Highly Informative Simple Sequence Repeat (SSR) Markers for Fingerprinting Hazelnut

K. Gürcan[1], S.A. Mehlenbacher[1] and N.V. Bassil[2]
[1] Department of Horticulture, Oregon State University, 4107 Ag. & Life Sciences Building, Corvallis, Oregon 97331, USA
[2] USDA-ARS National Clonal Germplasm Repository, 33447 Peoria Road, Corvallis, Oregon 97333, USA

Keywords: hazelnut, filbert, fingerprinting, microsatellite, polymorphic information content

Abstract

Simple sequence repeat (SSR) or microsatellite markers have many applications in breeding and genetic studies of plants, including fingerprinting of cultivars and investigations of genetic diversity, and therefore provide information for better management of germplasm collections. They are repeatable, co-dominant, highly polymorphic, and technically simple to use. We developed more than 150 SSR markers for hazelnut (*Corylus avellana*) and characterized them using a diverse set of 50 accessions. Some hazelnut SSR loci amplify other genera in the Betulaceae. In this study we identify two sets of a dozen markers each that are useful for fingerprinting hazelnut accessions. Loci were chosen based on high heterozygosity, PIC values, low frequency of null alleles, coverage of the hazelnut genome, and ease of scoring. The number of alleles per locus for these 24 primers ranged from 5-16. Polymorphic information content (PIC) values ranged from 0.50 to 0.87. Most loci segregated 1:1, 1:2:1 or 1:1:1:1 in our mapping population, allowing us to assign them to a linkage group.

INTRODUCTION

Conventional plant breeding has been used to develop improved cultivars of many species, including hazelnuts. Traditional methods can now be supplemented with molecular tools. These tools provide a better understanding of genetic diversity, and can be used for germplasm management, cultivar identification, marker-assisted selection, and parentage analysis. One commonly used marker type is the simple sequence repeat (SSR) or microsatellite, which consists of short tandem repeats in the DNA. SSRs are abundantly distributed throughout eukaryotic genomes, are highly polymorphic, locus-specific, can be shared by and are reproducible in different labs, and are suitable for semi-automation (fluorescent labeling, capillary electrophoresis, multiplexing, and computer-assisted size calling). However, the development of SSRs requires relatively high technology and expenditure. Several methods have been used for the development of SSRs, including microsatellite-enriched DNA libraries, PCR-based techniques such as ISSR-PCR, and public EST and genomic sequence databases. Large numbers of SSRs are now available, including >200 for strawberry (Govan et al., 2008), >5400 for cotton (Lacape et al., 2008). The genome database for the Rosaceae (2008) includes 335 for peach, 301 for apple, and 123 for almond. Once developed, SSRs can be used for multiple purposes and are often useful in related species. A small subset of markers is sufficient for fingerprinting and genetic diversity analysis. The chosen markers should be single-locus, highly polymorphic, easy to score, and give good coverage of the genome. The first 13 DNA typing multiplex sets became available for human research in 1998. Since then, many several "genotyping sets" have been selected; potato (Ghislain et al., 2004), cotton (Lacape et al., 2007), and strawberry (Govan et al., 2008). A tri-nucleotide (GAA) and two di-nucleotide (CA, GA) microsatellite-enriched libraries of the European hazelnut were constructed by Genetic Identification Services (GIS, Chatsworth, Calif.) were used to develop 43 polymorphic SSR loci (Bassil et al., 2005a, b; Boccacci et al., 2005) and have been used to fingerprint accessions and study diversity in collections (Gökirmak,

2008; Boccaci et al., 2006). 30 SSRs were placed on a linkage map (Mehlenbacher et al., 2006). In this report, we describe a genotyping set of 24 SSR loci suitable for fingerprinting hazelnut.

MATERIALS AND METHODS

We developed additional polymorphic SSR loci, 71 from our GA library and 13 from CA library. We also conducted simple sequence repeat-anchored PCR. This approach includes amplification of genomic DNA using ISSR primers, cloning and sequencing the resulting fragments, and then developing primers for SSRs in the fragments. We developed 23 SSRs from internal repeats and 50 from end-repeats. Additional polymorphic SSRs were developed from birch sequences submitted to the National Center for Biotechnology Information (NCBI). A few published birch SSR loci were also found to amplify hazelnut. A total 48 hazelnut accessions (Table 1) including the two parents of our mapping population were used for characterization. The transferability of markers was investigated using >35 species representing 6 genera in the family Betulaceae. In this study, we focus on choosing 24 genotyping set primers for fingerprinting hazelnut accessions. SSRs were selected based on five criteria: 1) single-locus, 2) coverage of the hazelnut genome, 3) high polymorphism, 4) ease of scoring, and 5) few null alleles.

