultivated strawberry (also octoploid).
- Mapping Population 1 surprised us by flowering in the fall, shortly after it was planted. We cannot tell yet if this population is segregating for flowering or if all the progeny will flower next fall. If the population does segregate for flowering in the fall, it will be another population that can be mapped for the remontancy trait (often referred to in strawberry as "day neutrality"). This will be important as we already expect multiple genes to contribute to this trait. Indeed, Jim Hancock's research has shown that different QTL are important to this trait in different environments.
- SCAR2, a marker associated with remontancy in *F. vesca*, was not linked to fall flowering in MP1
- Anomalies observed in testcross populations indicate the accepted theory that gender is controlled by three alleles at one locus may be too simple.

Deliverables:

Presentations:

- Ashman, T.-L. 2006. The ecological context for gender evolution: studies in wild strawberry. Center for Evolution, Genes and Genomics, Univ. of St. Andrews, Fife, Scotland. 11/13/2006.
- Ashman, T.-L. 2006. The ecological context for gender evolution: studies in wild strawberry. Department of Plant Biology, Oxford University, Oxford, England. 11/16/2006.

- Ashman, T.-L. 2006. The ecological context for gender evolution: studies in wild strawberry. Estación Biológica de Doñana, Seville, Spain. 11/23/2006.
- Ashman, T.-L. 2006. The ecological context for gender evolution: studies in wild strawberry. Department of Ecology and Evolutionary Biology, UCLA, Los Angeles, CA. 11/29/2006.
- K.S. Lewers. 2007. Genetic mapping with octoploid strawberry. Sponsored presentation. National Clonal Germplasm Repository and Oregon State University. Corvallis, Oregon. 2/14/2007.

Meeting Abstracts:

- Lewers, K.S., T.-L. Ashman, D. Main. 2007. Sex determination of strawberry genotypes: Preparation for genetic mapping of sex. Poster 666. *In* Plant and Animal Genome Conference XV. San Diego, California.

Training:

Undergraduates:

- Sheena Mathew. Summer 2006. PCR reactions.
- Danielle Feather. Fall 2006 through present. PCR reactions.
- Han Tam. 2006. Greenhouse cultivation and scoring of MC1 and MC2.
- Jaleah Robinson. Greenhouse cultivation and scoring of MC1 and MC2.
- Lynn Wright. Greenhouse cultivation and scoring of MC1 and MC2.

Postdoctoral Associates:

- Dr. Laurent Penet. 2006. Evaluating population progeny and parents for flowering traits. Identifying strawberry sequences related to flowering and designing primers.

Technicians:

- Kate Rappaport. 2006. Planted and extracted DNA from mapping population. Trained students in PCR, data organization, annotation, and extraction.
- Ernalyn Peralta. 2005-2006. Planted and extracted DNA from mapping population. Extracted and organized most of the data. Trained students in PCR, data organization, annotation, and extraction.
- Denise Cole. 2005-2006. Test crosses and scoring for MC1 and MC2.

Collaborations: List new interactions made possible from project support:

- Ann Callahan of the USDA-ARS station at Kearneysville, West Virginia kindly donated peach DNA for this project when we were unable to obtain any from other sources.

Project title: Genomic Resources to Improve Fruit Size and Quality in Sweet Cherry

PI: Amy Iezzoni

Co-PIs: Esther van der Knaap, Wayne Loescher, Dechun Wang

Accomplishments:

The overall goal of this project is to develop the genomic resources necessary to implement marker-assisted selection for fruit size and quality traits in cherry breeding programs.

Objective 1: Construct a sweet cherry genetic linkage map for comparative mapping with the *Prunus* reference map and other *Prunus* linkage maps.

- All 378 of the currently publicly available SSR markers mapped on the *Prunus* concensus map (TxE) were screened to determine if they could be placed on the NY and EF linkage maps. Only ~25% of these SSRs could be mapped as the remaining 75% either did not amplify or were not heterozygous in the parents.
- Due to the low level of SSR transferability and marker polymorphism, the van der Knaap lab is designing markers based on ESTs with known TxE locations that target gaps in our current maps. In addition, they have developed markers for one cherry vacuolar invertase gene and two sorbitol transporter genes.
- Currently, EF and NY54 maps consist of 8 and 9 linkage group, respectively. The EF map is 643 cM while the NY map is 419 cM. The average distance between marker loci for the EF and NY linkage maps are 5.7 cM and 6.6 cM respectively. All 16 of the linkage groups have been aligned with the *Prunus* map based on shared SSR markers.

Objective 2: Identify QTL for fruit size and quality traits.

- In 2006, fruit weight, diameter, and length were measured from five fruit from 153 progeny individuals and three highly significant QTL were identified using stringent criteria. On linkage group 2, for both EF and NY, fruit weight QTL were identified that explained 15.6% and 6.5%, respectively, of the phenotypic variance. In NY, another highly significant fruit weight QTL was identified on linkage group 6 that explained 19.5% of the phenotypic variance. Fruit diameter QTL were also identified for the same linkage map regions.
- From 149 of these progeny, the cellular components of fruit size are being measured: mesocarp cell number and cell length. The data available from 100 progeny indicates that both cell size and cell number are segregating among the progeny, and that high cell number and large cell size is not correlated.
- Fruit 70 progeny were evaluated for sorbitol, glucose, fructose and malic acid. These parameters varied among the NY x EF progeny individuals, however, the sample size in 2006 was too small to draw any conclusions regarding QTL loci.

Objective 3: Fine map the major QTL identified and design markers for marker assisted selections.

- For all three fruit weigh QTL identified in the NY x EF population, the significant marker trait associations extended over at least 20 cM regions. To increase the precision with which these QTL are located we are (1) increasing the marker density using ESTs from the TxE reference map that are predicted to map to our target regions and (2) extending the analysis to the entire population.

Objective 4: Validate the QTL across genetic backgrounds and identify QTL alleles.

- In collaboration with Dr. Marco Bink, we are utilizing QTL validation and allele mining software developed to analyze genotypic and phenotypic data from multigenerational pedigrees (FlexQTL). Our cherry pedigree set which consists of 41 cultivars representing up to four familial generations, has been genotyped for 51 of the mapped SSR markers and scored for the fruit weight, width, length, sorbitol, glucose, fructose, and malic acid. We are currently in the process of double checking our marker data as a prelude to running FlexQTL.

Broad Impacts:

- The cherry linkage map currently under construction will provide the first detailed intra-specific cherry linkage map with sufficient markers suitable for comparative mapping. Additionally this map will have markers that are robust and polymorphic within cherry germplasm used in breeder's collections so that it will finally be possible to conduct a routine genome scan in a breeding program to efficiently determine the map locations of genes controlling trait variation.
- Fruit size is the most important market drive trait in for fresh market sweet cherries, as the U.S. grower price is based solely on fruit size and the difference between profitability and net financial loss is frequently as little as 5 mm in mean fruit diameter. Our ability to identify QTL allele that control large fruit size will be critically important for cherry breeding programs that must release cultivars that meet a minimum fruit size standard, yet have improvements in other taste such as flavor and disease resistance.
- Fruit crops are highly heterozygous and breeding programs resemble family pedigrees. Therefore, breeding populations, due to their inherent family structure and integration of genetic variability, provide an excellent venue for QTL validation and allele mining. In this project we are testing this approach and adopting it to current efforts in cherry breeding.

Deliverables:

Publications:
Iezzoni, A. and J. Olmstead. 2006. Agricultural Biology in the 21st Century: It's about the genes. Compact Fruit Tree 39: 12-15.

Presentations with published abstracts:

Olmstead J, Iezzoni A, Whiting M. 2006. Fruit size QTL in sweet cherry: Cell number is under stronger genetic control than cell size. Prog. & Abst. 3rd Intl. Rosaceae Genomics Conference, Napier, NZ, pg 38.

Olmstead, J, A. Sebolt, and A. Iezzoni. 2006. Targeted mapping of a fruit size QTL to the S-locus in *Prunus cerasus*. Abstract Guide, Plant & Animal Genome XIV, San Diego, pg. 227.

Presentations without published abstracts:

Iezzoni A. 2006. Agricultural Biology in the 21st Century: It's about the genes. Conference of the Intl. Dwarf Tree Fruit Assoc., Hershey, PA.

Iezzoni, A. 2006. Maximizing fruit size and quality: a breeder and geneticists approach. Conference of the Intl. Dwarf Tree Fruit Assoc., Hershey, PA.

Posters:

Olmstead J, Iezzoni A. 2006. Alignment of sweet cherry linkage groups with the *Prunus* reference map. 3rd Intl. Rosaceae Genomics Conference, Napier, NZ.

Iezzoni, A, Olmstead, J., A. Sebolt, Y. Chen, and D. Wang. 2006. Construction of a sweet cherry linkage map suitable for comparative mapping in *Prunus*. Abstract Guide, Plant & Animal Genome XIV, San Diego, pg. 227.

Olmstead, J, A. Sebolt, Y. Chen S. Hammar, G. Iriarte, D. Wang, E. van der Knaap[4] and A. Iezzoni. 2007. Sweet cherry linkage maps constructed from a pseudo-testcross mapping population. Abstract Guide, Plant & Animal Genome XV, San Diego.

Web sites: www.cherrygenetics.org

Community resources generated (sequences, populations, plant materials): An F_1 mapping population consisting of 574 progeny from a cross between a domesticated sweet cherry and a wild forest trees sweet cherry is planted at the MSU research farm is mature and producing fruit. 190 of the progeny in this population are genotyped for the markers used to create the linkage map. The entire population was also genotyped for four segregating *S*-RNase alleles.

