Short-read DNA Sequencing Yields Microsatellite Markers for *Rheum*

Barbara S. Gilmore  
Formerly at USDA-ARS Arctic and Subarctic Plant Gene Bank, 1509 South Georgeson Road, Palmer, AK 99645

Nahla V. Bassil  
USDA-ARS National Clonal Germplasm Repository, 3347 Peoria Road, Corvallis, OR 97333

Danny L. Barney  
USDA-ARS North Central Regional Plant Introduction Station, Ames, IA 50010

Brian J. Knaus  
USDA-ARS Horticultural Crops Research Unit, 3420 NW Orchard Avenue, Corvallis, OR 97330

Kim E. Hummer  
USDA-ARS Arctic and Subarctic Plant Gene Bank, 1509 South Georgeson Road, Palmer, AK 99645; and USDA-ARS National Clonal Germplasm Repository, 3347 Peoria Road, Corvallis, OR 97333

**ABSTRACT.** Identifying and evaluating genetic diversity of culinary rhubarb (*Rheum ×rhabarbarum*) cultivars using morphological characteristics is challenging given the existence of synonyms and nomenclatural inconsistencies. Some cultivars with similar names are morphologically different, and seedlings may grow and become associated with the parental name. Morphological traits of one cultivar may vary when measured under different environmental conditions. Molecular markers are consistent for unique genotypes across environments and provide genetic fingerprints to assist in resolving identity issues. Microsatellite repeats, also called simple sequence repeats (SSRs), are commonly used for fingerprinting fruit and nut crops, but only 10 SSRs have previously been reported in rhubarb. The objectives of this study were to use short-read DNA sequences to develop new di-nucleotide-containing SSR markers for rhubarb and to determine if the markers were useful for cultivar identification. A total of 97 new SSR primer pairs were designed from the short-read DNA sequences. The amplification success rate of these SSRs was 77%, whereas polymorphism of those reached 76% in a test panel of four or eight rhubarb individuals. From the 57 potentially polymorphic primer pairs obtained, 25 SSRs were evaluated in 58 *Rheum* accessions preserved in the U.S. Department of Agriculture, National Plant Germplasm System. The primer pairs generated 314 fragments with an average of 12.6 fragments per pair. The clustering of many accessions in well-supported groups supported previous findings based on amplified fragment length polymorphisms (AFLPs). Cluster analysis, using the proportion of shared allele distance among the 25 SSRs, distinguished each of the 58 accessions including individuals that had similar names or the same name. Accessions that grouped in well-supported clusters previously belonged to similar clusters with high bootstrap support based on AFLP. In summary, our technique of mining short-read sequencing data was successful in identifying 97 di-nucleotide-containing SSR sequences. Of those tested, the 25 most polymorphic and easy-to-score primer pairs proved useful in fingerprinting rhubarb cultivars. We recommend the use of short-read sequencing for the development of SSR markers in the identification of horticultural crops.

The genus *Rheum* belongs to the family Polygonaceae and includes ≈55 species with 14 synonyms [U.S. Department of Agriculture (USDA), 2012]. Wang et al. (2005) reported that *Rheum* originated in the central and northern plateaus of Asia. In addition to its contemporary uses for cooking in a variety of dishes, the medicinal effects of several species in this genus have been described by multiple cultures. Rhubarb species are discussed in traditional Chinese medicine, *Shen Nung Pen Ts’ao Ching* (Hsu et al., 1986), in *Historia Naturalis* by Pliny the Elder (Jones, 1956), in *De Materia Medica* of Dioscorides (English translation by Beck, 2005), and in *Al-Qanun fi al-Tibb* (Canon of Medicine) by Ibn Sina in Syria (Barney and Hummer, 2012). From Asia and the Mediterranean region, medicinal rhubarb spread to Europe and North America (Beck, 2005; Foust, 1992; Sibly, 1790; Turner, 1938).

*Rheum* has a base chromosome number of *x* = 11 and includes species with ploidy ranging from diploid (*2n* = 2*x* = 22) through hexaploid (*2n* = 6*x* = 66). Culinary rhubarb, *Rheum ×rhabarbarum*, is mostly tetraploid, although its origin is complex, likely involving *R. undulatum*, either diploid or
Table 1. Fifty-three *Rheum ×rhabarbarum* accessions and five *Rheum* species representatives evaluated from the U.S. Department of Agriculture–Agricultural Research Service National Plant Germplasm System (NPGS).z