RESULTS AND DISCUSSION

We selected 24 SSRs that meet all five criteria for fingerprinting (Tables 2, 3). Six of these (CAC-B617, CAC-B619, CAC-B657, CAC-B753, CAC-B719, CAC-B776) were suggested by Bassil et al. (2009) for use in fingerprinting. CaT-504 was also mentioned by Boccacci et al. (2006) as promising for use in fingerprinting. These loci produced 230 alleles in 48 accessions. The number of alleles per locus ranged from 4 to 16, with a mean of 9.6. The mean values for expected heterozygosity (H_e), observed heterozygosity (H_o), polymorphism information content (PIC) and null allele frequency (r) were 0.78, 0.75, 0.74, and 0.01, respectively. All primer pairs amplify a single locus. All linkage groups are represented although groups 2 and 7 merged during mapping. All 24 loci are highly polymorphic, as indicated by their H_e, H_o, and PIC values. The loci are easy to score, as they show minimal stuttering. Non-template addition of an extra adenine nucleotide to the end of the true allele makes size calling difficult at some loci. We chose loci that were easy to score. We excluded primers showing any artifact, even if they are very polymorphic, included only primers producing very clear distinct bands. Our 24 loci also show a low frequency of null alleles, and the majority showed good transferability to other members of the Betulaceae. Allele size calling can vary slightly from one run to the next. We include 'Barcelona' and 'Negret' as controls in our fingerprinting. Allele sizes of hazelnut accessions maintained by the USDA-ARS National Clonal Germplasm Repository are available at http://www.ars-grin.gov/cgi-bin/npgs/crop/mrk_list.pl?cropno=105.

CONCLUSION

We hope that this well-characterized "genotyping set" for and the availability of reference genotypes and allele size data will be useful to hazelnut researchers around the world. We look forward to sharing data with our colleagues, comparing fingerprints of accessions, correcting mistakes in collections, and studying genetic diversity in hazelnut.

Literature Cited

Bassil, N.V., Botta, R. and Mehlenbacher, S.A. 2005a. Additional microsatellite markers of the European hazelnut. Acta Hort. 686:105-110.

Bassil, N.V., Botta, R. and Mehlenbacher, S.A. 2005b. Microsatellite markers in hazelnut: Isolation, characterization and cross-species amplification. J. Amer. Soc. Hort. Sci. 130:543-549.

Bassil, N.V., Postman, J., Hummer, K., Botu, M. and Sezer, A. 2009. SSR fingerprinting panel verifies identities of clones in backup hazelnut collection of USDA genebank.

Acta Hort. 845:95-102.
Boccacci, P., Akkak, A., Bassil, N.V., Mehlenbacher, S.A. and Botta, R. 2005. Characterization and evaluation of microsatellite loci in European hazelnut (*C. avellana*) and their transferability to other *Corylus* species. Molec. Ecol. Notes. 5:934-937.
Boccacci, P., Akkak, A and Botta, R. 2006. DNA typing and genetic relations among European hazelnut (*Corylus avellana* L.) cultivars using microsatellite markers. Genome 49:598-611.
Genome Database for Rosaceae. 2008. www.bioinfo.wsu.edu/gdr.
Ghislain, M., Spooner, D.M., Rodríguez, F., Villamón, F., Núñez, J., Vásquez, C., Waugh, R. and Bonierbale, M. 2004. Selection of highly informative and user-friendly microsatellites (SSRs) for genotyping of cultivated potato. Theor. Appl. Genet. 108:881-890.
Govan, C.L., Simpson, D.W., Johnson, A.W., Tobutt, K.R. and Sargent, D.J. 2008. A reliable multiplexed microsatellite set for genotyping *Fragaria* and its use in a survey of 60 *F.* × *ananassa* cultivars. Mol. Breeding DOI 10.1007/s11032-008-9206-2.
Lacape, J.M., Dessauw, D., Rajab, M., Noyer, J.L. and Hau, B. 2007. Microsatellite diversity in tetraploid *Gossypium* germplasm: assembling a highly informative genotyping set of cotton SSRs. Mol. Breeding 19:45-58.
Mehlenbacher, S.A., Brown, R., Nouhra, E., Gokirmak, T., Bassil, N.V. and Kubisiak, T. 2006. A genetic linkage map for hazelnut (*Corylus avellana* L.) based on RAPD and SSR markers. Genome 49:122-133.

Tables

Table 1. Fifty *Corylus avellana* genotypes used to characterize SSR loci.