Training:

Dr. Yiwu Chen, Post-doctoral Researcher with Dr. Wang, Linkage map construction, QTL analyses, FlexQTL

Suneth Sooriyapathirana, current PhD student with Dr. Iezzoni, QTL analysis of fruit size.

Antonio Cabrera, Current MS student with Dr. van der Knaap, marker development from EST sequences.

Dr. James Olmstead, Former PhD student with Dr. Iezzoni, genotyping and linkage map construction

Gloria Iriarte, former student with Dr. van der Knaap, marker generation and genotyping

Audrey Sebolt, research technician with Dr. Iezzoni, genotyping and fresh fruit phenotyping

Susan Hammar, ½ time research technician with Drs. Loescher/Iezzoni, genotyping and phenotyping of traits on the GC
Elizabeth Loconto, undergraduate student with Dr. Iezzoni, fresh fruit phenotyping

Collaborations:

- Dr. Marco Bink of Plant Research International, Wageningen, The Netherlands is a consultant on our project. Dr. Bink and his colleagues have led the way in the adaptation of human pedigree based software to test QTL associations in outbreeding plants. They are testing this software as part of the European HiDRAS project (involving 11 countries) whose goal is to breed high quality disease resistant apples. Therefore, we not only have access to the latest statistical software, but through Marco, we also benefit from the experiences and knowledge gained in the HiDRAS project.
- Dr. Cameron Peace (Wash. State Univ): screening cherry germplasm using candidate genes for flesh firmness.
- Drs. Jim Olmstead and Matt Whiting (Washington State Univ.): collaboration with the WSU breeding program regarding the capabilities of FlexQTL.
- Dr. Elisabeth Dirlewanger (INRA, France): comparison of linkage maps in sweet cherry and discussions of common phenotyping approaches.

Project title: Development Of Segregating Populations For Molecular And Genetic Analyses Of X-Disease Resistance In Chokecherry (*Prunus virginiana* L.)

PI: Wenhao Dai

Co-PIs: James A. Walla

Accomplishments:

- Development of three segregating populations of X-disease resistance in chokecherry. Total 909 hybrid seeds were obtained.
- Hybrid seedlings are grown in the greenhouse.
- Selection and confirmation of X-disease phytoplasma sources for inoculation.
- Chromosome analysis confirmed that *Prunus virginiana* is a tetraploid and will be a unique *Prunus* species for genetic research.
- Molecular marker analysis for chokecherry species, such as RAPD and SSR, have been attempted to develop reliable strategies of identifying molecular markers linked to X-disease resistance.
- A linkage map of chokecherry is being constructed using TRAP marker and other marker systems.

Broad Impacts:

X-disease and other phytoplasma-associated diseases are devastating to many high value cash crops and forest trees, such as stone fruits, vegetables, ornamentals, and plantation trees, causing billions of dollars of damage each year, yet there is no effective control method for any of these diseases. The purpose of this project is to develop a well-designed segregating population of chokecherry used to identify molecular markers linked to X-disease resistance genes. Therefore, the success of this project will assist us to understand host/pathogen interactions and genetics of resistance genes in chokecherry, and to develop effective methods of managing X-disease or other phytoplasma-associated diseases.

Deliverables:

Publications: None.
Presentations: None.
Meeting Abstracts: Development Of Segregating Populations For Molecular And Genetic Analyses Of X-Disease Resistance In Chokecherry (*Prunus Virginiana* L.). 2007 PAG conference.
Web sites: coming soon.
Community resources generated (sequences, populations, plant materials): Three chokecherry segregating populations.
Patents: None

Training:

Undergraduates: Sarah Gegne and Karissa Rothmeire
Graduates: Vicki Magnussion and Andrea Swanberg

Collaborations: This project is a collaborative effort in the multidisciplinary research field including plant molecular biology, plant pathology, plant genetics, and biotechnology. Additional collaborators:
Dr. J. Hu, Plant Molecular Genetics, USDA-ARS in Fargo, ND.
Dr. C.C. Jan, Plant Cytogenetics, USDA-ARS in Fargo, ND.

Project title: Candidate Genes for Fruit Softening in *Prunus*

PI: Cameron Peace

Co-PIs: Ann Callahan, Carlos Crisosto, Thomas Gradziel

Accomplishments:

The first goal is to elucidate the molecular genetic organization of the complex endopolygalacturonase (endoPG) locus in peach.

- Re-screening of two Clemson peach BAC libraries specifically for this gene by collaborators (Tatyana Zhebentyayeva and Bert Abbott) at Clemson University identified many BACs containing sequence with significant homology to the endoPG gene.
- From fingerprinting, subcloning, and sequencing of these BACs, the basic organization of the *Freestone-Melting flesh* (*F-M*) locus is emerging. The F allele appears to contain two copies of the gene, separated by less than 50-kb, and the f allele only one of the gene copies. EndoPG-like gene sequences also were detected immediately adjacent to the target endoPG genes. Further investigations are underway to determine the function of each gene sequence.

The second goal is to understand the role of endoPG in *Prunus* fruit softening.

- Complete co-segregation was verified on approximately 350 further progeny segregating for the freestone/clingstone and melting/non-melting flesh traits, greatly increasing our confidence that endoPG does indeed control *F-M*.
- Allelic diversity was assessed in *Prunus* via PCR of the microsatellite associated with the endoPG gene. Over 200 different alleles and gene copies of endoPG were detected from a survey of approximately 650 accessions of *Prunus* (peach, almond, apricot, plum, sweet and tart cherry, and many closely related species). This included over 100 peach varieties, for which the genotypic survey was capable of ascertaining the correct F-M fruit type for many varieties that previously had an unclear phenotype. Sweet cherry was the only species that did not show abundant polymorphism for endoPG.
- A subset of accessions was chosen for targeted sequencing. To date, endoPG sequences for about 20 accessions have been obtained and analyzed. EndoPG sequences were isolated through cloning, and for many of these accessions, multiple sequences were obtained as expected. These results are beginning to provide insights into the phylogeny of the locus in *Prunus*.
- To better understand the qualitative and quantitative effects of the various alleles and gene copies in peach, fruit softening characteristics were examined for two harvest seasons (2005 and 2006) for about 70 varieties each season. Optimization of the protocol was required for the second year's firmness to better assess functional groups. Statistical analyses are still underway to determine minor allele differences within each of the major phenotypic categories.
- Fruit tissue was obtained and frozen for physiological assays (gene expression, enzyme activity, and metabolite levels), which are currently underway.

The third major goal of the project is to provide molecular tools to facilitate identification and/or development of new fruit quality phenotypes of important *Prunus* crops.

- The endoPG PCR itself is already valuable in determining Freestone-Melting flesh fruit type of peach and nectarine. Surveys of varieties of these crops have validated the gene-phenotype association. We are still examining the extent to which the gene controls similar traits in other *Prunus* crops.
- The project is not just concerned with the endoPG gene, and to discover further genes that may affect fruit softening in *Prunus*, 40 additional candidate genes for softening were screened via PCR across *Prunus* accessions. Most genes were polymorphic between species. More than half were polymorphic within at least one of the *Prunus* crops. About a quarter of the genes tested were also detected in apple, indicating high homology of gene sequence, and potentially also function.
- The Stony hard trait (controlled by the *Hd* locus) was specifically investigated using the candidate gene approach based on size polymorphism of amplified gene sections. While we now have a much better understanding of the physiological interaction between *Hd* and *F-M*, no clear candidates for the controlling gene for *Hd* were identified, though some likely ones were all but ruled out.
- Candidate genes were also located on the *Prunus* reference map, mostly using the innovative bin-mapping approach. 80% of any gene sequences tested with simple PCR (and separated on large polyacrylamide gels) could be located on the *Prunus* genetic map using bin-mapping. This allowed construction of the "softening gene map" of *Prunus*, containing the genetic locations of more than 50 genes putatively involved in fruit softening.

Broad Impacts:

The massive genotypic diversity detected for endoPG in *Prunus* has strong implications for fruit evolution and crop domestication in this genus, and represents much potential for genetic improvement of flesh softening and stone adhesion attributes in stone fruit. The bin-mapping strategy, recently developed for *Prunus* at IRTA, Spain, is an excellent accompaniment to the candidate gene approach. Any DNA sequences of interest can now be readily located on the *Prunus* genetic map – we expect "gene maps" for every agronomic trait to soon appear from various labs. Such maps will greatly help the identification of important functional genes. The development of bin-mapping for other rosaceous crops is strongly recommended. The project's activities have allowed involvement in RosPOP, a new consortium for the sharing of data and physical resources of mapping populations. RosPOP represents a never-before-seen level of international intrageneric collaboration, and should result in rapid advances in functional and comparative genomics for Rosaceae. As an early example of RosPOP, efforts are currently underway to comparatively map the distal end of *Prunus* linkage group 4 where the endoPG gene and *F-M* locus reside, involving more than a dozen *Prunus* mapping populations. An intergeneric partnership to comparatively map softening candidate genes between peach and apple, unexpectedly arising from this NRI project, is hoped to bridge the genomic gap between these important rosaceous crops.