| Accession name          | Similarly named groups | PI no. | NPGS ID |
|-------------------------|------------------------|-------|---------|
| **R. ×rhabarbarum**     |                        |       |         |
| Ayakulik                |                        | RB63  |         |
| Ben Vanderweel 1       | Ben Vanderweel         | RB65  |         |
| Ben Vanderweel 2       | Ben Vanderweel         | RB66  |         |
| Canada Red              |                        | RB25  |         |
| Cawood Delight         |                        | 666012| RB20    |
| Cherry Red              |                        | RB23  |         |
| Chipman                 |                        | 666022| RB36    |
| Cooper                  |                        | 666021| RB35    |
| Coulter MacDonald       | Coulter MacDonald       | 666004| RB10    |
| Coulter McDonald        | Coulter MacDonald       |       |         |
| Crimson Cherry          | Crimson Cherry         | 665999| RB05    |
| Crimson Cherry          | Crimson Cherry         | 666024| RB39    |
| Crimson Delicious 136/30|                        | 666001| RB07    |
| Crimson Red             | Crimson Red            | 666007| RB13    |
| Crimson Red             | Crimson Red            | 666006| RB22    |
| Crimson Wine 115/55     |                        | 666003| RB09    |
| Cyclople                |                        | 666028| RB51    |
| German Wine (morphotype 1) | German Wine          |       |         |
| German Wine (morphotype 2) | German Wine          |       |         |
| Goiath                  |                        | 666011| RB18    |
| Johnsons St. Martin 426/27 |                  | 666017| RB29    |
| Kerwin                  |                        | 666027| RB42    |
| Kremlin Red             |                        | 666029| RB67    |
| Linnaeus 137/31         |                        | RB03  |         |
| Loher Blutz             |                        | 666026| RB41    |
| Mammoth                 |                        | 666016| RB38    |
| MacDonald               | MacDonald              | 666023| RB27    |
| McDonald                | MacDonald              | 666015| RB26    |
| Minn No. 8              |                        | 666019| RB33    |
| Moore’s Red-Right-Thru  |                        | 666009| RB16    |
| New Zealandz            |                        | 666005| RB11    |
| North Pole Variety      |                        | 666010| RB17    |
| OR 23                   |                        | 666029| RB52    |
| Oregon Giant            |                        | 666018| RB32    |
| Paragon 7/27            |                        | 666006| RB12    |
| Parsons Crimsonz        |                        | 665998| RB04    |
| Penn. State #3z         |                        |       |         |
| Plum Hutt               |                        | 666014| RB24    |
| Prince Albert           | Prince Albert          | 666036| RB60    |
| Prince Albert           | Prince Albert          |       |         |
| Red Right Through       |                        | 666037| RB34    |
| Reeds Early Superb      |                        | 666020| RB61    |
| Ruby 565/27             |                        |       |         |
| Strawberry              | Strawberry             | 666037| RB34    |
| Strawberry              | Strawberry             | 666020| RB61    |
| Strawberry              | Strawberry             | None  |         |
| Sunrise                 |                        | 666009| RB44    |
| Sutton                  | Sutton                 | 666000| RB06    |
| Sutton Seedless         | Sutton                 | 666013| RB21    |
| The Sutton              | Sutton                 | 666008| RB14    |
| Timperley Early         |                        | 666013| RB21    |
| Valentine               |                        | 666037| RB43    |
| Victoria                |                        | 665997| RB01    |

Table 1. Continued.

| Accession name          | Similarly named groups | PI no. | NPGS ID |
|-------------------------|------------------------|-------|---------|
| **R. officinale**       |                        |       |         |
| Rubra (R. palmatum)     |                        | 666002| RB08    |
| 2/24 UK (R. palmatum)   |                        | RB50  |         |
| **R. palmatum**         |                        |       |         |
| var. tanguticum         |                        | RB48  |         |
| **Unknown *Rheum* species** |                      |       |         |
| RK Lot 5405333/CC/JO   |                        |       |         |

zSimilarly name accessions were placed into nine groups. Plant introduction (PI) number and NPGS identity (ID) are listed, where available.

During 2 years of trials at Palmer, Pantoja and Kuhl (2009) evaluated the effectiveness of 15 morphological characters for identifying cultivars and other clonal rhubarb specimens. Significant variability between the 2 years was reported for petiole epidermis color, petiole internal color, and the number of petioles per plant. The results suggest limited usefulness of these descriptors for identifying rhubarb accessions (Pantoja and Kuhl, 2009).

A few molecular marker studies have been reported in rhubarb. Kuhl and DeBoer (2008) used AFLP markers to investigate the genetic diversity of 41 accessions from the Palmer *Rheum* collection. Hu et al. (2010, 2011) used inter-simple sequence repeat (ISSR) markers to assay genetic diversity in Chinese populations of *R. tanguticum*, whereas Wang (2011) optimized ISSR–polymerase chain reaction in *R. officinale*, *R. palmatum*, and *R. tanguticum*. Compared with AFLPs and ISSRs, microsatellite or SSR markers are more reproducible and can complement morphological trait assessments in this task.

Advances in short-read sequencing platforms have considerably decreased the cost of identifying microsatellite repeats.
and enabled the development of SSR markers for species of little economic importance (reviewed in Zalapa et al., 2012). For example, DNA sequencing using next-generation sequencing technology (Illumina platform; Illumina, San Diego, CA) allowed development of polymorphic SSR markers for alaska yellow cedar [*Callitropsis nootkatensis* (Jennings et al., 2011)] and mile-a-minute weed [*Mikania micrantha* (Yan et al., 2011)]. The objective of this study was to develop SSR markers from short-read DNA sequences and use them to fingerprint the NPGS rhubarb collection.

**Methods and Materials**

**PLANT MATERIALS AND DNA ISOLATION.** Plants were field grown in Palmer, AK. Young, actively growing leaves were collected from 58 clonal rhubarb accessions on 22 June 2011 and shipped for overnight delivery to the USDA-ARS National Clonal Germplasm Repository in Corvallis, OR. The accessions included 53 *Rheum rhabarbarum* and one additional accession each of: *R. officinale*, *R. palmatum*, *R. palmatum* ‘Rubra’, *R. palmatum* var. tanguticum, and an unknown *Rheum* species (Table 1). Within 24 h of receiving the leaves, 30 to 50 mg of leaf tissue from each accession was placed in a cluster tube, frozen in liquid nitrogen, and stored at –80°C as described by Gilmore et al. (2011).

