Genetic and Developing Genomic Resources in Black Raspberry

N. Bassil, B. Gilmore and K. Hummer
US Dept. Agriculture
Agricultural Research Service (USDA, ARS)
National Clonal Germplasm Repository (NCGR)
33447 Peoria Road
Corvallis, OR 97333
USA

M. Dossett
Agriculture and Agri-Food Canada
Agassiz, BC
Canada

T. Mockler
The Donald Danforth Plant Science Center
St. Louis, MO
USA

S. Filichkin
Center for Genome Research and Biocomputing
Oregon State University (OSU)
Corvallis, OR, 97331
USA

M. Peterson and C.E. Finn
USDA, ARS
Horticultural Crops Research Unit (HCRU)
Corvallis, OR 97330
USA

J. Lee
USDA, ARS
HCRU Worksite
Parma, ID, 83660
USA

G. Fernandez
North Carolina State University (NCSU)
Raleigh, NC
USA

P. Perkins-Veazie
Plants for Human Health Institute
NCSU
Raleigh, NC
USA

C. Weber
Cornell University
Ithaca, NY
USA

R. Agunga and E. Rhodes
Dept. Human and Community Resource Development
Ohio State University
Columbus, OH
USA

J.C. Scheerens
Dept. of Horticulture and Crop Science
Ohio Agricultural Research and Development Center
Ohio State University
Columbus, OH
USA

K. Lewers
USDA-ARS BARC
Genetic Improvement of Fruits and Vegetables Lab
Beltsville, MD
USA

J. Graham
The James Hutton Institute
Invergowrie, Dundee, Scotland
United Kingdom

F. Fernández-Fernández
East Malling Research
East Malling, Kent
United Kingdom

S.J. Yun
Department of Crop Science
Chonbuk National University
Jeonju
South Korea

Keywords: *Rubus occidentalis*, microsatellite markers, genome assembly, expressed sequence tags, breeding

Proc. IIth IS on Biotechnology of Fruit Species
Ed.: S.E. Gardiner
Acta Hort. 1048, ISHS 2014
Abstract

Over the last 75 years, the black raspberry industry in the United States has steadily declined due to lack of adapted and disease resistant cultivars. The high anthocyanin content of black raspberry and associated health benefits have revived interest in production and breeding new cultivars. The United States Department of Agriculture (USDA) Agricultural Research Service, National Clonal Germplasm Repository manages black raspberry germplasm and maintains a collection of over 175 accessions. Wild black raspberries collected in their native range from more than 130 locations across 27 US states and two Canadian provinces were recently added to this collection. Evaluation of this wild germplasm led to the identification of four sources of aphid resistance, two of which were introgressed into the elite breeding pool in two mapping populations. A major focus of this project is to develop, and make available, genomic tools including linkage and physical maps, a draft genome assembly, ESTs, SNP and SSR markers for use in black and red raspberry breeding. We will study genotype by environment interactions in this black raspberry germplasm in four different production regions across North America and apply the genomic tools to identify QTL important for breeding objectives. These tools will facilitate informed decisions regarding germplasm value and usage, crossing, and selection through marker-assisted breeding, and will be useful for breeding programs across the US. Here, we present the current status of global genetic resources and genomic research in black raspberry.

INTRODUCTION

Since the 1940s, the black raspberry (Rubus occidentalis L. subg. Idaeobatus, 2n=2x=14) industry in the United States has steadily declined due to increased disease pressure and lack of resistant cultivars. Black raspberries are economically important as a processed food (e.g., whole fruit, puree, juice) and as a natural food colorant. The high phytochemical content of black raspberry and associated health benefits have revived interest in production and breeding of new cultivars. Unfortunately, few publicly available genomic resources exist for black raspberry, and knowledge about them is limited. Genotypic and phenotypic information must be obtained to implement marker assisted breeding.

The goal of our research is to develop the genomic infrastructure for breeding improved black raspberries. Over the course of the next several years, our objectives include:
- Transcriptome sequencing and high throughput genomic sequencing;
- Molecular marker development from genomic and expressed sequence tag (EST) libraries;
- Studies of genotype by environment interaction in black raspberry mapping populations;
- Construction of a saturated linkage map;
- Evaluation of cross-species transferability of simple sequence repeat (SSR) markers and genomic synteny between black and red raspberry;
- Determination of consumer preferences and market expansion factors.

As a result, industry personnel and students will be trained in molecular approaches to breeding through a multifaceted outreach and extension program.

This report provides an initial review of the current status of global genetic resources and genomic research in black raspberry.