Deliverables:

Publications

Peace CP, Callahan AM, Ogundiwin EA, Potter D, Gradziel TM, Bliss FA, Crisosto CH. Endopolygalacturonase genotypic variation in *Prunus*. Acta Hort. (in press)

Zhebentyayeva TN, Jiwan D, Jun JH, Reighard GL, Lalli DA, Forrest S, Duncan J, Main D, Abbott AG, Callahan A, Scorza R. Exploitation of structural and functional genomics databases for gene identification in peach. Acta Hort. (in press)

Peace C, Crisosto C (2006). Revealing the genetic control of internal breakdown in peach. Summerfruit Australia Quarterly 8: 20-21

Peace C (2006). Long-term approaches to increase peach fruit consumption. Compact Fruit Tree 39: 15-17

Presentations

Oral Presentations

American Society for Horticultural Science Annual International Conference 2005, Las Vegas, Nevada, June 2005. "Candidate genes for fruit softening in *Prunus*". Cameron Peace

Plant & Animal Genome XIV Conference, San Diego, California, January 2006. "*Prunus* projects of the USDA CSREES National Research Initiative: synergies and progress". Cameron Peace

International Fruit Tree Association 49th Annual Educational Conference, Hershey, Pennsylvania, March 2006. "Long-term solutions to increasing peach fruit consumption". Cameron Peace

3rd International Rosaceae Genomics Conference, Napier, New Zealand, March 2006: "Fruit softening in *Prunus*: progress and prospects of the candidate gene approach". Cameron Peace

Posters

UC Davis Plant Sciences Symposium, Davis, California, September 2005. "Molecular genetics of fruit quality at Kearney Agricultural Center". Cameron Peace

ISHS Symposium on Biotechnology of Temperate Fruit Crops and Tropical Species, Daytona Beach, Florida, October 2005. "Endopolygalacturonase genotypic variation in *Prunus*". Cameron Peace

Plant & Animal Genome XIV Conference, San Diego, California, January 2006. "Candidate gene analysis of *Prunus* fruit softening: research update". Cameron Peace

International Rosaceae Genomics Conference, Napier, New Zealand, March 2006. "Endopolygalacturonase marker-assisted selection for novel fruit types in peach". Tom Gradziel

Meeting Abstracts

Peace CP, Callahan AM, Ogundiwin EA, Potter D, Gradziel TM, Bliss FA, Crisosto CH (2005). Endopolygalacturonase genotypic variation in *Prunus*. International Symposium on Biotechnology of Temperate Fruit Crops and Tropical Species Abstracts p114

Peace CP, Abbott AG, Dai W, Iezzoni AF, Arus P, Baird WV, Callahan AM, Crisosto CH, Gradziel TM, Loescher W, Main D, Reighard G, Sosinski B, Tomkins J, van der

Knaap E, Walla JA, Wang D (2006). *Prunus* projects of the USDA CSREES National Research Initiative: synergies and progress. Plant & Animal Genome XIV Abstracts W-131, p37

Peace CP, Callahan AM, Ogundiwin EA, Potter D, Gradziel TM, Bliss FA, Crisosto CH (2006). Candidate gene analysis of Prunus fruit softening: research update. Plant & Animal Genome XIV Abstracts P-493, p225

Gradziel T, Peace CP, Crisosto C (2006). Endopolygalacturonase marker-assisted selection for novel fruit types in peach. 3rd International Rosaceae Genomics Conference Abstracts J3, p117

Peace CP, Ogundiwin EA, Gradziel TM, Potter D, Weeks C, Badenes ML, Iezzoni AF, Bliss FA, and Crisosto CH (2006). Fruit softening in *Prunus*: progress and prospects of the candidate gene approach. 3rd International Rosaceae Genomics Conference Abstracts OP24, p44

Web sites
http://www.mainlab.clemson.edu/gdr/community/funding/peace.shtml

Community resources generated (sequences, populations, plant materials)
Candidate gene sequences, primers, and assay conditions, and DNA of population parent cultivars have been supplied to researchers requesting them.

Patents: None

Training:

UC Davis, Kearney Agricultural Center
Daniel Edge-Garza (technical, UC Davis) – contributed to the understanding of the functional role of endoPG alleles by locating and obtaining fruit samples, performing softening measurements, supervising summer work crew for the postharvest lab, and tissue preparation and dissemination to collaborators in 2006; contributed to the understanding of endoPG genotypic diversity by locating and obtaining leaf samples for DNA extraction, performing and training others in DNA extraction, PCR, and gel running. **Jose Soto** (technical, UC Davis) – contributed to the understanding of the diversity and function of PCR alleles by performing softening measurements, obtaining leaf samples for DNA extraction, and performing DNA extraction, PCR, and gel running. **Robbie James** (technical, UC Davis) – contributed to the understanding of the functional role of endoPG alleles by performing softening measurements, phenotypic data entry and preliminary analysis. **Michael Pitchford** (technical, UC Davis) – contributed to the understanding of the functional role of endoPG alleles by performing softening measurements. **Rebecca Cooper** (technical) – contributed to the understanding of the functional role of endoPG alleles by locating and obtaining fruit samples and performing softening measurements. **Andrea Gamberini** (graduate, visiting PhD student from University of Bologna, Italy) – contributed to the understanding of the endoPG genotypic diversity through candidate gene nomination and primer design, PCR, and gel running. **George Manganaris** (postdoctoral, UC Davis) – contributed to the understanding of the functional role of endoPG alleles by performing endoPG enzyme activity measurements.

Ebenezer Ogundiwin (postdoctoral, UC Davis) – contributed to the understanding of allelic diversity and function of endoPG and other candidate genes, by performing bin-mapping PCR and gels, supervising the molecular lab, and helping with manuscript and poster preparation. **Cameron Peace** (postdoctoral, UC Davis) – contributed to all aspects of the project by managing the project as PI, initiating and maintaining various collaborations, candidate gene nomination and primer design, PCR and gel running for endoPG genotypic diversity assessment, endoPG sequence analysis, poster preparation, presentation preparation and delivery, and manuscript preparation

UC Davis, Davis campus
Eric Wada (graduate, UC Davis) – contributed to the understanding of the endoPG genotypic diversity and phylogeny by training and supervising three undergraduate students. **Hoang Yen Nguyen**, **Paulina Walichiewicz**, and **Fei Yian Yoong** (undergraduate, UC Davis) – contributed to the understanding of the endoPG genotypic diversity and phylogeny by cloning and sequencing of endoPG PCR products, and editing and analysis of sequences.

USDA-ARS, Appalachian Fruit Research Station
Anthony Rugh (technical, AFRS) – contributed to the understanding of the physical organization of the endoPG locus by BAC mapping, BAC subcloning, and sequencing of peach. **Jessica Lauffer** (undergraduate, University of Washington) – contributed to the physical organization of the endoPG locus through PCR of several cultivars of peach to determine primers that could distinguish functional alleles. **Kenneth Smith** (undergraduate, West Virginia University) – contributed to the physical organization of the endoPG locus by BAC mapping of plum.

Collaborations:

- Interaction between all four PDs is new – Peace, Crisosto, and Gradziel were already collaborating on other projects, but formal collaboration with Callahan was enabled with this project
- Project support enabled the originally planned collaborations with Marisa Badenes (IVIA, Spain), Amy Iezzoni (Michigan State University, USA), Dan Potter, David Slaughter, and John Labavitch (UC Davis, USA)
- Tatyana Zhebentyayeva and Bert Abbott (Clemson University, USA) for peach endoPG locus physical isolation and sequencing
- Laura Georgi (Clemson University, USA) for endoPG locus physical isolation from plum
- Mallikarjuna Aradhya, Clay Weeks, and Ed Stover (National Clonal Germplasm Repository at Davis, USA) for endoPG diversity assessment in the Davis *Prunus* germplasm collection
- Closer interaction with the California stone fruit industry for obtaining fruit samples for firmness measurements and leaves for DNA extraction
- Dick Okie (USDA-ARS, Georgia) for endoPG and *Hd* analysis in additional peach germplasm

- John Clark (University of Arkansas, USA) for endoPG and other softening candidate gene analysis in additional peach germplasm
- Jim Olmstead and Matt Whiting (Washington State University, USA) for candidate gene testing of sweet cherry germplasm
- Werner Howad and Pere Arus (IRTA, Spain) for details and DNA for the bin-mapping of candidate genes
- Rozemarijn Dreesen (KU Leuven, Belgium), Fabrizio Costa (University of Bologna, Italy), and Pere Arus (IRTA, Spain) for apple-peach comparative mapping
- Maria Teresa Dettori, Ignazio Verde, and Elisa Vendramin (ISF, Italy) for verification of endoPG co-segregation with *F-M* in other *Prunus* populations, elucidation of *F-M* interaction with *Hd*, and candidate gene analysis and mapping of *Hd*
- Elisabeth Dirlewanger (INRA, France) for verification of endoPG co-segregation with *F-M* in other *Prunus* populations and mapping of further candidate genes
- More than a dozen institutions in the US and Europe for a comparative mapping consortium for the endoPG region in *Prunus* and apple
- Takashi Haji (NIFTS, Japan) for elucidation of *F-M* interaction with *Hd*, and candidate gene analysis and mapping of *Hd*
- Involvement in RosPOP – PD Cameron Peace is co-Liaison with Werner Howad (IRTA, Spain) for RosPOP-*Prunus*
- Hosted Andrea Gamberini, a PhD student from the University of Bologna, Italy, for screening endoPG and other candidate genes across additional germplasm

Project title: Genetic Diversity of Wild Apple Accessions in the National Plant Germplasm System

PI: Gayle M. Volk

Co-PIs: Christopher M. Richards

Accomplishments:

1) Quantify and apportion genetic diversity in collections of wild *Malus*.

Manuscript is near completion entitled "Genetic diversity and population structure in *Malus sieversii*; the wild progenitor species of domesticated apple". It will be submitted by 12/06.

Data has been collected and analyses are underway for genetic diversity analyses of *Malus orientalis* and Chinese *Malus* species.

Additional progress: data collection is underway to genetically characterize >500 European hard cider varieties and identify their genetic relationships with the wild *Malus* species we have characterized.

Identify a group of individuals that represents the overall genetic diversity of wild *Malus* collections with the fewest number of individuals.

Manuscript is in preparation that will describe core subsets of *Malus sieversii* in Kazakhstan.

Additional manuscripts (2007) will describe core subgroups of *Malus orientalis* and Chinese species.