**DNA SEQUENCING AND SSR DISCOVERY.** *Rheum rhabarbarum* accession RB14 ‘Timperley Early’ was one of eight Illumina/C210 sequencing libraries that were pooled and sequenced in one lane with the Illumina/C210 Genome Analyzer II at the Oregon State University Center for Genome Research and Biocomputing (CGRB). Each library had a different bar-coded adapter to allow assignment of sequences to the correct plant. The remaining seven libraries were representatives of *Peonia* species and hybrids (Gilmore et al., 2013). A single Illumina® library was prepared from DNA of RB14 ‘Timperley Early’ and consisted of: DNA sonication [Bioruptor XL (BR_XL); Diagenode, Denville, NJ]; DNA strand end repair; addition of an adenine base to the 3′ ends; ligation of a custom Illumina/C210 sequencing adapter containing a 4-bp barcode (AGGT) to the DNA fragments; purification of the ligation products to remove any unligated adapters or adapter–adapter products using a low melting point 1.5% agarose gel (Lonsa NuSieve, Rockland, ME); and

Table 2. Primer names and sequences (5′-3′), number of fragments (*A*), and observed size range at 25 di-nucleotide repeats in 58 *Rheum* accessions.

| Primer | Sequence (5′-3′) | *A* | Size range (bp) |
|--------|------------------|-----|----------------|
| Rhe01  | F: GGAGGCCAACAGTTATAGCTC | 17  | 192–239 |
|        | R: CCTAGTTTTCCCTTGTITTTTCC |       |       |
| Rhe03  | F: ACATTTCATCGTGGGGAGTG | 15  | 120–148 |
|        | R: TGATTTGTATCACTTTTTGCTG |       |       |
| Rhe09  | F: CTCATAGCGTCGTTGCACAC | 17  | 160–279 |
|        | R: GCCACGTTATAACACCTTGGC |       |       |
| Rhe12  | F: TAGAGTGTCGGAACACCTG | 5    | 178–188 |
|        | R: TCCATGGAATCATACGACCAA |       |       |
| Rhe13  | F: TGGCCTCTCTTTTTATTTGC | 7    | 212–226 |
|        | R: TTTTTCGAAAACACCTTTCA |       |       |
| Rhe15  | F: GGTTGTTGGAATACATGGTGA | 24   | 134–228 |
|        | R: GCTACACAACCACAAACAA |       |       |
| Rhe31  | F: CTTTGAAGAAATCTAAGGTG | 13   | 157–194 |
|        | R: GAGTGTATAGCCCCCAGCT |       |       |
| Rhe36  | F: TGGAAACCAAATAAATAACACTTC | 6    | 124–138 |
|        | R: TTTTCGGAATTAGTTTTTCTA |       |       |
| Rhe43  | F: AGTTGGTGATCTCCTGCGT | 7    | 162–185 |
|        | R: ACTTCGCTATTTGTGCTTG |       |       |
| Rhe45  | F: GGATCCACTCTCCTCAGATCA | 13   | 183–305 |
|        | R: GAACGGCAAGATGAACAGGG |       |       |
| Rhe47  | F: GCCTACCGGAGAGAACAAC | 8    | 161–179 |
|        | R: TGGAAATTGTGGGAGATAA |       |       |
| Rhe50  | F: CATGGATTGGATCCTCAAG | 12   | 113–184 |
|        | R: TGGGATCTCCTTAAAATTGCT |       |       |
| Rhe54  | F: TGACGCTCTCTCTCTCCTC | 6    | 124–214 |
|        | R: TACCTCGATTTGCTGCTCT |       |       |
| Rhe55  | F: CGTTTCTCCTACGGTTGATT | 19   | 328–390 |
|        | R: ATTTTGCTGCTGATTGCTT |       |       |
| Rhe56  | F: GATGGGAACGTGCCCCTTT | 14   | 173–235 |
|        | R: ACTGAATGTTGACATGGA |       |       |
| Rhe57  | F: AACAAGCTATGAGGCCCCAAG | 3    | 190–194 |
|        | R: CCTTCGACTTCTCCTTTCTA |       |       |
| Rhe61  | F: TGTCCTCTCTCTCTCCTCCT | 6    | 194–204 |
|        | R: GACGCAGGAAGAGGATCTCT |       |       |
| Rhe65  | F: TCGATACCTTATGCTCTGAAA | 11   | 190–354 |
|        | R: GTGCAGCAAGGTCTCCTA |       |       |
| Rhe72  | F: AAGTGGTGACCTCTGCCA | 30   | 125–501 |
|        | R: TCTCGCTCTCTCCTCTCTG |       |       |
| Rhe77  | F: CTCATAGGCGCTTGTCAC | 10   | 162–220 |
|        | R: GCCCACTGATAAACCCATT |       |       |
| Rhe80  | F: AGAGATGGCGATGGGCAATTA | 10   | 163–181 |
|        | R: GATTTTGCTATTTCTCCTTT |       |       |
| Rhe82  | F: GATCTATCGGCCCAATT | 18   | 143–183 |
|        | R: ACAGGCCCCATCGAAAGTTCC |       |       |
| Rhe86  | F: CATGCTATGCTGAAAGTGAC | 4    | 191–197 |
|        | R: TGGATTTGGAATGGGAAACC |       |       |
| Rhe90  | F: CAAAGGGCTGGAAGAGTTG | 13   | 72–96 |
|        | R: CTGGCTTTGATCAAGACCTCTC |       |       |
| Rhe93  | F: CACTCTAATACCACCTCGGTTCA | 26  | 187–343 |
|        | R: AGGCCATTCATCAAATCCTC |       |       |