MATERIALS AND METHODS

DNA was extracted from actively growing leaf tissue of three black raspberry selections (ORUS 3021-2, ORUS 4153-1, and ORUS 4158-2) using commercial Cell Lysis and Protein Precipitation Solutions (Qiagen, Inc., Valencia, CA, USA), as previously described (Gilmore et al., 2012). These three selections are parents of two mapping populations, ORUS 4304 (ORUS 4158-2 × ORUS 3021-20) and ORUS 4305 (ORUS 3021-2 × ORUS 4153-1) that segregate for different loci conferring resistance to
the large raspberry aphid, *Amphorophora agathonica* Hottes (Dossett and Finn, 2010). A total of 138 primer pairs, developed from expressed sequence tags (ESTs) of ‘Bristol’, were screened for amplification and polymorphism in these three samples. M13 (TGTTAAAACGACGGCCAGT) tails were added to the 5’ end of each forward primer to allow economical fluorescent labeling of PCR products (Shuelke, 2000) for separation of multiplexed samples by capillary electrophoresis. Polymerase chain reaction (PCR) (Dossett et al., 2012) was performed in 15 µl volume containing 1× GoTaq DNA Polymerase reaction buffer, (Promega Corp., Madison, WI, USA), 2 mM MgCl₂, 0.2 mM dNTPs, 0.12 µM forward primer, 0.5 µM reverse primer, 0.5 µM M13 fluorescent tag (WellRed D2, D3 and D4; purchased from Integrated DNA Technologies, Inc. San Diego, CA), 0.025 U of GoTaq DNA polymerase (Promega Corp., Madison, WI), and 4.5 ng genomic DNA. The PCR protocol consisted of one cycle of initial denaturation at 94°C for 3 min, followed by a ‘touchdown’ PCR program consisting of 10 cycles of 40 s at 94°C; 45 s at 62°C; and 45 s at 72°C. Annealing temperatures were decreased 1°C each cycle to 52°C. PCR continued for 20 additional cycles of 40 s at 94°C; 45 s at 52°C; and 45 s at 72°C and another 8 cycles of 40 s at 94°C; 45 s at 53°C; and 45 s at 72°C with a final elongation step of 72°C for 30 min. PCR products were pooled for fragment analysis on a Beckman CEQ 8000 genetic analyzer (Beckman Coulter Inc., Fullerton, CA) in the presence of CEQ-600 internal lane size standard (Beckman Coulter Inc., Fullerton, CA). Allele sizing and visualization were performed with the fragment analysis module of the CEQ 8000 software.

For genome sequencing and assembly, DNA was extracted from young leaves of a highly homozygous wild selection, ORUS 4115-3, raised from seed collected from Rich Mountain, SC. DNA was extracted with the Omega E-Z 96 Plant DNA Kit (Omega Bio-Tek Inc, Norcross, GA, USA) as described by Gilmore et al. (2012). The DNA was digested with Proteinase K and then precipitated with 0.8 volumes of 2-propanol in the presence of 2.5 M ammonium acetate. The DNA sequencing libraries were prepared using TruSeq DNA Sample Prep v.2 (Illumina Inc., San Diego, CA, USA) according to the manufacturer’s protocol. DNA sequencing and assembly will be described once it is completed.

Total RNA was isolated from leaves, stems, canes, green berries, red berries, and ripe berries of ‘Jewel’ according to a modified protocol as previously described (Filichkin et al., 2010). To ensure a complete and consistent homogenization of hard tissues (stems and canes) and seeds (berries) all the tissues were pulverized in liquid nitrogen using a Mixer Mill MM 301 (Retsch Inc., Germany). The RNA was isolated from approximately 200 mg of frozen tissue powder using a Concert Plant RNA Reagent (Invitrogen, Carlsbad, CA, USA) followed by RNA clean up with Plant RNeasy columns (Qiagen, Inc.) as described by Fox et al. (2009). To avoid contamination with genomic DNA, the RNA sample was treated with Turbo DNase (Invitrogen). The RNA-seq libraries were prepared and barcoded using TruSeq RNA Sample Prep Kits (Illumina Inc., San Diego, CA, USA) according to the manufacturer’s protocol. Equimolar amounts of each of the six libraries were pooled and 101 paired ends (PE) reads were generated in two replicate lanes using the Illumina HiSeq2000 instrument at Oregon State University’s Center for Genome Research and Biocomputing (CGRB).

**RESULTS AND DISCUSSION**

The USDA-ARS NCGR genebank in Corvallis maintains about 230 accessions of *R. occidentalis* which include more than 170 seed lots, over 100 clonal genotypes in pots in a screenhouse, and 27 in vitro cultures. The clonal genotypes include about 30 named cultivars and around 70 that were wild collected or are breeder selections. Breeder selections have been obtained from the North Carolina (NC numbers) or the USDA small fruit breeding program in Oregon (ORUS numbers). These include selections that were found to carry resistance to the large raspberry aphid [ORUS 3778-1 (PI 658505), ORUS 3817-1 (PI 658506), and ORUS 4109-1 (PI 659143)] (Dossett and Finn, 2010). In a recent study, Dossett et al. (2012) used 21 SSR loci to examine genetic diversity in 148 wild and
cultivated black raspberry accessions. The majority of these accessions are maintained at the NCGR in Corvallis. They found that black raspberry cultivars clustered tightly and averaged higher than expected heterozygosity while heterozygosity of wild accessions was much lower than expected. SSR analysis could not distinguish between six of the cultivars (‘Bristol’, ‘Cumberland’, ‘Munger’, ‘New Logan’, ‘Plum Farmer’, and ‘Shuttleworth’) and did not support stated pedigree for some of the cultivars (‘Jewel’, ‘Allen’, ‘Haut’, and ‘Earlysweet’) (Dossett et al., 2012). In contrast to black raspberry cultivars, wild black raspberry germplasm appears to be relatively diverse and presents an untapped resource for future breeding. Information about black raspberry accessions maintained at the NCGR is publicly available at https://www.ars.usda.gov/Main/docs.htm?docid=11370.