Broad Impacts:

In 2006, we spearheaded an effort to develop the framework to put our genetic SSR data into the GRIN database. As we started looking at the GDR and GRIN databases, we realized that neither was suited to the diversity data that our laboratory is generating for Rosaceae species. The "molecular data in GRIN" framework is currently being programmed into the database. When we have submitted the diversity manuscript for Malus sieversii, this data will be deposited into both GDR and GRIN databases.

When the PI attended the International Rosaceae Genomics Conference in NZ, contacts were made with Europeans working on the HiDRAS project. As a result of those contacts, US programs are now using the same *Malus* DNA standards as the European groups. The accessions chosen as standards in the HiDRAS project were not even available through the NPGS. Efforts are underway to add those accessions to the NPGS so they are available to US researchers. We have also adopted the use of a set of

European SSR primer sets (in addition to a set of US SSR primers) so the genetic diversity data can be compared across continents.

Deliverables:

Publications

Genetic diversity and population structure in *Malus sieversii*; the wild progenitor species of domesticated apple. To be submitted to Theoretical and Applied Genetics 12/06

Presentations

Volk, GM. Characterization of wild Malus populations using genotypic and phenotypic traits. 3rd International Rosaceae Genomics Conference, NZ, 19-22 March 2006.

Genetic diversity and population structure in Malus sieversii, a wild progenitor species of the cultivated apple. Seminar, Horticulture Dept., Cornell University Sept. 12, 2006

Meeting Abstracts

John L. Norelli[1], Schuyler S. Korban[2], Gayle M. Volk[3], Zhao-Bang Zeng[4], Herb S. Aldwinckle[5], Carole L. Bassett[1], Jonathan Beever[2], Robert E. Farrell, Jr.[6], Ksenija Gasic[2], Yuepeng Han[2], Sunnee Kertbundit[2], Brandy Marron[2] and Christopher M. Richards[3:] PAGXIV 2006 USDA-CSREES-NRI Projects Developing Genomic Resources for the Rosaceae (*Malus*)

Christopher M. Richards, Adam D. Henk, Ann A. Reilley, Philip L. Forsline, and Gayle M. Volk 2006 PAG XIV Patterns of Molecular and Quantitative Trait Variation in *Malus sieversii* from Kazakhstan

Volk, GM, CM Richards, PL Forsline. Characterization of wild Malus populations using genotypic and phenotypic traits. 3rd International Rosaceae Genomics Conference, NZ, 19-22 March 2006.

Web sites none

Community resources generated (sequences, populations, plant materials)

SSR data for wild Malus species to be deposited into GDR and GRIN starting in12/07

Patents none

Training:

Adam Henk, Ann Reilley, Ann Caspersen, Technical personnel training: SSR data collection

Collaborations:

Jimmie Mowder, GRIN, USDA-ARS
Quinn Sinnott, GRIN, USDA-ARS
Ian Merwin, Cornell University
Herb Aldwinckle, Cornell University
Gennaro Fazio, USDA-ARS

Santiago Perira Lorenzo, Spain
Kanin Routson, Graduate Student, U. Arizona
Luca Gianfranceschi, Italy
Andrea Patocchi, Switzerland
Eric van de Weg, The Netherlands
Charles Eric Durel, INRA, France
Sook Jung, GDR, WSU

Project title: Functional Genomic Response of Apple to Fire Blight

PI: Jay Norelli

Co-PIs: Herb Aldwinckle, Carole Bassett, Bob Farrell

Accomplishments:

- cDNA suppression subtractive hybridization (SSH) was applied to study the temporal progression of gene expression in susceptible ('Gala') apple in response to infection. In the early hours post-infection (1h, 2h and 24 h), host genes associated with photosynthesis and signaling, as well as a few host defense-related genes, were down-regulated. Some PR genes were up-regulated late (48h and 72 h) post-infection, including chitinase and Mal d1.
- In a second study, SSH and cDNA-AFLP analyses were used to identify genes expressed in resistant (Geneva 41) and susceptible (M.26) apple in response to infection. Two cDNA-AFLP kits (Licor and Infobiogen), which employ different restriction enzymes, were used in order to maximize discovery of EST's.
- A multi-plasmid transformation approach has been evaluated for the high-throughput generation of RNAi mutants in apple. M.26 apple tissue was transformed with three single pHellsgate8-derived plasmids and a mixture of all three. Transformation frequency was not reduced by use of multiple silencing plasmids. Universal PCR primers were developed for pHellsgate8-derived plasmids that can: 1) detect the presence of single or multiple EST silencing insertions in the RNAi transgenic, and 2) provide sequencing template to determine the EST contained in the silencing insertion.
- pHellsgate8-derived silencing plasmids were constructed for 19 *Malus* ESTs identified as most likely to be associated with response to infection.

Broad Impacts:

The genomics research undertaken in this project will elucidate the poorly understood mechanisms responsible for the resistance and susceptibility of apple and pear to fire blight disease and, thereby identify new opportunities for improving natural resistance. The research project is identifying apple genes responsible for resistance to fire blight disease and is developing technologies to determine the biological function of those genes. This will facilitate new methods of marker-assisted selection to efficiently breed and/or genetically engineer superior apple varieties with improved fire blight resistance resulting in: 1) reduced dependence on synthetic pesticides, 2) greater stability in production levels over time due to enhanced ability to better tolerate fire blight, and 3) improved competitiveness of the US fruit industry in the global market due to increased tree fruit productivity and reduced inputs.

Deliverables:

Publications

- Norelli, J. L., R. E. Farrell Jr., C. L. Bassett, A. M. Baldo, H. S. Aldwinckle, M. E. Wisniewski. 2007. Rapid transcriptional response of apple to fire blight disease revealed by cDNA suppression subtractive hybridization analysis. *In preparation*

Presentations

- Plant and Animal Genome (PAG) XV Conference, January 13-17, 2007:
  1. Functional Genomic Response of Apple to Fire Blight. Norelli JL, Bassett CL, Farrell Jr. RE, Baldo AM, Malnoy M, Borejsza-Wysocka EE, Lalli DA, Korban SS, Ksenija Gasic K, Wisniewski ME, Aldwinckle HS
  2. Identification of Genes Expressed in Apple (*Malus* X *domestica*) in Response to *Erwinia amylovora* Infection. Farrell Jr RE, Bassett CL, Norelli JL, Baldo AM, Aldwinckle HS, Wisniewski ME
  3. Differential Apple Gene Expression During Interactions of Resistant and Susceptible Apple Cultivars with the Necrogenic Bacterium *Erwinia amylovora.* Malnoy M, Borejsza-Wysocka EE, Farrell Jr. RE, Norelli JL, Bassett CL, Aldwinckle, HS
  4. A Multi-Plasmid Transformation Approach for the High-Throughput Generation of RNAi Mutants of Apple. Borejsza-Wysocka EE, Norelli JL, Malnoy MM, Farrell Jr. RE, Basset CL, Aldwinckle HS
  5. Computational Identification of Candidate Genes Involved in Response to Fire Blight in Apples. Baldo AM, Bassett CL, Malnoy M, Korban SS, Farrell Jr RE, Aldwinckle HS, Wisniewski ME, Norelli JL
- PAG XIV Conference January 14-18, 2006: Temporal Response of Apple (*Malus*) to Fire Blight Disease. Norelli JL, Farrell Jr. RE, Bassett CL, Baldo AM, Aldwinckle HS, Wisniewski ME
- XII International Congress on Molecular Plant-Microbe Interactions, December 14-19, 2006, Merida, Mexico: Rapid Genomic Response of *Malus* to Infection by *Erwinia amylovora*. Norelli JL, Bassett CL, Farrell Jr. RE, Wisniewski ME
- American Phytopathological Society, July29-August 2, 2006, Quebec City, Canada: Using Functional and Applied Genomics to Identify Genes that Confer either Resistance or Susceptibility to Fire Blight. Norelli JL, Farrell Jr. RE, Leder EH, Bassett CL, Baldo AM, Malnoy M, Borejsza-Wysocka EE, Aldwinckle HS, Ksenija Gasic K, Korban SS, Wisniewski ME
- Gordon Research Conference - Plant Molecular Biology, July 17, 2006, Plymouth, NH: Rapid genomic response of *Malus* to infection by *Erwinia amylovora.* Farrell Jr. RE, Norelli JL, Bassett CL, Wisniewski ME
- Mid Atlantic Plant Molecular Biology Society (MAPMIBS) August 11-12, 2005, Laurel, MD: Rapid genomic response of *Malus* to infection by *Erwinia amylovora.* Farrell Jr. RE, Norelli JL, Bassett CL, Wisniewski ME
- MAPMIBS August 17-18, 2006, Laurel, MD: Transcriptional Response in Apple to Fire Blight Disease. Farrell Jr. RE, Norelli JL, Bassett CL, Baldo AM, Aldwinckle HS, Wisniewski ME

Meeting Abstracts

- Norelli, J. L., R. E. Farrell, E. H. Leder, C. L. Bassett, A. M. Baldo, M. Malnoy, E. Borejsza-Wysocka, H. S. Aldwinckle, K. Gasic, S. S. Korban, M. E. Wisniewski. 2006. Using Functional and Applied Genomics to Identify Genes that Confer either Resistance or Susceptibility to Fire Blight. Phytopathology 96:S86 (Abstract).

Web sites: none

Community resources generated (sequences, populations, plant materials):

- Approximately 450 fire blight-associated Malus ESTs were deposited into GenBank, 60% of the ESTs were not previously associated with fire blight resistance and 5 are new Malus ESTs.

Patents: none

Training:

Undergraduates:

- Greg Richart, current undergraduate, Pennsylvania State University (PSU), York: contributed to SSH analysis.
- John McGraw, former undergraduate, PSU, York: contributed to the bioinformatics analysis of SSH ESTs.