*The fragment sizes listed include the M13 tail.

and the development of SSR markers for species of little economic importance (reviewed in Zalapa et al., 2012). For example, DNA sequencing using next-generation sequencing technology (Illumina platform; Illumina, San Diego, CA) containing a 4-bp barcode (AGGT) to the DNA fragments; purification of the ligation products to remove any unligated adapters or adapter–adapter products using a low melting point 1.5% agarose gel (Lonsa NuSieve, Rockland, ME); and

24 J. AMER. SOC. HORT. SCI. 139(1):22–29. 2014.
excision of a 350-bp DNA band from the gel smear with a scalpel (5-Prime SafeXtractor-25; Fisher, Waltham, MA). The adapter-modified DNA fragments that had adapter molecules on both ends were selectively enriched by polymerase chain reaction (PCR) for 18 cycles. The PCR products were purified using a QIAquick PCR Purification Kit (Qiagen, Valencia, CA) and then separated on a 1.5% agarose gel to verify library size and to visually estimate the fragment concentration. The extracted DNA was quantified, and the 260/280 absorbance ratio was determined using a ultraviolet-Vis spectrophotometer (NanoDrop ND-1000; Thermo Scientific, Wilmington, DE). The library size was determined with a Bioanalyzer (Agilent 2100; Agilent Technologies, Wilmington, DE) at the CGRB. After diluting and pooling in equimolar amounts, the bar-coded Rheum library and seven other libraries (7.5 pM) (Gilmore et al., 2013) were submitted for paired-end 80-bp sequencing in a single lane by the Illumina/C210 Genome Analyzer II at the CGRB. Illumina/C210 Version 3.0 reagents were used for cluster generation and sequencing, image acquisition, and base calling using the Illumina/C210 pipeline Version 1.5. The resulting microreads were sorted by barcodes using a custom PERL script BCSORT (Knaus, 2011). Sorted reads were searched only for di-nucleotide motifs. The di-nucleotide motif guidelines were microreads containing at least four perfect repeats. Both nucleotides were represented at least four times and had fewer than eight ambiguous (N) bases. Paired-end microsatellite-containing reads were joined into a single sequence (contig) by concatenating read 1, the reverse complement of read 2, and separated by 50Ns. This was used to identify the break between microreads. A stringent filter was applied to identify microreads with microsatellites located near the center of the sequence. The output was filtered for redundant sequences (identity ≥95% or greater) to a single unique microread by clustering the reads using the program cd-hit-454 (Niu et al., 2010). This stringent filter yielded the largest possible

Table 3. List of 29 additional simple sequence repeat primers screened for polymorphism in eight *Rheum rhabarum* accessions ‘Crimson Delicious’ (RB7), ‘Coulter MacDonald’ (RB10), Penn. State 3 (RB12), ‘New Zealand’ (RB19), ‘Parson’s Crimson’ (RB32), ‘Crimson Cherry’ (RB39), ‘Lober Blut’ (RB41), and *R. officinale* (RB45) either by 2% agarose gel electrophoresis (AGE) or by capillary electrophoresis (CE).z

