

## FINAL PROJECT REPORT

**Project Title:** Pear genome project

**PI:** Amit Dhingra  
**Organization:** Washington State University  
**Telephone/email:** 5093353625/adhingra@wsu.edu  
**Address:** 149 Johnson Hall  
**Address 2:** PO Box 646414  
**City:** Pullman  
**State/Zip:** WA 99164

**Cooperators:** Ananth Kalyanaraman, WSU; Todd Einhorn, OSU; Marie-Helene Simard (Pear Breeder at INRA), Yves Lespinasse, Charles-Eric Durel, Elisabeth Chevreau, INRA at Angers, France; Richard Bell, USDA; Kate Evans, WSU; Robert Bogden, Amplicon Express, Riccardo Velasco, IASMA, Italy; Jasper Rees, Univ of Western Cape, SA.

### Other funding sources

**Agency Name:** Andres Bello University – Herman Silva and Lee Meisel  
**Amount awarded:** \$10,000  
**Notes:** Funds to generated additional genome sequence

**Agency Name:** IASMA, Italy  
**Amount awarded:** \$20,000  
**Notes:** Funds being used in Italy to generate additional sequence from BAC DNA library constructed as part of this project.

**Agency Name:** Roche Inc.  
**Amount awarded:** \$30,000  
**Notes:** Funds being used at 454 to generate scaffold DNA libraries and sequencing to enable efficient assembly of the genome.

**Agency Name:** USDA - NRI  
**Amount awarded:** \$ 224,000  
**Notes:** Supplemental funding provided by USDA for scaffold sequencing in apple. The method developed with the apple funds will be utilized for rapid and efficient assembly of the pear genome.

**Total Project Funding:** \$ 57,000

### Budget History:

Item	Year 1:	Year 2:	Year 3:
Equipment			
Supplies	44,000		
Travel			
Miscellaneous	13,000		
<b>Total</b>	<b>57,000</b>		

## ORIGINAL OBJECTIVE

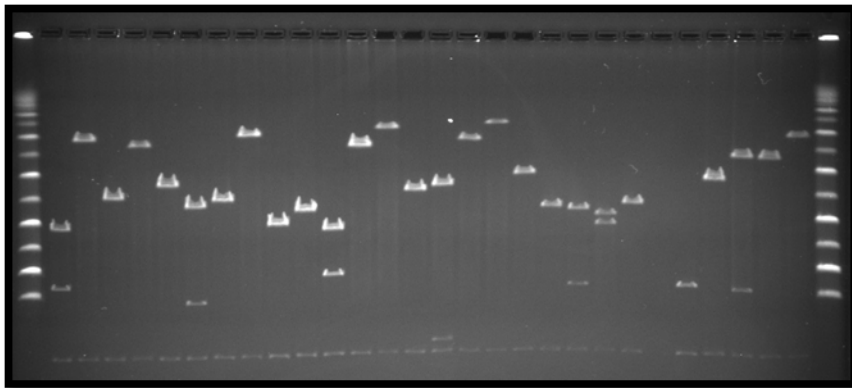
1. Generate sequence information from the double haploid (DH) Comice pear to establish a high quality draft sequence of the pear genome.

## SIGNIFICANT FINDINGS

Definition: BAC Library – It is a method of capturing the large mass of the genome or DNA from any cell and breaking it into small pieces. This needs to be done so that the genome can be studied in manageable pieces.

The basic purpose of this project was to generate a draft assembly of the DH Comice pear genome sequence. Support from this project has allowed us to construct a high quality BAC DNA library that has large pieces of genomic DNA captured in a way that we can multiply them individually. The average size of the genomic DNA in these clones is 145 Kb, which is far above the average size other libraries have (Fig. 1). In comparison, apple BAC library has 130 kb average size of DNA fragment in it.

Figure 1:



**Pear BAC Library. The figure shows DNA as bright bands. Using a measure of size on the sides, we can determine the average size of genome captured in a library.**

Our original plan was to generate sequence information from random DNA pieces derived from the pear genome. Thereafter, we would use the redundancy or sequence similarity at the ends to develop the entire genome.

While the work was in progress, we obtained additional funding from USDA-NRI program to develop a new method that allows us to generate sequence in a way that not only represents the entire genome; it also builds a scaffold simultaneously. This requires a computational program that we are refining for apple and will be directly applicable to pear once completed in the next coming months.

The objective of generating the sequence information is currently in progress. We are also in the process of refining the computational methods to integrate random and scaffold sequencing data for building a complete assembly of the pear genome.

We have also provided the DH Pear genome DNA library to IASMA for generating sequences from

the ends of DNA fragments. All this data will aid in our final goal of assembling the DH pear genome. They have committed their own funds to generate this information as part of our ongoing collaboration.