Postdoctorals:

- Dr. Zuping Yang, USDA-ARS, Kearneysville, WV; October, 2005 – January, 2006: contributed to the cloning of fire blight-associated *Malus* ESTs deposited into GenBank; Dr. Yang had to terminate his appointment due to personal family circumstances.
- Dr. Erica H. Leder, USDA-ARS, Kearneysville, WV; February, 2006 – April, 2006: contributed to the construction of pHellsgate8-derived silencing plasmids for 19 *Malus* ESTs identified as highly probable to be associated with response to infection; worked on a temporary basis to continue progress on the project while a permanent replacement for Dr. Yang was located.
- Dr. Donna A. Lalli, USDA-ARS, Kearneysville, WV; August 2006 – present: contributed to deposit of fire blight-associated *Malus* ESTs in GenBank.

Collaborations:

- Collaboration between the Project Co-Directors has increased. The Cornell lab of Herb Aldwinckle now has new interactions with the labs of Carole Bassett at USDA-ARS Kearneysville WV and of Bob Farrell, Pennsylvania State University, York, PA. Interactions with the lab of Jay Norelli USDA-ARS Kearneysville WV have increased significantly.
- Collaboration was established with the ongoing *Malus* microarray project of Tim McNellis and Philip Jensen, Pennsylvania State Univ., University Park, PA and the ongoing EST project of Schuyler Korban and Ksenija Gasic, Univ. of Illinois, Urbana, IL.
- Established a collaboration with Angela Baldo, USDA-ARS in bioinformatics research to: 1) custom-modify vector-screening tools to identify EST contamination, and 2) develop computational methods to identify apple genes expressed uniquely in disease-challenged tissue or that are similar to *Arabidopsis* ESTs associated with *Pseudomonas syringae* pv. *tomato* infection.

- Established collaboration with Gennaro Fazio, USDA-ARS to facilitate vertical integration between NRI project and ARS/Cornell apple rootstock breeding program.
- Established Non Funded Cooperative Agreement among USDA-ARS, Cornell University and HortResearch, New Zealand to identify molecular markers for fire blight ESTs and determine if the markers are associated with QTLs for fire blight resistance.
- Established collaboration with Rich Jorgensen and Chonglie Ma, Univ. of Arizona, Tucson to develop sense-RNAi for functional genomic analysis in *Malus*. Rich Jorgensen has an current NSF project titled 'Functional Genomics by Sense-RNAi: A Forward Genetic Approach for Cell-Type-Targeted Mutagenesis and for Polyploids' (# 05018240).
- Bob Farrell obtained research support from NutriCore, N.E. to facilitate the study of the molecular biology of fire blight disease in apple. $10,000.00 received in 2006.

Project title: High-Resolution Physical Mapping of the Apple Genome by BAC Fingerprinting

PI: Schuyler S. Korban

Co-PIs: Jonathan E. Beever

Accomplishments:

1. Constructed a genome-wide framework for a physical map of the apple genome by BAC fingerprinting.
2. A total of 82,503 BAC clones (average insert size of 110 kb), derived from the two complementary BAC libraries (a *Bam*HI library and a *Hind*III library), were fingerprinted.
3. A total of 74,281 clones, representing ~10.5× haploid genome equivalents, were assembled into contigs. Following automatic assembly, 68,058 BAC clones (92%) were assembled into 3,943 contigs. Then, this was followed by hand assembly, thus collapsing the number of contigs to 2,702.
4. Preliminary efforts to develop EST-SSR molecular markers that will be used to construct/saturate the apple genetic map have indicated that from a total of 122 unique apple sequences, 28 were found to produce polymorphic profiles. This paves the way in identifying a large number of polymorphic markers (1,000) for saturating the apple genetic map.

Broad Impacts:

The apple is highly self-incompatible thus contributing to its high level of heterozygosity. In our project, more than 92% of the apple BAC clones have been successfully assembled into contigs, with an average estimated size of ~340 kb in physical length. This suggests that heterozygosity has little influence on contig assembly of the apple genome.

Deliverables:

Publications

Han, Y., K. Gasic, B. Marron, J.E. Beever, and S.S. Korban. 2007. A BAC-based physical map of the apple genome. *Genomics* (in press).

Naik, S., C. Hampson, K. Gasic, G. Bakkeren, and S.S. Korban. 2006. Development and linkage mapping of E-STS and RGAs for functional gene homologues in apple. *Genome* 49: 959-968.

Presentations

1. A presentation entitled “Apple genomics research”, University of Illinois, Urbana, IL, Oct. 5, 2006.
2. A seminar on “New resources for apple genomics” at the Katholike University of Leuven, Leuven, Belgium, July 3, 2006.

3. A seminar on "Advances in apple genomics resources" at the Universidade Federal de Santa Catarina, Santa Catarina, Brazil, April 12, 2006.
4. A presentation on "Developing a physical map for the apple genome" - 3rd International Rosaceae Genomics Conference, Napier, Hawkes Bay, New Zealand, March 21, 2006.

Meeting Abstracts

1. Han, Y., K. Gasic, S. Kertbundit, B. Marron, J.E. Beever, and S.S. Korban. 2007. A first draft of a genome-wide physical map for the apple. Plant & Animal Genome Conf. XV Abstracts (in press).
2. Han, Y., K. Gasic, S. Kertbundit, B. Marron, J.E. Beever, and S.S. Korban. 2006. Development of a genome-wide physical map for the apple. 3rd International Rosaceae Genomics Conference Abstracts, p. 52.
3. Han, Y., K. Gasic, S. Kertbundit, B. Marron, J.E. Beever, and S.S. Korban. 2006. Developing a genome-wide physical map of the apple genome. Plant & Animal Genome Conf. XIV Abstracts –P-496, p. 225.

Web sites: http://korban.nres.uiuc.edu/

Community resources generated (sequences, populations, plant materials): BAC clones, EST-SSRs, mapping population

Patents: none

Training:

- Yuepeng Han - Postdoctoral - he has been involved in BAC fingerprinting and assembly of BAC clones into contigs, including hand assembly.
- Kesenija Gasic - Postdoctoral - she has been involved in developing EST-SSRs
- Brandy Marron - Research Specialist - she has been involved in BAC fingerprinting
- Mark Danielson - undergraduate student - he has been working as a student helper in preparing media for BAC DNA extractions.
- Suresh Naik - Postdoctoral (with collaborator Dr. Cheryl Hampson) - developing EST-SSRs and assessing their utility in mapping.

Collaborations:

- Interactions with Dr. Cheryl Hampson (Agri-Canada) on EST-SSRs and mapping
- Interactions with HortResearch scientists (Dr. Sue Gardiner) on developing molecular markers
- Interactions with Brazilian scientists (Dr. Adriana Dantas) about molecular markers useful for mapping
- Interactions with various US and European scientists about apple mapping efforts, and possible genome sequencing efforts for comparative genomics.

Project title: Identifying the Genes Associated with Day-Neutrality in Strawberries Using a QTL Approach

PI: J. F. Hancock

Co-PIs: K. Lewers, S. Van Nocker and D. Wang

Accomplishments:

Our first objective was to develop a molecular map and identify QTL associated with day-neutrality in a population of 'Honeoye' × 'Tribute' after phenotyping them in multiple environments. Our intent was to begin mapping with the AFLP markers that we had already developed, but to incorporate SSRs as fast as we could identify them. We have now placed 387 single dose restriction fragments (SDRFs) on a consensus map of 1310.7 cM with 42 linkage groups. All of these are AFLPs, but we have about 60 new SSR markers that are ready to place on the map. Individuals from the mapping population were observed for their flowering habit throughout the growing season in Michigan (MI), Minnesota (MN), Maryland (MD), Oregon (OR) and California (CA). Eight QTL were found that were either location specific or shared among locations. None of these QTL explained more than 36 % of the phenotypic variation, indicating that the inheritance of day-neutrality is likely a polygenic trait. At all three eastern sites, one QTL was identified on Linkage group (LG) 17 that was a strong regulator of day-neutrality; additional QTL were identified in MI on LG 17 and in MN on LGs 7, 20 and 28. In the western states, only one significant QTL was identified on LG 7 which accounted for 22 % of the phenotypic variation in CA. This same peak was uncovered with a P value less than 0.2 in MI.

Our second objective was to develop a molecular map and identify QTL associated with day-neutrality in another large segregating population with the same SD parent ('Honeoye'), but a different DN *F. virginiana* parent (RH 30). We have now generated a family of 167 individuals of this cross and phenotyped them for their photoperiod sensitivity in the greenhouse. Over 80 % of the hybrids appeared to be day-neutral, compared to 50 % of the 'Tribute' x 'Honeoye' population. DNA has been extracted from this new hybrid population and we will begin mapping it this winter. We will also replicate the progeny individuals via runners and set them in the field next spring to further test their photoperiodic responses.

Our final objective was to anchor our two new octoploid linkage maps with existing maps of *F.* ×*ananassa* and diploid *F. vesca*. We have just begun this work with a set of 50 diploid SSR primers. So far, it appears that a high percentage of them will prove useful in mapping the octoploids.

Broad Impacts:

It is interesting that the QTL on LG 17 was so prominent in all three eastern states, but absent in CA and OR. Different loci may regulate day-neutrality in the various areas, as there is a strong temperature/photoperiod interaction that determines flowering in the

strawberry. We speculate that the QTL on LG 17 may be required for floral initiation under the hot summer conditions found in eastern continental climates.