| Name  | Sequence | Polymorphism |
|-------|----------|--------------|
| Rhe14-F | F: TGGCTTCAGAAAGAGAAGA | CE |
| Rhe14-R | R: ATGAAATACCCCAACCGACA | |
| Rhe17-F | F: ACACAACAAGTCCACACCA | CE |
| Rhe17-R | R: TTTCTTTTCCGTGGAGGAC | |
| Rhe20-F | F: TGAATATGGAAAGACACACA | CE |
| Rhe20-R | R: TCCATAAAGAAAGTAATCCTGA | |
| Rhe44-F | F: TCAAATTTGGAACCTTAAGTCT | CE |
| Rhe44-R | R: TTTGAGTAAATTTGAGAGAGA | |
| Rhe59-F | F: TGGTTCGGAAATTTGAGATGG | CE, null |
| Rhe59-R | R: CCCCAATAACAAACCTAAT | |
| Rhe83-F | F: GGAAGCCGCTGCAAGCTCTC | CE, null |
| Rhe83-R | R: TAGGGCTTCAAGGGCCTTCC | |
| Rhe91-F | F: GCAATTATGTTAGCGGATT | CE, null |
| Rhe91-R | R: TCTCGGTTATTTGAGGTTT | |
| Rhe07-F | F: AAGGACCTGAAACTGGGAAGA | CE, null |
| Rhe07-R | R: GAAAGCCCTCATGTTTGGGA | |
| Rhe18-F | F: GATTTTGGAGACCCAGATG | CE, null |
| Rhe18-R | R: AAAGGGCTTCTTATTTGACTCTC | |
| Rhe21-F | F: GCCCTATACACAGTCCCCAAA | CE, null |
| Rhe21-R | R: CCACAGCCCTATATGAGT | |
| Rhe22-F | F: CCACTAAGAGCGCTTATC | CE, null |
| Rhe22-R | R: AGCAATTCTGAAAGCCAATC | |
| Rhe29-F | F: TTTCGAAGAACACTAAATCCTAAA | CE, null |
| Rhe29-R | R: GATATATTATGATGCTGGCCATA | |
| Rhe33-F | F: TGGGAAGCCCTCAAATATGAG | CE, null |
| Rhe33-R | R: GGTTGATAACACACAGCAC | |
| Rhe85-F | F: GTATCTTCGCAAAGCCTAATC | CE, null |
| Rhe85-R | R: GGCTTGCGCTGAATGAAAATA | |
| Rhe87-F | F: CTGAATTTGCGCCCCTTGT | CE, null |
| Rhe87-R | R: AGAGCCTTATCATCCGCAT | |
| Rhe95-F | F: GTTATTAATGGAGGGCGTGTC | CE, null |
| Rhe95-R | R: GAAACCGAGAGCCACCTTAAAT | |
| Rhe06-F | F: CGAAATTTCTCCTTTCTCTCT | AGE |
| Rhe06-R | R: AAGGAAAGAGATGCCTGCACAC | |
| Rhe24-F | F: TGTTTGGTTGAGAATATCATG | AGE |
| Rhe24-R | R: CTCGCGGATACCTCATGCTA | |
| Rhe27-F | F: TGGTAAATGTTGGGAGCAG | AGE |
| Rhe27-R | R: TTTCAAACTGCAACACCTA | |
| Rhe30-F | F: TCACACATCTCAACACAGAC | AGE |
| Rhe30-R | R: TGAATTGTGCGGGTTG | |
| Rhe35-F | F: CCCTCAAGGCAAGCATATA | AGE |
| Rhe35-R | R: TGGGAATTGATATTTATGCTG | |
| Rhe38-F | F: ATCCACATCAAGTGCCCTCAG | AGE |
| Rhe38-R | R: TCCGCTGAAATAGTTGTTG | |
| Rhe63-F | F: TGGATTTTTGTGATCCATCA | AGE |
| Rhe63-R | R: AAAAATGGGTTGGGATAAATCAGT | |
| Rhe67-F | F: TAGTTTCCCCTCAACCTACG | AGE |
| Rhe70-F | F: RCAGCTATATGGTTTGTGCT | |
| Rhe70-R | R: CCAAAAGAAAGACTCCAACGA | |
| Rhe71-F | F: ATTATGCGGATGTTGCT | |
| Rhe71-R | R: TCCAAAAAGAAAGACTCCAACGA | |
| Rhe73-F | F: CCTTTAGCACAATCTACGTTG | |
| Rhe73-R | R: GAAGCTTATAGCCCACACTG | |
| Rhe78-F | F: GATGAGACCTTGCAGAAGAGA | |
| Rhe78-R | R: TTTAAGCTCGCTTCAGGGT | |

Continued next page
flanking sequences for subsequent primer design (Jennings et al., 2011).

**Microsatellite marker development.** PCR primers were designed for each SSR-containing singleton cluster and contig sequence with BatchPrimer3 (You et al., 2008). Default settings were used except for product size, which was increased to a maximum of 300 bp to the 5′ end of each forward primer; an M13 sequence (TGTAAAACGACGGCCAGT) was added to allow for an economic method of fluorescent labeling of PCR products following the procedure outlined by Schuelke (2000). The 97 M13-tagged forward primers and corresponding reverse primers, along with universal fluorescently labeled M13 (-21) were ordered from Integrated DNA Technologies (IDT, San Diego, CA). The primers were tested for amplification in four rhubarb accessions, RB7, RB10, RB12, and RB45, and products were separated by 2% agarose gel electrophoresis and visualized with ethidium bromide. Thermocycler amplification of the PCR products was performed under conditions of an initial denaturing step of 94 °C for 3 min, then 10 cycles of 94 °C for 40 s, 62 °C for 45 s (decreasing the annealing temperature by 1.0 °C per cycle), and 72 °C for 45 s followed by 20 cycles of 94 °C for 40 s, 52 °C for 45 s, and 72 °C for 45 s; eight cycles of 94 °C for 40 s, 53 °C for 45 s, and 72 °C for 45 s; and a final extension of 72 °C for 30 min. The 15-μL PCR reaction mix contained: 3 μL of 5X GoTaq DNA Polymerase Buffer (Promega Corp., Madison, WI); 1.2 μL of 2.5 mM dNTPs; 1.2 μL of 25 mM MgCl₂; 0.075 μL of 5U/μL GoTaq DNA Polymerase (Promega Corp.); 0.18 μL of 10 μM forward primer; 0.75 μL of 10 μM reverse primer; 0.75 μL of 10 μM M13-fluorescent tag, WellRED D2, D3, or D4; and 1.5 μL of 3 ng-μL⁻¹ template DNA.

Ninety SSR markers generated a PCR product, which were then evaluated for polymorphism in eight rhubarb accessions, ‘Crimson Delicious’ (RB7), ‘Coulter MacDonald’ (RB10), Penn. State 3 (RB12), ‘New Zealand’ (RB19), ‘Parson’s Crimson’ (RB32), ‘Crimson Cherry’ (RB39), ‘Lober Blu’ (RB41), and R. officinalis (RB45). The PCR products were separated by 2% agarose gel electrophoresis and visualized with ethidium bromide. Of those 90, 75 that appeared polymorphic on agarose were screened by capillary electrophoresis using the CEQ 8000 (Beckman Coulter, Brea, CA) to confirm polymorphism and evaluate ease of scoring.