Accession	Genotype	Origin
PI 557060	OSU 54.039	Turkey
PI 617269	OSU 556.027	Turkey-Istanbul
	OSU 408.040	Univ. Minnesota
	Albania 55	Albania
	Alli	Estonia
PI 557108	Artellet	Spain
PI 557050	Aurea	Germany
PI 557122	B-3	Macedonia
PI 557156	Barcelonner Zellernuss	England
PI 557037	Barcelona	Spain-Tarrogona
PI 557114	Bergeri	France
PI 557089	Buttner's Zellernuss	Germany
PI 557033	Casina	Spain-Asturias
PI 557049	Contorta	England
PI 557039	Cosford	England
PI 557006	Cutleaf	England
PI 557423	Des Anglais	France
PI 557099	DuChilly	England
PI 557080	COR 187	Finland
PI 557047	Fusco Rubra	Germany
PI 557042	Gasaway	USA-Washington
PI 557029	Gem	USA-Oregon
	Gulshisvela	Georgia
PI 557191	Gunslebert	Germany
PI 557085	Gustav's Zellernuss	Germany
PI 557027	Hall's Giant	Germany
	Iannusa Racinante	Italy-Sicily
PI 271105	Imp. de Trebizonde	Turkey
PI 557090	Kadetten	Germany
PI 339723	Mortarella	Spain
PI 270340	Negret	Spain-Tarrogona
	OSU 495.049	S. Russia
PI 304632	Palaz	Turkey
PI 271110	Pellicule Rouge	France
PI 557048	Pendula	England
PI 557167	Ratoli	Spain-Tarragona
PI 271280	Rode Zeller	Netherlands
PI 557171	Romische Nuss	Unknown
	OSU 26.072	Russia-North C.
	OSU 681.078	Russia-Moscow
PI 557103	Sant Jaume	Spain-Tarragona
PI 557166	Simon	Spain-Tarragona
PI 304634	Tombul Ghiaghli	Turkey-Greece
PI 296206	Tonda Bianca	Italy-Campania
PI 296207	Tonda di Giffoni	Italy-Campania
PI 557075	Tonda G. d. Langhe	Italy-Piemonte
PI557025	Tonda Romana	Italy-Viterbo
PI 617239	Tapparona di SCC	Italy-Liguria
None	OSU 252.146	USA-Oregon
None	OSU 414.062	USA-Oregon

Table 2. SSR loci for fingerprinting hazelnut accessions.

Locus	Primers (5'-3')	Repeat motif	Size range	T_a (ºC)
CAC-B749	F-GGCTGACAACACAGCAGAAA R-TCGGCTAGGGTTAGGGTTTT	$(TC)_{12}$	200-210	60
CAC-B029b	F-CAATTTACACCTCAGGGAAGAG R-AAGTTCACCCAAGAAATCCAC	$(GA)_{13}$	116-136	58
CAC-A635	F-TTACCCAATGGATGATGGACTAGCATT R-GGATCTGTGGTTGGCTTTTTGGTACTAT	$(CT)_{16}$	225-241	67
CAC-B005	F-CAAACTTATGATAGGCATGCAA R-TGTCACTTTGGAAGACAAGAGA	$(GA)_{22}$	277-297	62
CAC-B753	F-AAGGGTTGTTACCCATGCAC R-GGTGCATTTAGTGCTTCTGG	$(GA)_{15}$	224-254	60
CAC-B733	F-CACCCTCTTCACCACCTCAT R-CATCCCCTGTTGGAGTTTTC	$(TC)_{15}$	161-183	60
CAC-B751	F-AGCTGGTTCTTCGACATTCC R-AAACTCAAATAAAACCCCTGCTC	$(GA)_{15}$	141-153	60
CaT-C504	F-CGCCATCTCCATTTCCCAAC R-CGGAATGGTTTTCTGCTTCAG	$(CT)_{18}$	152-168	60
CAC-B619	F-AGTCGGCTCCCCTTTTCTC R-GCGATCTGACCTCATTTTTG	$(TC)_{21}$	146-180	60
CAC-B791	F-CACCAGGACCCTGATACCAT R-TCCACAATGATTTTGTGAAAAC	$(AG)_{14}$	205-241	60
CAC-C115	F-CATTTTCCGCAGATAATACAGG R-GTTTCCAGATCTGCCTCCATATAAT	$(TAA)_5(GAA)_{12}$	173-225	60
CAC-K74-2-36	F-GGAAGGTGAGAGAAATCAAGT R-AGGCATCAGTTCATCCAA	$(AGG)_7$	333-345	55
CAC-A014a	F-GGTTTGTTACAGAAATTCAGACG R-GCGTGTGGTTAATGTTTTCTTT	$(CA)_{13}$	209-225	60
CAC-B625	F-CGCAAGTCATTGCACATTTT R-GTGTGCTGTGCTCCTTTGAA	$(TC)_{13}$	237-283	67
CAC-B776	F-TGTATGTACACACGGAGAGAGAGA R-TGAGGGGAAGAGGTTTGATG	$(GA)_{17}$	134-148	60
CAC-A614	F-TGGCAGAGCTTTGTCAGCTT R-GCAGTGGAGGATTGCTGACT	$(TC)_{17}(CA)_{10...}(CA)_6$	125-156	60
CAC-B617	F-TCCGTGTTGAGTATGGACGA R-TGTTTTTGGTGGAGCGATG	$(GA)_{15}$	280-298	60
CAC-K89-2-3	T-GAATACTTGCGACTCTAATCT R-TTTCTTTCTTCACGACCC	$(AAT)_8$	320-342	54
CaT-B507	F-CTA AGCTCACCAAGAGGAAGTTGAT R-GCTTCTGGGTCTCCTGCTCA	$(CT)_{16...}(CT)_5$	180-190	62
CAC-B719	F-CCGACTCCTTAGCATCCTTG R-AGCTAGTCTAAACATCAATAGAGCAA	$(TC)_{18}$	283-297	60
CAC-B706	F-TGCATGAAATGGAATCACAGA G-AGCAAAGAGGTAAGCAAATTCA	$(CT)_{28}$	168-206	67
CAC-B502	F-CTCATGACTGCCCATTTCTCG R-AGGCATGCAGGCTTCACAC	$(GA)_{14}$	185-214	62
CAC-B657	F-GAGAGTGCGTCTTCCTCTGG R-AGCCTCACCTCCAACGAAC	$(AG)_{15}$	210-228	60
CAC-K74-1-19	F-AAGCAAGAAAGGGATGGT R-CTTACAGATAAATGGCTCAAA	$(AAAT)_5$	226-248	54