Deliverables:

Publications

C.K. Weebadde, D. Wang, C.E. Finn, K.S. Lewers, J.J. Luby, J. Bushakra, T.M. Sjulin and J.F. Hancock. 200-. Using a linkage mapping approach to identify QTL for day-neutrality in the octoploid strawberry (Fragaria ×ananassa Duch ex Rozier). Plant Breeding (submitted)

Jim Hancock, Cholani Weebadde and Sedat Serçe. 200-. Challenges faced by day-neutral strawberry breeders in the continental climates of the eastern USA and Canada. HortScience (submitted).

Meeting Abstracts

C. K. Weebadde and J. F. Hancock. 2005. QTL associated with day-neutrality in strawberry. HortScience 40:1121

Community resources generated (sequences, populations, plant materials)

Two populations segregating for day-neutrality are available for distribution, 'Tribute' × 'Honeoye' and 'Honeoye' × RH 30 (*F. virginiana*)

Training:

- Cholani Weebadde (PhD. student) developed the linkage map for 'Tribute' × 'Honeoye'.
- Sue Hammer (technician) and Elena Sierra-Ponce (undergraduate assistant) contributed to the map building.
- Kate Rappaport (technician) assembled and processed PCR reactions and extracted sizing data from the ABI 3730 Genetic Analyzer.
- Ernalyn Peralta (biological science aid) assembled and processed PCR reactions and extracted sizing data from the ABI 3730 Genetic Analyzer. Ms. Peralta will be resigning in early December to pursue a career in her chosen field, biological engineering.
- Sheena Mathew (biological science aid), currently attending U. C. Davis as a sophomore, assembled and processed PCR reactions.

Collaborations:

The grant itself initiated a strong collaboration between the co-PIs Kim Lewers and Jim Hancock.

Serious interactions will soon begin with Tom Davis (Univ. New Hampshire) and Kevin Folta (Univ. Florida) to determine if the linkage maps developed in *F. vesca* are transferable to the octoploids.

Project title: Increasing the diversity of EST sequences for *Fragaria*

PI: J. P. Slovin

Co-PIs: P. D. Rabinowicz

Accomplishments:
The cold stressed tissue cDNA library was constructed and 9,600 high quality sequences were produced and deposited to GenBank. An assembly (~2000 contigs and ~ 6,000 singletons) of all *F. vesca* ESTs and other cDNA sequences in GenBank is now available at a TIGR website. Tissue and RNA has been obtained for the remaining libraries and library construction is underway.

Broad Impacts:
About 50% of the existing *Fragaria* ESTs in GenBank were generated by this project so far. Dr. T. Davis has already found these sequences to be helpful in annotation of his genomic sequences. About one quarter of the cold-stress library sequences do not show similarity to a Rosaceae unigene in GDR, and therefore contribute totally new information about genes in this family. Sequences relating to abiotic stress responses in the Rosaceae are under-represented in existing databases, so we our initial data suggests that the sequences from the remaining libraries will also contain new information.

Deliverables:

Publications:
Slovin, J.P., Park, S., Cohen, J.D. Strawberry fruit protein with a novel post-translational indole-acyl modification. Acta Hort. (Submitted)

Presentations/Meeting Abstracts:
Slovin, J.P. and Rabinowicz, P.D. 2005. Increasing the diversity of EST sequences for Fragaria. ASHS, Las Vegas, NV

Slovin, J.P. and Rabinowicz, P.D. 2006. Increasing the diversity of EST sequences for Fragaria. Abstract Guide: Plant and Animal Genome XIV, San Diego, CA

Slovin, J.P., Davis, T., Rabinowicz, P.D., Shulaev, V., Sargent, T. 2006. *Fragaria vesca*, a reference plant for the Rosaceae family. Arthur M. Sackler Colloquia of the National Academy of Sciences, Washington, D.C.

Slovin, J., Park, S. Cohen, J. 2006. Strawberry fruit protein with a novel post-translational indole-acyl modification. International Horticultural Congress, Seoul, Korea

Slovin, J.P. and Rabinowicz, P.D. 2007. Increasing the diversity of EST sequences for Fragaria. Abstract Guide: Plant and Animal Genome XV, San Diego, CA

Web sites:
TIGR Transcript Assemblies http://plantta.tigr.org

Community resources generated (sequences, populations, plant materials): 9,600 *Fragaria vesca* ESTs. Accession # DY666641-DY676279

Patents: none

Training:
Undergraduate student: Matthew Barton, student at University of Maryland, learned aseptic technique and assisted in growing, treating and harvesting plant tissues for the cDNA libraries. Mr. Barton also learned PCR techniques for developing gene specific markers for several of the sequences in the library.

Collaborations:
- Dr. Zhongchi Liu, University of Maryland, co-wrote an NRI grant proposal with Dr. Slovin to examine the heat stress response in reproductive structures using diploid *Fragaria*.
- Dr. Shunyuan Xiao, University of Maryland has established a collaboration with Dr. Slovin to determine whether RPW8, an Arabidopsis gene involved in powdery mildew resistance, is effective in *Fragaria vesca.*
- Dr. Zoran Ristic, University of Kansas, is collaborating with Dr. Slovin to study EF-tu expression in response to elevated temperatures in *Fragaria.*
- Dr. Kevin Folta, University of Florida, collaborated with Dr. Slovin and Dr. Rabinowicz in construction of the cold stress cDNA library.
- Dr. Agnes Chan, TIGR, collaborated on analysis of the ESTs.

Project title: Gene Pair Haplotypes and Sequence Samples from Strawberry (Rosaceae): Multi-Purpose, Transferable Resources for Genomics and Variety Improvement.

PI: Tom Davis

Co-PIs: Kevin Folta, Phillip SanMiguel (Cooperator: Jeff Bennetzen).

Accomplishments:

We have completed the sequencing and preliminary contig assembly for 50 fosmid inserts from a genomic library of the strawberry diploid model species *Fragaria vesca* (ssp. *americana* var. Pawtuckaway). Based upon sequence assembly lengths, the fosmid inserts ranged in size from about 28 kb to 45 kb, with an average insert size of about 35 kb. A cumulative total of about 1.75 Mb of genomic sequence was obtained (slightly under 1% of the *F. vesca* genome). Of the 50 sequenced fosmid clones, 20 were selected via library probing on the basis of targeted gene content, while 30 were selected at random (after pre-screening to exclude organellar inserts). The targeted clones were selected using probes for the following genes: (flowering-related) CONSTANS, Suppressor of CONSTANS, Leafy, PhyA, Hy5, Pistillata, Apetala3; (fruit color and flavor) Chalcone Synthase, Chalcone Isomerase, Dihydrofolate Reductase, RAN (Regulator of Anthocyanin Synthesis), Linalool Synthase (two clones); (other metabolic) Alcohol Dehydrogenase, Granule-Bound Starch Synthase I; (disease resistance) gRGA1, gRGA2, gRGA3, cRGA1, Ve-like. We are currently using a variety of markers, including SSRs that occur at a frequency of about 1 per 6 kb, and gene pair markers (see below), to place all 50 fosmid clones onto the diploid strawberry reference map.

Among the 20 gene-targeted fosmids, gene density is about 1 protein-encoding gene per 6 kb. Concomitantly, intergenic (gene-to-gene) distances average about 3-4 kb. Based on the obtained sequences, we have begun developing novel "gene pair" markers and testing their utilities in diploid and octoploid strawberries, and in various other Rosaceae species. Here, the term "gene pair" refers to two immediately adjacent genes. The gene pair marker concept is a novel development and important deliverable of the present grant. The citing of a "forward" PCR primer in a proximal exon in one member of the gene pair, and the "reverse" primer in a proximal exon of the second member of the pair, allows amplification of the "gene pair interval", which includes the intergenic region as well as any exon (and intron) sequences flanked by the PCR primers. Because the primers are cited in conserved exon sequences, and because considerable conservation of micro-colinearity is expected among rosaceous genomes, gene pair markers have the potential to be highly transferable among rosaceous genera.

Using DNA samples from various rosaceous species provided by thirteen collaborating investigators (see collaborator list, below), we have begun to assess gene pair marker transferability in the Rosaceae. An initially developed gene pair locus based on the tandemly duplicated chalcone synthase gene can be amplified in strawberry, cherry, peach, rose, and *Rubus*, but was problematic in apple. A second gene pair locus, provisionally termed CDPK-BHLH, has been amplified in strawberry, cherry, rose, and *Rubus*, but not apple. We anticipate that gene pair sites can be established from each

fosmid sequence, providing the potential for development of up to 50 gene pair anchor loci for comparative mapping among rosaceous species.

An initial gene pair locus, at site RGA-subtilase, has been examined in depth in *Fragaria*. Three distinct haplotypes have been defined at this locus in octoploid *Fragaria* species, and results indicate that it will be possible to differentiate multiple haplotype forms in individual octoploid varieties, supporting the expectation that gene pair haplotypes can be used as an effective mapping tool in the octoploid, cultivated strawberry. Five additional gene pairs have been amplified and analyzed from the octoploid cultivar Strawberry Festival. Several of these are related to genes associated with photoperiodic flowering and constitute the basis for functional studies around the trait. The variability identified within and between genomes is being compared to diploid species to test subgenome origin and provide a basis for mapping in the octoploid species. The same gene pairs have been assigned to linkage groups in the East Malling population using gene-pair spanning polymorphisms. Several more have been analyzed for variability and stand to be mapped in the near future.

Broad Impacts:

This project has leveraged the complementary strengths and backgrounds of the Co-PIs (Tom Davis and Kevin Folta) and their collaborator (Jeff Bennetzen) in an integrated effort to advance the state of *Fragaria* and Rosaceae genomics and the understanding of plant genome composition in general. The resulting genomics resources will be of both immediate and long term value relevant to breeding and marker assisted selection for strawberry variety improvement. A welcome and only partially anticipated outcome of the project is the extent to which genomic resources developed for strawberry - specifically genomic DNA sequence data and concomitant development of novel "gene pair markers" - will provide a comparative mapping resource that is proving highly transferable to other rosaceous species. The potential appeal of this resource has attracted cooperators from six different countries representing almost all of the major Rosaceae crop species, thereby strengthening Rosaceae community interaction and commonality of interest and effort to the benefit of all.