**Fingerprinting and genetic diversity evaluation.** Twenty-five primer pairs were selected for fingerprinting and genetic diversity assessment of the 58 accessions in the rhubarb collection (Tables 2 and 3). The PCR products were pooled into 12 multiplexes before capillary electrophoresis to maximize the number of PCR fragments separated in one capillary channel. Each multiplex contained fragments that did not overlap in size and had different fluorescent colors. Allele sizing and visualization were performed using the fragment analysis module of the CEQ 8000 software (Beckman Coulter). Alleles were scored by fitting peaks into bins that were less than one nucleotide wide. The exact copy number of each allele was not inferred in the rhubarb individuals as a result of unknown ploidy. Therefore, individuals were scored for the presence or absence of each allele and each allele was treated as a separate locus.

**Results and Discussion**

**Microsatellite locus discovery.** In all, 48.5 million barcode reads were generated in the Illumina® flow cell lane that contained one rhubarb library and seven peony libraries. The rhubarb library comprised 1.3 million reads compared with a high of 14 million bar-coded reads for Paeonia rockii and a low of 2.8 million reads for P. delavayi var. lutea × P. ×suffruticosa that were included in the same lane (Gilmore et al., 2013). The DNA used to prepare the rhubarb Illumina® library was viscous, indicating poor quality, which may have led to fewer sequence reads for rhubarb than was previously found for the seven peony libraries (Gilmore et al., 2013). The number of primer pairs designed to amplify di-nucleotide SSR repeats was proportional to the number of bar-coded sequence reads obtained from each library. Whereas only 97 primer pairs were designed from the rhubarb library that generated the smallest number of sequences, 2219 primer pairs were designed from the library that generated the largest number of sequence reads, P. rockii. Still, at an estimated reagent and sequencing cost of $215, we obtained many new SSRs in rhubarb, a crop in which only 10 SSRs were previously described (Chen et al., 2009).

Of the 97 rhubarb primer pairs designed to amplify di-nucleotide motifs present in the sequenced library, 25 primer pairs were polymorphic and were evaluated further in this study. Of the remaining 72 primer pairs, 22 failed to amplify or produced unclear results and were discarded; 18 were monomorphic; three were monomorphic in the R. ×rhabarum accessions but had a unique allele within R. officinalis; seven were polymorphic but were not evaluated further; nine were polymorphic but failed to amplify in one or more of the eight tested accessions; and 13 generated polymorphic DNA fragments based on agarose gel analysis but need further evaluation.
by capillary electrophoresis to confirm polymorphism (Table 3). Therefore, observed and potential polymorphism in the 29 primer pairs that exhibited polymorphism in the seven *R. schabarum* test sample panel used needs further evaluation and is provided (Table 3). It is possible that a larger number of SSR markers could be generated from this library by optimizing PCR amplification conditions for the 22 primer pairs that failed to amplify and by mining the sequence data for tri- and tetra-nucleotide repeats. The raw sequence data from this study are available at the sequence read archive at the National Center for Biotechnology Information (NCBI), accession number SRA054037 (NCBI, 2012).

**Microsatellite marker diversity.** The 25 SSR primer pairs generated 314 fragments for an average of 12.56 fragments per SSR primer pair in the 58 accessions (Table 2). The number of fragments, *A*, ranged from three for Rhe57 to 30 for Rhe72 and on average was larger than the 4.9 previously reported by Zhang et al. (2008). This difference in *A* is not surprising given the small number of samples (12 *R. tanguticum*) used by Zhang et al. (2008) compared with the 58 accessions in this study. Ploidy may also have been a factor. However, *A* is comparable to the 14.6 obtained in the 114 *R. tanguticum* from 10 geographically separate populations evaluated with seven SSRs (Chen et al., 2009).

**Fingerprinting and genetic diversity in Rhubarb.** Based on NJ cluster analysis using the proportion of shared allele distance for 25 SSRs, each of the 58 accessions was uniquely identified (Fig. 1). Even the nine sets of accessions that had similar or identical names had unique SSR profiles (Table 1). Accessions from all but four (‘Strawberry’, ‘Crimson Red’, ‘Prince Albert’, and ‘German Wine’) of these similarly named individuals were grouped together in the dendrogram indicating a close relationship (Fig. 1). One unnumbered ‘Strawberry’ from Palmer did not group with the other two accessions designated as ‘Strawberry’ (RB34 and RB61) and had a pairwise genetic