Table 3. Characterization of hazelnut SSR loci.

Locus	No. Alleles	H_e	H_o	PIC	r	LG	Amplification (%) in	
							Corylus	Other Genera
CAC-B749	5	0,59	0,62	0,50	-0,02	1	100	49
CAC-B029b	13	0,85	0,83	0,83	0,00	1	100	11
CAC-A635	11	0,70	0,70	0,66	0,00	1	100	30
CAC-B005	7	0,73	0,79	0,70	0,00	2--7	100	98
CAC-B753	12	0,84	0,78	0,82	0,03	2--7	85	9
CAC-B733	8	0,69	0,68	0,64	0,00	2--7	86	28
CAC-B751	7	0,80	0,78	0,77	0,01	2--7	96	65
CaT-C504	5	0,76	0,63	0,72	0,00	2--7	100	98
CAC-B619	16	0,88	0,88	0,87	0,00	3	100	74
CAC-B791	13	0,81	0,74	0,80	0,04	3	82	30
CAC-C115	10	0,78	0,71	0,75	0,00	4	92	40
CAC-K74-2-36	5	0,66	0,64	0,60	0,01	4	100	95
CAC-A014a	8	0,73	0,67	0,69	0,00	5	99	28
CAC-B625	14	0,8518	0,8	0,835	0,03	5	50	0
CAC-B776	7	0,71	0,60	0,67	0,07	6	100	61
CAC-A614	15	0,8642	0,82	0,8523	0,00	6	100	38
CAC-B617	9	0,80	0,76	0,77	0,02	8	100	72
CAC-K89-2-3	9	0,76	0,68	0,73	0,04	8	100	0
CaT-B507	12	0,87	0,85	0,86	0,02	9	100	84
CAC-B719	8	0,77	0,68	0,73	0,05	9	92	47
CAC-B706	14	0,88	0,86	0,87	-0,02	10	100	20
CAC-B502	11	0,79	0,83	0,77	0,00	10	96	28
CAC-B657	8	0,84	0,98	0,83	-0,07	11	100	49
CAC-K74-1-19	4	0,67	0,78	0,60	-0,07	11	100	93
Total	230	18,64	18,09	17,87	0,15			
Mean	9,58	0,78	0,75	0,74	0,01			

Expected heterozygosity (H_e), observed heterozygosity (H_o), polymorphic information content (PIC), frequency of null alleles (r), and linkage group (LG).
Transferabilitity: Percentage amplification in *Corylus* species and in non-*Corylus* species in the Betulaceae