The work demanded improvement of available DNA extraction protocols. Optimization and modification of existing protocols generated new means of isolating materials suitable for most downstream applications, with greater yield and quality. Additionally, a simple yet powerful gene-pair analysis software suite is in development. The program predicts potential colinearity between an unknown genome (in this case *Fragaria*) and a known genome (Arabidopsis) using EST sequences. The program will speed further predictions of possible relationships between species and allow a basis for wet-lab testing for gene-pair variability.

Deliverables:

Publications

Davis TM , Denoyes-Rothan B, and Lerceteau-Kohler E. 2006. Strawberry. In: Kole C (ed) Genome Mapping & Molecular Breeding in Plants IV: Fruits and Nuts. Springer, Heidelberg, Berlin, New York, Tokyo.

Sargent, D.J., J. Clarke, D.W. Simpson, K.R. Tobutt, P. Arus, A. Monfort, S. Vilanova, B. Denoyes-Rothan, M. Rousseau, K.M. Folta, N.V. Bassil, N.H. Battey 2006. An enhanced microsatellite map of diploid *Fragaria*. *Theoret Appl Genet* 112:1349-59.

Bassil, N.H., Dunn, M., Folta, K.M. and Lewers, K 2006. Microsatellite markers for *Fragaria* from 'Strawberry Festival' ESTs. *Mol Ecol Notes,* 6:473-476.

Folta KM and Davis TM. 2006. Strawberry genes and genomes. Crit. Rev. Plant Sci. 25:399-415.

Folta KM and Dhingra A 2006. Transformation of Strawberry: The basis for translational genomics in the Rosaceae. In Vitro Developmental and Cellular Biology-Plant (in press).

Presentations

Oral

Davis TM. 2005. A diploid platform for strawberry genomics. Plant and Animal Genome XIII. San Diego, CA, January 15-19 (Fruit and Nut Workshop).

Davis TM. Folta KM, Bennetzen J, and SanMiguel P. 2005. Gene Pair Haplotypes and Sequence Samples in Strawberry. American Society for Horticultural Science, Annual Conference.

Tombolato, DCM. 2005. Determining linkage associations in strawberry using "gene-pair haplotypes" Florida Plant Molecular Biology Workshop, Daytona Beach FL, May 6-8.

Folta KM. 2005. Genomics Research and Florida Strawberries, Florida Strawberry Growers Agritech Conference, Plant City, FL, August 16.

Davis TM, Folta KM, Bennetzen JL, SanMiguel P, Slovin J, Rabinowicz PD, Hancock JF, Lewers KS, van Nocker S, Wang D, Ashman T-L, Main D, Staton M. 2006. Functional and structural genomic resource development, and QTL mapping of day-neutrality and sex-determination, in strawberry (*Fragaria*). International Plant & Animal Genome Conference XIV, January 14-18, San Diego, CA. Fruit and Nut Workshop

Davis TM. 2006. Sequence Samples and Gene Pair Haplotypes in Strawberry. 3rd International Rosaceae Genomics Conference, Napier, New Zealand, March 19-22.

Folta KM. 2006. Photoperiodic Flowering Control in Strawberry, Tropical Research and Education Center, University of Florida, Homestead, FL, April 6.

Folta KM. 2006. Novel Approaches to Plant Signal Transduction, Developmental Transitions, Structural Genomics and Phylogenetic Analyses. University of Illinois, Chicago, April 19-20.

Tombolato, DCM. 2006. Gene-pair haplotypes: novel, complex markers for linkage mapping in wild diploid and cultivated octoploid strawberries. Florida Plant Molecular Biology Workshop, Jacksonville Beach FL, May 9-11.

Folta KM 2007. Transformation of strawberry- the basis of translational genomics in *Fragaria*. International Plant & Animal Genome Conference XIV, January 13- 17, San Diego, CA.

Presentations - Posters

Brese RL and Davis TM. 2005. Construction and analysis of a diploid strawberry (*Fragaria vesca*) cDNA library from developing flower buds. Plant and Animal Genome XIII. San Diego, CA, January 15-19 (poster # 14).

Shields M, Zhang Q, and Davis TM. 2005. Large-insert library resources for strawberry (*Fragaria*). Plant & Animal Genome XIII. San Diego, CA, January 15-19 (poster #50)

Tombolato DCM, Davis TM, Folta KM. 2005. Gene-pair haplotypes: novel, complex markers for linkage mapping in octoploid strawberry. Florida Genetics Symposium, University of Florida, Gainesville FL, November 30-December 1. (Poster #15)

Davis TM, Shields ME, Zhang Q, Bennetzen JL, Pontaroli AC, SanMiguel P, Folta KM, Tombolato DCM. 2006. Sequence samples from the diploid strawberry model species, *Fragaria vesca*.). International Plant & Animal Genome Conference XIV, January 14-18, San Diego, CA.

Dhingra A, Stewart PJ, Tombolato DCM, Madzima TF, Colquhoun T, Howard L, Folta KM. 2006. *Fragaria* Genomics at the University of Florida. 3rd International Rosaceae Conference, Napier, New Zealand, March 19-22.

Davis TM, Shields ME, Zhang Q, Bennetzen JL, Pontaroli AC, SanMiguel P, Tombolato DCM, Folta KM. 2007. Sequence samples and gene pair haplotypes in strawberry. 3rd International Rosaceae Conference, Napier, New Zealand, March 19-22.

Stewart PJ, Winslow AR , Folta KM. 2006. The photoperiod pathway in *Fragaria* (Rosaceae). American Society of Plant Biologists, Boston MA, August 3-7.

Stewart, PJ, Winslow AR, Folta KM. 2006. The photoperiod pathway in *Fragaria* (Rosaceae). American Society for Horticultural Sciences Annual Conference, New Orleans, LA, July 21-24.

Tombolato DCM**,** Shields ME**,** Zhang Q**,** Bennetzen JL, SanMiguel P**,** Folta KM**,** Davis TM. 2007. A Diploid Platform For Strawberry Genomics – II. International Plant & Animal Genome Conference XIV, January 13-17, San Diego, CA.

Davis TM, Shields ME, Zhang Q, Tombolato DCM, Folta KM. 2007. Gene Pair Markers: An Innovative Tool For Comparative Linkage Mapping In The Rosaceae Family And In Other Taxa With Small Genomes. International Plant & Animal Genome Conference XIV, January 13-17, San Diego, CA.

Web sites

Davis: http://www.strawberrygenes.org
Folta: http://arabidopsisthaliana.com/strawberry/

Community resources generated (sequences, populations, plant materials)

Large insert (fosmid) genomic library (and library filter sets) of the strawberry diploid model species, *Fragaria vesca*.

1.75 Mb of genomic sequence (~1% of basic strawberry genome).
- includes complete genomic sequences of about 300 protein-encoding genes, including 20 targeted genes associated with traits of economic interest.
- includes about 300 new SSR loci.

F. vesca cDNA library from unopened flower buds.

~2700 total EST sequences including ~1900 unigenes, deposited to GenBank (accession numbers DV438013 - DV440729).

F. vesca linkage mapping population: F2 generation 'Yellow Wonder' x 'Pawtuckaway'.
F. virginiana mapping population: F2 generation of L1 x BC6.
F. ×ananassa cDNA libraries constructed from developing flowers, developing/ripening fruits and roots, in Gateway vectors. Libraries are being sequenced with grower-organization funding.
The COMPAIR software will be made available to the public via a web-based interface in Q1 of 2007.

Patents: None

Training - Undergraduate research projects

Davis lab:

- Poulsen E. 2005. Genome composition and evolution in *Fragaria*. UNH Undergraduate Research Conference. April 30, 2005. Poster # 24. Supported by SURF award, 2004.
- Orcheski, Ben. The genetic determination of sex expression in strawberry. Supported by SURF award, 2005.
- Danton, Ben. Witnessing Evolution: Cellular Organelle and Nuclear Genomic Incompatibilities as a Source of Reproductive Isolation and Divergence of Species. Supported by SURF award, 2006.

Folta lab:

- Jones, Mark. Analysis of intergenic variation in genes relevant to agriculturally-important traits in cultivated strawberry (*F. ×ananassa)*
- Frantz, Amanda. Expanding efficient selection of the diploid strawberry accession Hawaii-4
- Steeves, Cody. High-throughput identification, cloning, and transformation of abundantly expressed strawberry genes of unknown function

Training - Graduate thesis projects

Davis lab:

- Brese, RL. 2006. The development and utilization of EST (expressed sequence tag) resources in the diploid strawberry model system. M.S. Thesis. University of New Hampshire, Durham, NH 03824. 95 pp.
- Shields, ME. 2005. Construction and characterization of a large-insert genomic library for *Fragaria* (Rosaceae). M.S. Thesis. University of New Hampshire, Durham, NH 03824. 141 pp.

Folta lab:

- Tombolato, DCM. Structural genomics in *Fragaria* spp.- new methods in DNA extraction, linkage mapping, gene prediction and annotation. Ph.D. Thesis. University of Florida, Gainesville FL (August 2007 anticipated).
- Gonzalez, WF. Development of molecular markers for disease resistance in wild accessions of octoploid strawberry (*F. ×ananassa*). M.S. Thesis. University of Florida, Gainesville, FL. (December 2007 anticipated).

- Stewart, PJ. The *Fragaria* CONSTANS gene family and the regulation of photoperiodic flowering. Ph.D. Thesis. University of Florida, Gainesville FL (May 2007 anticipated).

