Plant Molecular Biology Reporter **22:** 23–28, March 2004

Genetic Resources

Construction of a Sugar Beet BAC Library From a Hybrid With Diverse Traits

J. MITCHELL McGRATH*, R. SCOTT SHAW, BENILDO G. de los REYES[1] and JOHN J. WEILAND[2]

USDA-ARS Sugarbeet and Bean Research Unit, Department of Crop and Soil Sciences, 494 Plant and Soil Sciences Building, Michigan State University, East Lansing, MI 48824-1325; [1]Deptartment of Biological Sciences, 5751 Murray Hall, University of Maine, Orono, ME 04469-5751; [2]USDA-ARS, Northern Crop Science Lab, 1307 18th St N, Fargo, ND 58105-5677

Abstract. A bacterial artificial chromosome (BAC) library of the 750-Mbp sugar beet genome represented in hybrid US H20 was constructed from *Hin*d III–digested DNA, with an average insert size of 120 kbp. US H20 is a variety grown in the eastern United States. It exhibits heterosis for emergence and yield, presumably because of its hybridity between eastern and western US germplasm sources. Filter arrays were used to assess the abundance and distribution of particular nucleotide sequences. An rRNA gene probe found that 1.2% of the library carried sequences similar to these highly repetitive and conserved sequences. A simple sequence repeat element $(CA)_8$ thought to be predominantly distributed throughout centromere regions of all chromosomes was present in 1.7% of clones. For more than half of the 28 randomly chosen expressed sequence tags (ESTs) used as probes, a higher-than-expected number of single-copy hybridization signals was observed. Assuming 6× genome coverage, this suggests that many duplicate genes exist in the beet genome.

Key words: BAC library, gene copy number, genetic resource, sugar beet, physical mapping

Abbreviations: BAC, bacterial artificial chromosome; CMS, cytoplasmic male sterility; EST, expressed sequence tag; SNP, single nucleotide polymorphism.

Introduction

Breeding and genomics must integrate 3 lines of investigation in order to predict agronomic performance on the basis of DNA sequences. First, agronomically relevant plant populations for genetic analyses, including progeny sets from controlled crosses, should define agronomic loci in the existing germplasm. Second, facile means to correlate genes with phenotypes, guided by heredity (e.g., molecular marker mapping) and gene expression profiling, should augment gene locus information with putative biochemical mechanisms and should indicate where

*Author for correspondence. e-mail: mitchmcg@msu.edu; fax: 1-517-337-6782; ph: 1-517-355-0271 ext. 207.

opportunities exist for genetic intervention. Third, and perhaps most challenging, methods of agronomic trait phenotyping that are easy to conduct and reproduce and that predict crop responses under highly variable field conditions should enable greater analytical precision and increase selection intensity for problematic traits such as those with low heritability.

Germplasm resources are well developed for sugar beet (Doney, 1995). The majority of beet germplasm is self-sterile and highly heterozygous, and inbred lines would facilitate genomic analyses, but are not yet widely available. Genomic resources for sugar beet (*Beta vulgaris* L.) include molecular genetic maps (Halldén et al., 1998; Schneider et al., 2002), 21,523 *B. vulgaris* expressed sequence tags (ESTs; as of December 2003) (http://www.ncbi.nlm.nih.gov), and large insert libraries (Gindullis et al., 2001; Hohmann et al., 2003). The ESTs represent mRNAs expressed in seedlings, roots, leaves, flowers, and inflorescences and various pathogen-infected beets; most represent a unigene set of more than 10,000 ESTs covering the important developmental stages of beet growth (Herwig et al., 2002).

Here we report the construction and availability of a BAC library for sugar beet US H20 as a genetic resource to the plant community. BAC libraries are preferred for creating an ordered array of genomic DNA fragments that represent chromosomes (e.g., physical mapping) because of their relatively large size and the ease of manipulating and screening BAC clones (Luo et al., 2003; Peters et al., 2003).

Materials and Methods

For BAC library construction, we used sugar beet US H20 (Coe and Hogaboam, 1971a), seed lot WC990379, from a resynthesis of US H20 in Oregon in 1999 and deposited as voucher seed in the National Plant Germplasm System (GRIN) as PI631354.