![Fig. 1. Neighbor joining (NJ) dendrogram of 58 Rheum accessions based on 25 simple sequence repeats (SSRs) using the proportion of shared allele distance. Bootstrap support based on 1000 iterations is indicated above each branch if it exceeded 50%. Five groups ('Cherry Red', 'Canada Red', 'Parson’s Crimson', 'Valentine', and 'Sutton') with bootstrap support 72% or greater were named following Kuhl and DeBoer (2008) in the amplified fragment length polymorphism (AFLP)-based unweighted pair group method with arithmetic mean (UPGMA) dendrogram. Three additional groups with strong bootstrap support were also observed: 'Ben Vanderweel' (72%), 'MacDonald' (74%), and 'Coulter MacDonald' (82%).](image-url)
distance of 0.176 and 0.170 to RB34 and RB61, respectively. This finding suggests that the unnumbered ‘Strawberry’ accession is mislabeled as speculated by personnel at the ASPGB (Fig. 1). The genetic distance was 0.296 between the two ‘Crimson Red’ (RB13, RB22) accessions, 0.239 between the two ‘Prince Albert’ (RB24, RB61), and 0.298 between the two ‘German Wine’ morphotypes (RB30, RB31). The distances indicated that these three similarly named accessions were genetically different. Genetic distance among these 58 accessions ranged from 0.012 between ‘Sutton Seedless’ (RB47) and ‘Cooper’ (RB35) to 0.44 between ‘Paragon’ (RB52) and R. palmatum (RB50) and averaged 0.113 (Fig. 1). Rhubarb is clonally propagated so the genetic distance should be zero among plants of the same genotype. However, rhubarb readily outcrosses, is wind-pollinated, and the seeds easily germinate. Contamination from seedlings could explain the genetic variation we observed among these similarly named accessions. These results agree with Kuhl and DeBoer’s (2008) suggestion that the two ‘German Wine’ or ‘Prince Albert’ or ‘Crimson Red’ accessions could have arisen from seedlings. Mislabeling could also have occurred during exchange of these rhubarb clones.

The majority of the well-supported groups (72% or greater bootstrap support) in this study was also well-supported based on AFLP analysis of 37 culinary rhubarb accessions and four species representatives (greater than 90% bootstrap support; Kuhl and DeBoer, 2008), suggesting close genetic relationships. These five well-supported groups were named following Kuhl and DeBoer (2008) and include: 1) ‘Cherry Red’; 2) ‘Canada Red’; 3) ‘Parson’s Crimson’; 4) ‘Valentine’; and 5) ‘Sutton’. Three additional groups with strong bootstrap support were also observed: ‘Ben Vanderweel’ (72%), ‘MacDonald’ (74%), and ‘Coulter MacDonald’ (82%) (Fig. 1). The Palmer Genebank obtained ‘Coulter MacDonald’, ‘Coulter McDonald’, and ‘Plum Hutt’ from the University of Minnesota. These cultivars originated from the Morden Experiment Station, Morden, Manitoba, Canada. ‘Kremlin Red’ also grouped with these accessions. This ‘Kremlin Red’ accession was found near the original USDA building (commonly referred to as the “Kremlin” building) in downtown Palmer, AK, and may have been brought from Minnesota by one of the 55 founding families that colonized Palmer, AK, in 1935 (Estelle, 2009).

There is limited historical and pedigree information about these rhubarb accessions and this severely limits discussion of the observed clustering. Furthermore, the deep branching observed in the dendrogram that was generated from binary analyses of SSR data across varying levels of ploidy in Rheum precludes inference of relationships among these accessions. Still, comparable well-supported grouping based on Kuhl and DeBoer’s (2008) AFLP-based analyses and the current SSR analyses indicates close relationships between cultivars that are present in these groups.

Similarities in clustering were observed between the SSR-based dendrogram (Fig. 1) and the unweighted pair group method with arithmetic mean dendrogram generated for 36 culinary rhubarb cultivars and two rhubarb species based on 19 morphological and phytochemical characters (Pantoja and Kuhl, 2009). Many accessions from well-supported groups in this study also grouped together based on morphological and phytochemical traits with some exceptions; ‘Moore’s Red-Right-Thru’ was in the ‘Cherry Red’ group; ‘Canada Red’ and ‘Coulter MacDonald’ grouped together; ‘Minn No. 8’ grouped with ‘Parson’s Crimson’, whereas ‘Strawberry’ grouped with ‘The Sutton’ and ‘German Wine’. Based on morphological and phytochemical differences, ‘Sutton’ and ‘The Sutton’ were different and did not group together (Kuhl and DeBoer, 2008); however, they grouped closely in the ‘Sutton’ group with a genetic distance of 0.038 based on SSR analysis (Fig. 1). On the other hand, ‘Mammoth’ and ‘Cooper’ could not be separated based on morphological and phytochemical characters but were easily separated in the ‘Sutton’ group based on SSR analysis (Fig. 1). Therefore, morphologically similar cultivars can be differentiated with SSR markers.

**Conclusion**

In our study, short-read DNA sequencing and filtering generated 97 di-nucleotide-containing sequences for rhubarb. Further examination of the raw data would likely identify sequences with larger (tri- to hexa-) repeat motifs. The raw sequence data (accession number SRA054037) are available at the sequence read archive at NCBI (NCBI, 2012).

We found that 25 of the new SSRs were effective for fingerprinting rhubarb accessions. The SSRs distinguished each of the rhubarb accessions and no duplicates were identified within the nine groups of accessions with the same or similar cultivar names.

Rhubarb SSRs may be valuable to medical researchers assayng Rheum species for species identification or for purity testing of dry medicinal material. Furthermore, these SSRs could be used in future genetic studies of rhubarb such as linkage map construction, evaluation of genetic diversity, and population structure of wild rhubarb populations from different species including R. officinale, R. palmatum, and R. tanguticum. We recommend the use of short-read sequencing for the development of SSR markers in the identification of horticultural crops because of the extreme economic efficiency over traditional techniques.

**Literature Cited**

Barney, D. and K. Hummer. 2012. Rhubarb: Botany, horticulture, and genetic resources. Hort. Rev. 40:147–182.