Training - Post Doctoral

Davis lab:

- Zhang, Qian. Postdoctoral Research Associate. Fosmid library probing, fosmid handling, sequencing, genotyping, sequence analysis and annotation.

Folta lab:

- Dhingra, A. Former Postdoctoral Research Associate. Rapid sequence capture from *Fragaria* plasmids using the ASAP technique.

Bennetzen lab:

- Pontaroli, Ana. Post doctoral training. Subcloning of *Fragaria* fosmids, sequence assembly, annotation and analysis.

Collaborations:

Janet Slovin, USDA, Beltsville. Utilization of EST and genome library resources.
Kim Lewers, USDA, Beltsville. Mapping and marker development.
Vladimir Shulaev, Virginia Bioinformatics Institute. EST sequencing, other.
Dan Sargent, East Malling Research, UK. Marker development, mapping.
Nir Dai, Volcani Institute, Israel. Grant proposal submission.
Bert Abbott, Jahn Davik, Clemson. Strawberry-peach comparative genomics.
Sue Gardiner, HortReseach, NZ. Gene pair marker testing.
Pere Arus, Amparo Monfort, IRTA, Spain. Marker development, mapping.
Jim Hancock, Michigan State University. Anticipated comparative mapping.
Tia-Lynn Ashman, University of Pittsburgh. Anticipated candidate gene testing.
Timo Hytonen, University of Helsinki. Functional analysis of genes associated with photoperiodic flowering.
Fumi Takeda, USDA Kearneysville. Functional analysis of genes associated with photoperiodic flowering.
Natalia Peres, GCREC, University of Florida. Development of molecular markers for disease resistance in wild strawberry accessions.
Gerco Angenent, PRI, Wagenenginen, The Netherlands. Development of transgenic strawberry lines.
Erzebet Kiss, University of Budapest, Development of transgenic strawberry lines
Bobby Phills, Florida A&M, Translation of strawberry markers to red raspberry.

Contributors of Rosaceae DNA samples to Davis lab for use in assessing gene pair marker transferability: 13 investigators, six countries.

Irene Tierney, Scottish Crop Research Institute (*Rubus*)
Dan Sargent, East Malling Research, UK (*Fragaria*)
Amy Iezzoni, Michigan State University (*Prunus* - cherry)
Thomas Debener, University of Hannover, Germany (*Rosa*)

Elisabeth Dirlewanger, INRA, France (*Prunus* - cherry, peach)
Gennaro Fazio, Cornell, Geneva, NY (*Malus* - rootstock)
Schuyler Korban, University of Illinois at Urbana-Champagne (*Malus*)
Fabrice Faucher, INRA, France (*Rosa*)
Cameron Peace, Washington State University (*Prunus* - peach)
Kim Lewers, USDA Fruit Lab, Beltsville (*Rubus*)
Werner Howad/Pere Arus, IRTA, Spain (*Prunus* - peach)
Patrick Lambert, INRA, France (*Prunus*)
David Byrne, Texas A&M University (*Rosa*)

Project Title: Algorithms and Programs for Gene Expression QTL analysis

PI: Zhao-Bang Zeng

The goal of this proposal is to develop statistical methods, bioinformatics tools and computer software to perform gene expression QTL (eQTL) mapping analysis and to interpret the mapping results. Significant progress has been made in several directions.

- **Multiple interval mapping for eQTL (MIM-eQTL)**: We have developed an efficient procedure specifically for gene expression QTL analysis (Zou and Zeng 2006). The method uses our previously developed MIM to search for eQTL and uses the false discovery rate (FDR) (the estimated proportion among the declared eQTL for all expression profiles that are falsely positive) to justify the model selection procedure. In this method, we proceed to scan the genome for one or multiple QTL on each expression trait stepwisely. In each step, we make the decision whether to continue the search for more QTL or stop the process based on a criterion that can be tuned up from FDR calculation. This process is similar to that proposed by Storey et al (2005, PLoS Biology 3:e267), but with a few critical differences or improvements. (1) The search is not restricted to markers, but covers the whole genome in the fashion of interval mapping. This improvement may not be very significant for dense markers, but is a nice generalization and can be important for less dense markers. (2) Our search is not restricted to two steps, which is the case of Storey et al and can proceed to multiple steps (eQTL) as justified. (3) The search in the second and subsequent steps is restricted to those expression traits that the previous search step is significant. This is drastically different from Storey et al that advocates the two-dimensional (or potentially three or higher dimensional search) for all expression traits. In Zou and Zeng (2006), we show that our conditional search is actually much more powerful statistically than Storey et al and we found more eQTL and eQTL epistasis on the same yeast data (Brems and Krugleyak (2005 PNAS 102:1572-1577) that Storey et al also analyzed with the same FDR level.
- **QTL Cartographer for eQTL**: We are currently developing a new package based on QTL Cartographer, that implements the procedures of MIM for eQTL. The package will contain the procedures for data input and editing of large amount of marker data and gene expression data; various data checking and quality control procedures; several methods for eQTL mapping, including our MIM-eQTL. MIM-eQTL includes procedures for threshold estimation based on permutation; stepwise eQTL search based on MIM; FDR calculation for the selected eQTL; and eQTL epistasis estimation. The new package will be released at our QTL Cartographer web site (http://statgen.ncsu.edu/qtlcart/index.php) as soon as possible.
- **eQTL Viewer**: The eQTL analysis will produce a list of eQTL (i.e. genomic regions) for all the typed and analyzed expression traits. The genomic region for each eQTL can be defined by a 1.5 LOD-support interval calculated from MIM-eQTL, and the genes in each region can be listed if the genome is sequenced and annotated. So, essentially the final results of eQTL analysis could be summarized in a gene list for each eQTL that are matched to its expression gene. Then it would be necessary to come up an efficient and informative way to display, annotate and interpret the results. Using the Scalable Vector Graphics (SVG) technology, we have developed a

very informative and useful tool, called eQTL Viewer (http://statgen.ncsu.edu/eQTLViewer/), for displaying eQTL mapping results (Zou, Aylor and Zeng 2006). The tool is a dynamic database of the gene lists with a graphic 2D display with x-axis for the genome location of eQTL genes and y-axis for the genome location of expression genes. Each gene in the database can be linked to the public genome databases. The scalable feature allows us to zoom-in to look at the detail of a particular region and zoom-out to look at the overall patterns.

- **Prioritize candidate genes in eQTL regions**: There are several ways to annotate the information of eQTL mapping results. One way is to numerically prioritize the genes in each eQTL gene list as potential causal genes for the eQTL. Recently, we have attempted to develop a Bayesian algorithm (Zou 2006; W. Zou and Z.-B. Zeng, unpublished) for such a purpose. The algorithm first uses the information from prior studies and annotations on gene relationships (such as GO classifications, KEGG relationships, chip-on-chip study information) and weighs the information to come up a raw gene-pair relationship score matrix. Then the algorithm uses this relationship score matrix as a prior and combines it with the whole gene pair information from the eQTL mapping study in a repeated recursive re-weight scheme to provide the final priority scores. Information can be reinforced particularly for those eQTL that have effects on multiple expression traits. Although this information is only suggestive, it can play a very important role for a variety of applications in advancing testable hypotheses.
- **MT-MIM for eQTL mapping**: We are currently working on adapting our multiple trait multiple interval mapping (MT-MIM) for eQTL analysis (J. Maia and Z.-B. Zeng, unpublished). This is a very complex procedure and can be used for several purposes. First, it can be used to estimate the contribution of individual eQTL to genetic correlations between expression traits. Second, MT-MIM has the potential to further improve the search for eQTL for a pair or multiple expression traits simultaneously. The main research is still in progress.
- **Using the QTL shielding test (QST) to infer genetic pathways**: For the potential pathway inference, we recently worked out a QTL shielding test (QST) that tries to infer whether the relationship between a QTL and multiple (expression and/or trait) phenotypes can be described by a pathway network or a star network (C. Woods and Z.-B. Zeng, unpublished.). The test focuses on a particular QTL at a time, say Q, and two or more expression and trait phenotypes that share the same QTL, say Y1 and Y2, to see whether Y1 can shield the effect of Q on Y2, i.e. whether Y1 is in the pathway from Q to Y2. We studied various statistical issues for the test, and showed that the test works very effectively for one Q and two Y's. Currently we are working on extending the test and analysis to multiple Q's and Y's. This is a very promising approach to robustly infer sub-networks that have relatively strong causal relationshipsand pathway structures for eQTL and their target traits.

We also made progress in developing new theory for modeling QTL with epitasis and linkage disequilibrium (Zeng et al 2005; Wang and Zeng 2006) and a multiple interval mapping method for categorical trait (CT-MIM) (Li and Zeng 2006). CT-MIM has been implemented and released in the current version of Windows QTL Cartographer (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm).

References:

Li, J., S. Wang and Z.-B. Zeng (2006) Multiple interval mapping for ordinal traits. Genetics 173: 1649-1663.

Wang, T., and Z.-B. Zeng (2006) Modeling quantitative trait loci with epistasis and linkage disequilibrium in experimental and natural populations. BMC Genetics 7:9.

Zeng, Z.-B., T. Wang and W. Zou (2005) Modeling quantitative trait loci and interpretation of models. Genetics 169:1711-1725.

Zou, W. 2006. Transcriptional regulatory pattern in yeast revealed through expression quantitative trait locus mapping. PhD Thesis, Bioinformatics Graduate Program, NC State University.

Zou, W., D.L. Aylor., Z.-B. Zeng. 2006. eQTL Viewer: Visualizing genomics control of transcriptome using SVG. BMC Bioinformatics (submitted).

Zou, W., Z.-B. Zeng. 2006. Multiple interval mapping for gene expression QTL analysis. Genetics (submitted).