Research Report

Development of Marker-Based Breeding Technologies for Pear Improvement

David B. Neale, Sara Montanari, Pedro J. Martínez-García and Randi Famula, UC Davis Department of Plant Sciences, Davis, CA
Rachel Elkins, UC Cooperative Extension, Lake and Mendocino Counties
Nahla Bassil and Joseph Postman, USDA/ARS National Clonal Germplasm Repository, Corvallis, OR
Richard L. Bell, USDA/ARS Appalachian Fruit Research Station, Kearneysville, WV
Amit Dhingra, Washington State University Department of Horticulture, Pullman, WA
Kate Evans, Washington State University Tree Fruit Research & Extension Center, Wenatchee, WA

ABSTRACT

Pear production can be increased by developing new varieties with improved agronomic characteristics, such as disease/insect resistances and dwarfing stature, which can be combined with high fruit quality and many other traits. In traditional breeding the selection of such elite cultivars is based on visual evaluation of the phenotype, and in woody perennial crops, including pear, this process is time consuming and expensive, because of the trees' long juvenile phase, laborious trait assessment, and large land requirement. Marker-assisted selection (MAS) technologies are currently routinely and successfully applied for several plant crops, and they can potentially increase pear breeding efficacy. In this project, we aimed at developing a high number of molecular markers to be used to screen the entire germplasm collection held at the USDA Clonal Germplasm Repository in Corvallis, OR. In particular, we re-sequenced 55 pear cultivars and hybrids and we discovered thousands of single nucleotide polymorphisms (SNPs), which will be included in a custom genotyping array. The large set of genotypic data produced with this array will be useful to find marker-trait associations to be applied for MAS in pear.

OBJECTIVES

1. Design a re-sequencing project and a SNP genotyping assay.
2. Collect leaf samples from ~2000 *Pyrus* spp. accessions from the National Clonal Germplasm Repository (NGCR) in Corvallis, OR.
3. Extract DNA from the collected samples.

4. Conduct bioinformatics analysis of the re-sequencing data and design a SNP array.
5. Genotype all the collected samples.
6. Submit the re-sequencing and genotypic data to the Genome Database of Rosaceae (https://www.rosaceae.org/).

PROCEDURES

Design a re-sequencing project and a SNP genotyping assay for pear

Researchers working on pear breeding and genomics in the U.S., their extension collaborators, and the pear marketing boards created the Pear Genomics Research Network (PGRN), with the aim of bringing together their efforts for the enhancement of the pear-growing industry in the U.S. Within this collaboration, we started a re-sequencing project for the evaluation of *Pyrus* genetic diversity. We selected 55 pear accessions, representing founding cultivars and a total of 29 species and hybrids, within the NGCR in Corvallis, OR, and the Appalachian Fruit Research Station (AFRS) in Kearneysville, WV, to constitute the polymorphism discovery panel in this project (Table 1). These accessions were processed for whole-genome, low-coverage sequencing.

Sample collections and DNA extraction

During the summer 2014 we collected leaves from 1870 different *Pyrus* spp. cultivars and hybrids maintained at NGCR and AFRS. For the 55 samples included in the discovery panel, we extracted DNA from freeze-dried leaves using the DNeasy Plant Mini Kit (Qiagen®). For each sample, paired-end libraries were constructed using the Nextera DNA Sample Preparation kit (Illumina®) at the UC Davis Dept. of Evolution and Ecology. Libraries were sent to the Institute for Genomic Medicine at UC San Diego for sequencing on an Illumina® HiSeq2500 in high output mode with v4 chemistry and 2x100 bps runs.

The remaining collected leaf samples were lyophilized for long-term preservation in the Neale Lab at UC Davis.

Bioinformatics analyses of re-sequencing data

We verified the quality of the sequences with FastQC and we calculated the sequencing depth for each sample. We then aligned the sequences to the published *P. communis* 'Bartlett' v1.0 reference genome (Chagné et al., 2014) using the Burrows-Wheeler Alignment (BWA) software. Finally we used the software SAMtools to identify the polymorphic sites (variants) in each of the 55 samples and we reported all the discovered variants in a unique file (VCF file format). We are now processing the VCF file through a three-stage filtration pipeline (Fig. 1), in order to remove artifacts and guarantee a final set of high-quality SNPs. Afterwards, we will submit the filtered SNPs to Affymetrix® for the construction of a custom genotyping array, according to the Axiom myDesign™ protocol, which will include between 50k and 675k molecular markers (depending on the success of the filtration process).

SNP validation and genotyping of the whole NGCR collection
We will select a subset of ~200 samples to be genotyped with the newly developed Axiom array, with the objective of validating the chosen molecular markers and finally use them to genotype the entire set of 1870 samples.

RESULTS

The Pear Genomics Research Network
The University of California (UC) Davis, UC Cooperative Extension, the NGCR in Corvallis, OR, the AFRS in Kearneysville, WV, Washington State University (WSU) and Oregon State University (OSU), have teamed up under the new Pear Genomics Research Network (PGRN), which also involves the industry organizations California Pear Advisory Board (CPAB), Pear Pest Management Research Fund (PPMRF), Pear Bureau Northwest (USA Pears), and Washington Tree Fruit Research Commission (WTFRC). A website for the PGRN (http://ucanr.edu/sites/peargenomics/) was developed in March 2015, and since then there have been 3596 unique visits with 4593 page views.