Beck, L.U. 2005. Pedanius Dioscorides of Anazarbus: De materia medica. Olms-Weidmann, Hildesheim, Germany.

Chen, F., A. Wang, K. Chen, D. Wan, and J. Liu. 2009. Genetic diversity and population structure of the endangered and medically important Rheum tanguticum (Polygonaceae) revealed by SSR markers. Biochem. Syst. Ecol. 37:613–621.

Chin, T.C. and H.W. Youngken. 1947. The cytotaxonomy of Rheum. Amer. J. Bot. 34:401–407.

Englund, R. 1983. Odlad rabarber och taxonomiska problem inom sla‘ket Rheum speciellt sektionen rhapontica. Report, Dept. Systematic Bot., Uppsala Univ., Uppsala, Sweden.

Estelle, R. 2009. Palmer’s original colonists. 26 Nov. 2012. <http://www.palmermuseum.org/palmer-history>.

Foust, C. 1992. Rhubarb: The wondrous drug. Princeton Univ. Press, Princeton, NJ.

Foust, C. and D. Marshall. 1991. Culinary rhubarb production in North America: History and recent statistics. HortScience 26:1360–1363.

Gilmore, B., N. Bassil, A. Nyberg, B. Knaus, D. Smith, D. Barney, and K. Hummer. 2013. Microsatellite marker development in peony using next generation sequencing. J. Amer. Soc. Hort. Sci. 138:64–74.

Gilmore, B., K. Hummer, and N. Bassil. 2011. DNA extraction protocols from dormant buds of twelve woody plant genera. J. Amer. Pomol. Soc. 65:201–207.
Hsu, H., Y. Chen, S. Shen, S. Hsu, C. Chen, and H. Chang. 1986. Oriental materia medica: A concise guide. Oriental Healing Arts Inst., Taipei, Taiwan, China.

Hu, Y., L. Wang, and Y. Li. 2011. Inter-simple sequence repeats (ISSR) primer screening and preliminary evaluation of genetic diversity in *Rheum tanguticum*. J. Medicinal Plants Res. 5:5907–5911.

Hu, Y., L. Wang, X. Xie, J. Yang, Y. Li, and H. Zhang. 2010. Genetic diversity of wild populations of *Rheum tanguticum* endemic to China as revealed by ISSR analysis. Biochem. Syst. Ecol. 38:264–274.

Jennings, T., B. Knaus, T. Mullins, S. Haig, and R. Cronn. 2011. Multiplexed microsatellite recovery using massively parallel sequencing. Mol. Ecol. Resources 11:1060–1067.

Jones, W. (ed.). 1956. *Pliny, natural history*. Loeb Classical Library, London, UK.

Knaus, B.J. 2011. Short read toolbox. 1 Oct. 2012. <http://brianknaus.com>.

Kuhl, J. and V. DeBoer. 2008. Genetic diversity of rhubarb cultivars. J. Amer. Soc. Hort. Sci. 133:587–592.

Liu, K. and S. Muse. 2005. PowerMarker: An integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129.

Morse, J. 1901. The new rhubarb culture. Orange Judd, New York, NY. National Center for Biotechnology Information. 2012. NCBI accession number: SRA054037. 29 June 2012. <http://www.ncbi.nlm.nih.gov/sra>.

Niu, B., L. Fu, S. Sun, and W. Li. 2010. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. BMC Bioinformatics 11:187.

Pantoja, A. and J. Kuhl. 2009. Morphologic variation in the USDA/ARS rhubarb germplasm collection. Plant Genet. Resour.; Characterization Util. 8:35–41.

Powell, W., G. Machray, and J. Provan. 1996. Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 1:215–222.

Ruirui, L., A. Wang, X. Tian, D. Wang, and J. Liu. 2010. Uniformity of karyotypes in *Rheum* (Polygonaceae), a species-rich genus in the Qinghai-Tibetan plateau and adjacent regions. Caryologia 63:82–90.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nat. Biotechnol. 18:233–234.

Sibly, E. 1790. Culpeper’s English physician; and complete herbal. C. Stalker, London, UK.

Turner, D.M. 1938. The economic rhubarbs: A historical survey of their cultivation in Britain. J. Royal Hort. Soc. 63:355–370.

U.S. Department of Agriculture. 2012. USDA, ARS, National Genetic Resources Program, Germplasm Resources Information Network. 9 Sept. 2013. <http://www.ars-grin.gov/cgi-bin/npgs/html/tax_search.pl>.

Wang, A., M. Yang, and J. Liu. 2005. Molecular phylogeny, recent radiation and evolution of gross morphology of the rhubarb genus *Rheum* (Polygonaceae) inferred from chloroplast DNA trnL-F sequences. Ann. Bot. (Lond.) 96:489–498.

Wang, X. 2011. Inter-simple sequence repeats (ISSR) molecular fingerprinting markers for authenticating the genuine species of rhubarb. J. Medicinal Plants Res. 5:758–764.

Zalapa, J., H. Cuevas, H. Zhu, S. Steffan, D. Senalik, E. Zeldin, B. McCown, R. Harbut, and P. Simon. 2012. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. Amer. J. Bot. 99:193–208.

Zhang, D., N. Chen, Y. Yang, Q. Zhang, and J. Liu. 2008. Development of 10 microsatellite loci for *Rheum tanguticum* (Polygonaceae). Conserv. Genet. 9:475–477.