Seeds were sown in flats at high density (>1 seed/cm^2) in commercial soilless mix for 3 wk in the greenhouse with supplemental lighting. Young leaves (<0.5-cm length, 300 g total harvest) were harvested, and nuclei were isolated, embedded in low-melting-point agarose, partially digested with *Hin*d III, and size fractionated on agarose gels twice. Fragments were ligated into vector pECBAC1 as described (Frijters et al., 1997). BAC library construction and subsequent large-scale manipulations (e.g., spotting the library on nylon filters and pooling the library in a PCR-ready format) were performed by Amplicon Express, Pullman WA (library designated SBA).

Subsequent analyses followed the manufacturer's protocols for filter hybridization (Hybond-N+, Amersham Biosciences) with ^{32}P-labelled probes. All BAC clones were represented twice on the filters, and positive signals were unambiguous. A 774-bp putative mitochondrial and a 1190-bp chloroplast hybridization probe used to assess possible organelle contamination were amplified and gel-purified from US H20 genomic DNA. This was done by using primers designed from a sugar beet cytoplasmic male sterility (CMS)–associated mitochondrial sequence (Z12825; forward primer 5′-ATTTGCCGGCTATGTTTCTTCT, reverse primer 5′-GTTGGCCTTTTTACTTTCCCTA) and from the beet chloroplast *pet*G,

*ycf*7, *psb*E, *psb*F, ORF32 gene cluster (X87636; forward 5′-TTAGACTCAGCA-ATCGGGTTCATC, reverse 5′-ATTCGGATTCGGTTTTGTTTATC). Centromeric-associated repeat units were similarly amplified and gel-purified by using $(CA)_8$ and telomere units from $(TTTAGGG)_7$ as the primers. All other probes were excised from their cognate plasmids.

Results and Discussion

We are publicly releasing a germplasm resource BAC library for sugar beet, constructed from a former commercial hybrid expected to contain high levels of heterozygosity relevant to agronomic performance. This library was constructed (1) to isolate genomic DNA corresponding to genes involved in enhanced germination and field emergence of sugar beet (e.g., seedling vigor; de los Reyes et al., 2003), (2) to construct a physical map of sugar beet, and (3) to develop single nucleotide polymorphism (SNP) markers at loci of agronomic importance (Schneider et al., 2001). Constructing a physical map and developing SNP markers are complimentary in that genic regions are targeted with ESTs.

The BAC library, constructed from *Hin*d III–digested sugar beet hybrid US H20, has 6.1-fold genome coverage (36,864 clones) of the ca. 750-Mbp sugar beet genome (Arumuganathan and Earle, 1991). Insert size averaged 120 kbp among 12 randomly chosen clones ranging from 90-160 kbp and is sufficient to mathematically recover any sequence of interest with greater than 99% probability (Clark and Carbon, 1976). Filter arrays containing all clones were prepared and used to assess the abundance and distribution of particular types of nucleotide sequences. Using rRNA gene probe pTA71 (Gerlach and Bedbrook, 1979), 1.2% (450 clones, Table 1) of the library carried sequences similar to these highly conserved sequences located on chromosome 1 of the Butterfass trisomic series (Schondelmaier and Jung, 1997). A simple sequence repeat element $(CA)_8$ thought to be predominantly distributed throughout centromere regions of all chromosomes (Schmidt and Heslop-Harrison, 1996) was present in 1.7% of clones. A probe for the plant telomere canonical sequence (TTTAGGG) only hybridized with 7 BAC clones. However, this region at the end of chromosomes is difficult to clone and was not expected to be well represented in this library, and the positive signals may represent interstitial relics from previous inversion events. Organelle DNA (plastid, mitochondria) contamination was assessed with organelle-specific DNA probes. Chloroplast DNA contamination was greater than CMS-associated mitochondrial DNA (1.7% of clones vs. 0.01%, respectively), although other potentially less variable mitochondrial probes were not tested.

Twenty-eight randomly chosen ESTs were screened against nylon filter arrays of the BAC library (Table 1). In this limited sample, all tested probes showed a hybridization signal. The number of gene copies similar to a particular EST in the beet genome was estimated from the number of hybridization signals as a first approximation. For more than half of the ESTs used as probes, a higher-than-expected number of hybridization signals was observed for those of a single copy sequence. This finding suggests that many duplicate genes exist in the beet genome or representation in the BAC library is unequal. The latter appears likely in some cases because only a single signal was observed for one clone (AW697745).