In our study 121 of the 138 SSR primer pairs screened over DNA extracted from the parents of the mapping populations generated PCR products. Thirty four and 28 SSRs were polymorphic in parents of the ORUS 4304 and ORUS 4305 mapping populations, respectively. Twenty-one primer pairs generated polymorphic PCR products in both populations. These SSRs are being used to genotype progeny of both populations and construct a linkage map in R. occidentalis. The monomorphic PCR products are being tested by high resolution melting (HRM) analysis for “blind mapping” of the segregating haplotypes as has been recently demonstrated in black raspberry (Bushakra et al., 2012; Dossett et al., 2009). High throughput single nucleotide polymorphism (SNP) markers will also be developed and used to saturate these linkage maps. Up to 200 plants of each of these two populations have been clonally propagated and planted at four research stations (OSU Lewis-Brown Farm (Corvallis, OR), New York Agricultural Experiment Station (Geneva, NY), Ohio State University (Columbus, OH) and North Carolina State University (Raleigh, NC)) and at six commercial farms (Oregon Berry Packing (Hillsboro, OR), Riverbend Farm (Scio, OR), Sandy Farms (Boring, OR), Townsend Farms (Forest Grove, OR), Orchard Dale Farm (Waterport, NY) and SunnyRidge/Dole’s Owl Den Farm (Lincolnton, NC)), where they are being managed by commercial growers. Studying the performance of these seedling populations segregating for adaptation and other important traits in four production regions, Oregon, New York, Ohio, and North Carolina will provide valuable information on relative performance for these traits and effectiveness of selection in different locations with strong small fruit industries and an interest in improved black raspberry cultivars.

Currently, 48 R. occidentalis sequences, mostly from the chloroplast, are available at the National Center for Biotechnology Information (NCBI) nucleotide database (accessed May 24, 2012). The six cDNA libraries for RNA-seq were generated from different tissue types including young leaves, stems, canes, and three developmental stages of berries (i.e., green, red, and black ripe berries) of ‘Jewel’. The number of 101 bp PE reads ranged from 96.7 Mbp in canes to 141.45 Mbp in green berries and was ~704 Mbp in total (Fig. 1). These RNA sequences will be assembled and then used to develop gene models for the black raspberry genome assembly. Initial assembly of the ORUS 4115-3 genome (Table 1) is based on Illumina HiSeq PE 216bp library, 2×101 bp reads at an expected coverage of ~72× and estimated genome size of 268 Mbp (based on Kmer analysis). Low percent SNP bases (0.06%) confirms very low heterozygosity in this selection while high percent of read assembly is consistent with high quality sequence data.

Sequence and marker data generated from this project will be made publicly available through NCBI and the Genome Database for the Rosaceae (GDR) websites. Genetic and genomic resources developed throughout this project will allow identification of candidate genes or markers responsible for many traits of interest for development of improved black raspberry cultivars. These markers will facilitate informed decisions regarding germplasm value and usage, crossing, and selection and will be useful for breeding programs across the US.
ACKNOWLEDGEMENTS

We would like to thank Liz Alperin and April Nyberg for technical assistance in DNA preparation and SSR analysis. We would also like to acknowledge the OSU CGRB staff for their help in Illumina sequencing and data processing and storage. This project is being funded by the USDA-NIFA Specialty Crop Research Initiative and by USDA-ARS CRIS numbers 5358-21000-038-00D, 5358-21000-037-00D, and 5358-21000-041-00D.

Literature Cited


Tables

Table 1. Genome assembly statistics of ORUS 4115-3, a highly homozygous wild selection, made from seed collected from Rich Mountain, SC.

<table>
<thead>
<tr>
<th>Assembly (Mbp)</th>
<th>258.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of scaffolds</td>
<td>42,000</td>
</tr>
<tr>
<td>Average scaffold length</td>
<td>4,994</td>
</tr>
<tr>
<td>N50</td>
<td>24,693</td>
</tr>
<tr>
<td>N90</td>
<td>2,632</td>
</tr>
<tr>
<td>Percent Ns</td>
<td>0.14%</td>
</tr>
<tr>
<td>Percent SNP bases</td>
<td>0.06%</td>
</tr>
<tr>
<td>Percent assembled reads</td>
<td>98.9%</td>
</tr>
<tr>
<td>No. of genes (Augustus)</td>
<td>40,777</td>
</tr>
</tbody>
</table>
Fig. 1. Number of raw 101 bp paired end (PE) reads obtained from the six RNA-seq libraries prepared from green, red and ripe berries, young leaves, stems and canes of ‘Jewel’.