Re-sequencing of the SNP discovery panel
The SNPs discovery panel included 55 accessions, of which 19 were *P. communis* and *P. communis* subspecies cultivars, 6 were samples from *P. communis* ancestors and close relatives, 8 were varieties from the most widely cultivated Asian species (*P.* x *bretschneideri*, *P. pyrifolia,* and *P. ussuriensis*), 14 were samples from wild East Asian species, and 8 were interspecific hybrids (Table 1). Most of these accessions are founders in the breeding programs at WSU and AFRS, as well as in pear breeding programs carried out in other countries. The wild species were included because of traits of particular interest to breeding programs.
We extracted high quality DNA from these 55 samples. Sequencing resulted in a total of 731.2 Million read pairs, with a per sample coverage of 3.3x to 5.4x. The quality of the sequences was high; hence no manipulation of the reads was necessary before the alignment to the reference genome.

Variants discovery and filtration pipeline
We discovered a total of 66,787,567 unique variants, including 62,176,050 SNPs and 4,611,517 insertions and deletions (indels). Variants passing the Stage 1 filtration will be used to evaluate the genetic diversity among the 55 accessions of the discovery panel. This is expected to give us information about relatedness among the different re-sequenced species and about *Pyrus* domestication.
Afterwards, we will subject variants to Stage 2 and Affymetrix filters, which is expected to drastically reduce their number to a set of high-quality SNPs useful for large-scale genotyping.

DISCUSSION

All partners in the PGRN will take advantage from this new collaboration, as they will share old and new data produced from their individual research projects. In particular, to the end of starting a new pear breeding program at UC Davis, aimed at addressing the main concerns of Californian pear growers, we will benefit from NGCR, AFRS, and WSU researchers' expertise and established resources. On the other hand, we will produce a highly-dense SNP array for pear, which is a fundamental tool for the enhancement of MAS in this crop. Currently, a SNP array including about 1000 European pear SNPs is available (Montanari et al., 2013), and it has been proved useful for the construction of dense genetic maps and application in quantitative trait locus (QTL) mapping projects. However, sequencing technologies have progressed at a very fast pace in the last few years, and it is now possible to design arrays with a much greater number of SNPs at a relatively small cost. Such a high-throughput genotyping tool will enhance genome wide association studies, pedigree-based analysis, and MAS in pear.

The 55 accessions included in the discovery panel are extremely diverse. By detecting species-specific variants, we may be able to identify subgroups of closely related species, thus elucidating their ancestry and natural distribution area, which in some cases is poorly understood. More interestingly, we may identify genomic regions highlighting diversity between cultivated and wild pears; these regions could have been selected during domestication, and thus are associated with important agronomic features.

REFERENCES

Chagné, D., Crowhurst, R. N., Pindo, M., Thrimawithana, A., Deng, C., Ireland, H., … Velasco, R. (2014). The draft genome sequence of European pear (Pyrus communis L. "Bartlett"). *PLOS ONE*, *9*(4), 1–12. http://doi.org/10.1371/journal.pone.0092644

Montanari, S., Saeed, M., Knäbel, M., Kim, Y., Troggio, M., Malnoy, M., … Chagné, D. (2013). Identification of Pyrus Single Nucleotide Polymorphisms (SNPs) and evaluation for genetic mapping in European pear and interspecific Pyrus hybrids. *PLOS ONE*, *8*(10), 1–11. http://doi.org/10.1371/journal.pone.0077022

Table 1: List of the pear cultivars and hybrids included in the polymorphism discovery panel.

List of re-sequenced accessions		
EUROPEAN SPECIES	**ASIAN SPECIES**	**INTERSPECIFIC HY**
Pyrus communis	*Pyrus* × *bretschneideri*	(*Pyrus ussur*
'Anjou'	'Ta-Shian-Sui Li'	*P. communis*
'Bartlett'	'Xuehuali' (Snowflake)	NJ487
'Bosc'	'Ya li'	NJA2
'Coscia'	*Pyrus pyrifolia*	*Pyrus commu*
'Gem'	'Dan bae' (Olympic)	NJB9
'Gin'	'Nijisseiki'	NY 10
'Harrow Delight'	'Zao su'	NY 10
'Harrow Sweet'	*Pyrus ussuriensis*	'Takis
'Old Home'	'Pai Li' (Beijing White Pear)	*Pyrus ussurie*
'Para de Zahar de Bihor'	*P. ussuriensis* No. 2 (Korea)	Illinois
'Roi Charles de Würtemburg'	*Pyrus pashia* 'Naspati'	
'Seckel'	*Pyrus elaeagrifolia* MSU6768	
US 309	*Pyrus glabra*	
US76128-009	*Pyrus regelii*	
US82720-002	*Pyrus sachokiana* GE-2006-115	
Pyrus communis subsp. *pyraster*	*Pyrus salicifolia* GE-2004-141	
'Erabasma'	*Pyrus spinosa* (*amygdaliformis*)	
'Mednik'	*Pyrus syriaca*	
ALB-2011-024	*Pyrus betulifolia* 2291.002	
Pyrus communis subsp. *caucasica*	*Pyrus betulifolia* 2291.006	
Pyrus cordata pure	*Pyrus fauriei*	
Pyrus cordata (Turkey)	*Pyrus koehnei*	
Pyrus cossonii (Russia)	*Pyrus hondoensis*	
Pyrus gharbiana No. 1	*Pyrus pseudopashia*	
Pyrus mamorensis	*Pyrus* × *sinkiangensis* 'Ho mon'	
Pyrus nivalis		

Figure 1: Filtering pipeline applied to discovered variants (work in progress).