Table 1. Representation of sequences similar to genes and other landmarks in the sugar beet genome of hybrid US H20, representing 6.1-fold coverage distributed among 36,864 BAC clones.

Genbank Accession No.	Putative Function of Probe	No. Positive Signals on Filter	Percent of Library	Estimated Copy No., In
B1543270	14-3-3-like protein	7	0.019	1.15
B1543560	ABC transporter	7	0.019	1.15
B1543538	Adenine triphosphatase	14	0.038	2.30
AW697786	Alcohol dehydrogenase	5	0.014	0.82
B1095948	Allergen	8	0.022	1.31
B1643109	Aquaporin	18	0.049	2.95
BF011005	Beta amyrin synthase	5	0.014	0.82
B1096069	Calmodulin	7	0.019	1.15
BE590278	Cysteine protease	4	0.011	0.66
B1543290	Enolase	3	0.008	0.49
AF310017	Germinlike protein	40	0.109	6.56
B1095991	Glyceraldehyde 3P dehydrogenase	18	0.049	2.95
B1543424	Heat shock protein	7	0.019	1.15
AW697750	Heat shock protein 81-2	2	0.005	0.33
BI 543276	Hexokinase	2	0.005	0.33
B1095900	Hydroxymethyltransferase	33	0.090	5.41
B1095941	Isocitrate lyase	6	0.016	0.98
B1073206	Malate dehydrogenase	11	0.030	1.80
BQ060614	MAP kinase-like protein	2	0.005	0.33
AW697745	PEP mutase	1	0.003	0.16
B1096032	Phosphofructokinase	12	0.033	1.97
BI 073208	Pyruvate dehydrogenase-1	13	0.035	2.13
81096005	Pyruvate dehydrogenase-2	28	0.076	4.59
B1073233	Ribulose phosphate 3-epimerase	12	0.033	1.97
81643066	RUBISCO	14	0.038	2.30
B1543240	Sucrose synthase	39	0.106	6.39
B1073142	UDP-glucose glucosyltransferase	15	0.041	2.46
B1096068	UDP-glucose pyrophosphorylase	8	0.022	1.31
pTA71	Ribosomal RNA repeat unit	450	1.221	
Z12825	Mitochondrial CMS-associated	2	0.005	
X87636	Chloroplast DNA	629	1.706	
$(CA)_8$	Centromeric-associated repeat unit	619	1.679	
$(TTTAGGG)_7$	Telomeric repeat unit	7	0.019	

However, the estimated number of germinlike protein genes (AF310017) is similar to the number of ESTs recovered from GenBank and estimated by means of Southern blot analyses, suggesting good representation of some genome regions.

In sugar beet (*Beta vulgaris* L., Chenopodiaceae), public breeding programs have released numerous germplasm lines (Doney, 1995). Each release is selected for one or more unique characters or combinations of characters that are used by commercial breeders for variety development. In the United States and northern Europe, varieties are exclusively hybrids developed from a single original source of CMS (Owen, 1945). Tools constructed for genomic analyses toward eventual breeding applications should recognize the hybrid nature of the crop. Hybrid US H20 was released by the United States Department of Agriculture's Agricultural

Research Service (USDA-ARS) for commercialization in 1972 and quickly became a favored variety in the eastern United States because of its consistently high yields. These high yields of hybrid US H20 are attributed to its early season vigor and its ability to persist against seedling diseases, particularly Aphanomyces seedling disease caused by *Aphanomyces cochlioides*. US H20 combines discrete eastern and western US germplasm pools through an Aphanomyces-resistant pollinator (SP6822, PI615525; Coe and Hogaboam, 1971b) and a 2-way hybrid between western US CMS and maintainer seed parent lines (EL44CMS, PI590856 [Hogaboam and Schneider, 1982a]; EL45/2, PI590720 [Hogaboam and Schneider, 1982b]). Western US germplasm is generally characterized as resistant to curly top virus.

The US H20 sugar beet BAC library has no user restrictions. Clones are available individually, as a library in liquid culture or nylon filters or as pools of BAC clones. The library will be used to create a physical map, and local landmarks of interest to the plant community can be incorporated into the physical mapping strategy.