## **RESULTS & DISCUSSION**

Sequence information can be rapidly utilized for developing molecular markers for the pear improvement program. It can also provide complete sequence information for genes where we only have partial information. Over the last year we have utilized the preliminary assemblies for mining such information for various colleagues at WSU. Most importantly this information has been the basis of identifying the complete coordinates and sequence for a putative gene believed to regulate the 2<sup>nd</sup> stage ethylene biosynthetic burst associated with the onset of ripening in winter pear that we identified within another project funded by WTFRC. Knowledge of genes underlying important traits can also serve as targets for improving existing varieties using controlled sports induction (CSI) using non-transgenic approaches. We have a continued emphasis on refining the CSI approach in our program to improve existing varieties thereby circumventing marketing and retail shelf space issues.

The significance of this information will far outlive the duration of this project. Each economically important trait or desirable quality in the fruit tree is controlled at some level by genes. An accessible genomic blueprint of pear enables us to pinpoint what gene or group of genes are responsible for agriculturally important traits. This information will guide pear improvement in both the short and long term future. Another testimony to this fact is that scientists have now discovered the gene underlying skin and lung cancer in humans utilizing human genome information. As in the case of humans, the potential economic benefits to the industry are apparent. With the pear genome sequence in hand, we can develop unique varieties for the PNW combining all priority traits that can create lucrative economic opportunities ranging from production to post-harvest stages.

## **BROADER IMPACTS**

*Presentations:* The pear genome information has been highlighted at several forums over the last year including WSHA meetings. In 2009, the PI was invited to speak at the Hort Show about Enabling Economic Resilience through Genomics Research. Besides that, the work has been shown as poster presentations at annual international meetings like American Society of Plant Biology and Plant and Animal Genome Meeting.

*Publications:* The data generated from WTFRC and WSU-supported DH pear genome will be integrated with the sequence information generated at IASMA and a manuscript will be submitted by June 2010.

*Research:* The apple and pear genome sequencing projects have enabled us to now begin sequencing work in order to obtain the cherry genome. We are also a part of the strawberry and peach genome project consortia.

*Training opportunities:* This project has been steered by graduate student Christopher Hendrickson. Our program has graduated a computer science student (Vandhana Krishnan) who utilized both apple and pear genome data for her MS thesis.

*Community building:* A survey was conducted to gauge the ongoing interest of several pear researchers worldwide and assess their interest in the use of pear genome information. A template of the survey is provided below. Out of 12 groups worldwide, 10 responded and have committed to participating not only in increasing the genomic information for pear but collaborating on providing genotypes, that will include dwarfing trait, other types of genomic information that will help in

advancing the science of pear improvement further.

#### **“Pear Genome Project Collaborators**

The PNW Pear Research Bureau has provided support to initiate genome sequencing of the Double Haploid Pear Genome. You are receiving this message as you had indicated to participate as a collaborator on this project.

There are two main objectives of the approved project

- 1. Prepare a 6X BAC library – *The BAC library has been constructed.***
- 2. Obtain 4X genome coverage using de novo sequencing on the 454 platform**

While these objectives will establish the much needed nucleus for the project, there is a need for additional support to complete the genome. In particular, following resources are needed

1. Genome Scaffolding using 3kb and 20 kb paired-end reads.
2. Solexa-based paired-end read sequencing for fine-scaffolding.
3. BAC-end sequencing.
4. Transcriptome sequencing for functional annotation.

We will be happy to send the 454 Titanium library to any of the colleagues who are capable of performing the sequencing in house. Alternatively, a bulk run on 454 Titanium post-library preparation costs \$7193 and we have a mechanism in place to charge the collaborators if they decide to sponsor some runs. I would like to present a report to the PNW Pear Research Bureau and to facilitate the preparation of this report your feedback is important. I have prepared a feedback form that you can fill out indicating how the pear genome sequence will help your research. This will enable us to apply for funding to federal agencies in the coming months. Please feel free to add as much information as you would like to add. Thanks!

1. Name and area of research in pear biology
2. How do you plan to utilize the pear genome sequence in your ongoing research?
3. Please list synergistic projects and funding amount of these projects that will benefit from the Pear genome sequence.
4. Would you be able to contribute to the pear genome project? If yes, in what capacity?
5. Would you be interested in performing or sponsoring any of the additional tasks listed above?”

## **EXECUTIVE SUMMARY**

*Significant progress:* The objective of generating pear genome sequence coverage is currently in progress. We have devised a new method of generating far more useful and complete information using a novel scaffold-sequence approach. At present we continue to refine the computational methods for creating a complete and efficient pear genome assembly. It is a reiterative process owing to the computational constraints that involves testing different parameters to arrive at the best and most accurate genome assembly possible. Collaborations with IASMA, Italy; Roche Inc. and Andres Bello University have provided extra funds to develop a much finer assembly of the pear genome.

*Outcomes and summary of finding:* Preliminary DH pear genome sequence data are available that are being used by our program to identify coordinates and sequence information of important genes linked to desirable traits; one important one being the cold-induced ripening gene. In summary this is just the start of the most efficient way of connecting traits to genes, an emphasis of our fruit crop genomics program.

*Future directions:* We have two proposals under review at NSF and USDA, and others at various stages of writing to build upon this foundational information. Our programmatic approach is to connect traits with genes using function information. Future projects are aimed at applying this approach in new and novel ways to the improvement of pears.