Acknowledgments

We thank Drs Lee Panella and Bob Lewellen for helpful discussions; Susan Myers, Quanzhou Tao, Robert Bogden, and Keith Stormo for technical assistance; and member companies of the Beet Sugar Development Foundation for financial support.

References

Arumuganathan K and Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9: 208-18.

Clarke L and Carbon J (1976) A colony bank containing synthetic Col El hybrid plasmids representative of the entire *E. coli* genome. Cell 9: 91-100.

Coe GE and Hogaboam GJ (1971a) Registration of USH20 sugar beet. Crop Sci 11: 942.

Coe GE and Hogaboam GJ (1971b) Registration of sugarbeet parental line SP 6322-0. Crop Sci 11: 947.

de los Reyes BG, Myers SJ, and McGrath JM (2003) Differential stress-induction of glyoxylate cycle enzymes as a marker for seedling vigor in sugar beet (*Beta vulgaris*). Mol Gen Genomics 269: 692-8.

Doney DL (1995) USDA-ARS sugar beet releases. J Sugar Beet Res 32: 229-57.

Frijters ACJ, Zhang Z, van Damme M, Wang G-L, Ronald PC, and Michelmore RW (1997) Construction of a bacterial artificial chromosome library containing large *Eco*RI and *Hin*dIII genomic fragments of lettuce. Theor Appl Genet 94: 390-9.

Gerlach WL and Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucl Acids Res 7: 1869-85.

Gindullis F, Dechyeva D, and Schmidt T (2001) Construction and characterization of a BAC library for the molecular dissection of a single wild beet centromere and sugar beet (*Beta vulgaris*) genome analysis. Genome 44: 846-55.

Halldén C, Ahrén D, Hjerdin A, Säll T, and Nilsson N-O (1998) No conserved homoeologous regions found in the sugar beet genome. J Sugar Beet Res 35: 1-13.

Herwig R, Schulz B, Weisshaar B, Hennig S, Steinfath M, Drungowski M, Stahl D, Wruck W, Menze A, O'Brien J, and others (2002) Construction of a 'unigene' cDNA clone set by oligonucleotide fingerprinting allows access to 25,000 potential sugar beet genes. Plant J 32: 845-57.

Hogaboam GJ and Schneider CL (1982a) Registration of EL44 and EL44CMS sugarbeet parental lines. Crop Sci 22: 700.

Hogaboam GJ and Schneider CL (1982b) Registration of EL45/2 sugarbeet parental lines. Crop Sci 22: 700.

Hohmann U, Jacobs G, Telgmann A, Gaafar RM, Alam S, and Jung C (2003) A bacterial artificial chromosome (BAC) library of sugar beet and a physical map of the region encompassing the bolting gene B. Mol Genet Genomics 269: 126-36.

Luo MC, Thomas C, You FM, Hsiao J, Shu OY, Buell CR, Malandro M, McGuire PE, Anderson OD, and Dvorak J (2003) High-throughput fingerprinting of bacterial artificial chromosomes using the SNaPshot labeling kit and sizing of restriction fragments by capillary electrophoresis. Genomics 82: 378-89.

National Center for Biotechnology Information Home Page. Accessed 22 Mar. 2004 <http://www.ncbi.nlm.nih.gov>.

Owen FV (1945) Cytoplasmically inherited male-sterility in sugar beets. J Ag Res 71: 423-40.

Peters JL, Filip Cnudde F, and Gerats T (2003) Forward genetics and map-based cloning approaches. Trends Plant Sci 10: 484-91.

Schmidt T and Heslop-Harrison JS (1996) The physical and genomic organization of microsatellites in the sugar beet. Proc Nat Acad Sci USA 93: 8761-5.

Schneider K, Weisshaar B, Borchardt DC, and Salamini F (2001) SNP frequency and allelic haplotype structure of *Beta vulgaris* expressed genes. Mol Breeding 8: 63-74.

Schnieder K, Schafer-Pregl R, Borchardt DC, and Salamini F (2002) Mapping QTLs for sucrose content, yield and quality in a sugar beet population fingerprinted by EST-related markers. Theor Appl Genet 104: 1107-13.

Schondelmaier J and Jung C (1997) Chromosomal assignment of the nine linkage groups of sugar beet (*Beta vulgaris* L.) using primary trisomics. Theor Appl Genet 95: 590-6.