Variability for Restriction Fragment-length Polymorphisms (RFLPs) and Relationships among Elite Commercial Inbred and Virtual Hybrid Onion Populations

Joseph J. King,¹ James M. Bradeen,² and Michael J. Havey³

U.S. Department of Agriculture, Agricultural Research Service, Department of Horticulture, 1575 Linden Drive, University of Wisconsin, Madison, WI 53706

ADDITIONAL INDEX WORDS. *Allium cepa*, genetic distance, UPGMA

ABSTRACT. Nuclear RFLPs were used to estimate relationships among 14 elite commercial inbreds of bulb onion (*Allium cepa*) from Holland, Japan, and the United States. Variability for known alleles at 75 RFLP loci and 194 polymorphic fragments revealed by 69 anonymous cDNA probes and a clone of alliinase were scored to yield genetically characterized and uncharacterized data sets, respectively. The inbred onion populations possessed more than two alleles at 20 of 43 (46%) codominant RFLP loci. Relationships among the inbreds were estimated by cluster analysis of simple-matching (genetically characterized data) and Jaccard (genetically uncharacterized data) coefficients using the unweighted pair group method and agreed with known pedigrees. RFLPs confidently distinguished among elite inbreds within and between specific market classes. RFLP profiles for virtual hybrids were computer-generated by combining gametic arrays among inbreds of the same market class and analyzed as described above. Allelic and genetically uncharacterized RFLPs confidently distinguished among these hybrids, even though heterozygosity for many markers produced a majority of monomorphic fragments. We randomly sampled decreasing numbers of RFLPs from the complete data sets and calculated simple-matching and Jaccard distances, noting the numbers of probes that were unable to distinguish any two inbreds or hybrids. As few as 10 polymorphic probe-enzyme combinations distinguished among all the inbreds and samples of 20 genetically characterized or 10 genetically uncharacterized clones distinguished all the virtual hybrids. This study demonstrated that the previously reported few RFLPs observed among open-pollinated (OP) onion populations were due to the highly heterozygous nature of the OP population.

An understanding of the genetic variation among and within germplasm resources can enhance the efficiency of evaluating, conserving, and using this diversity for plant improvement. The degree of variation among individuals or populations is expressed as genetic distance. Morphological variation has been used traditionally to distinguish taxa at many levels, estimate genetic distances, and select parents for breeding programs (Dudley, 1994). While morphological differences are readily observable, they can be a misleading measure of diversity due to uncharacterized genetics and variable environmental effects on trait expression. For this reason, molecular markers are becoming the standard approach to measuring genetic distance.

Variability at molecular-marker loci can be used to group lines or populations based on genetic distance, select desirable genotypes in breeding programs, develop genetic maps, identify and locate quantitative trait loci, fingerprint lines for protection of intellectual property rights, and check purity of inbreds or hybrids (Dudley, 1995). Genetic relationships estimated by RFLPs have been consistent with relationships based on breeding behavior and pedigrees in numerous species, including *Zea mays* L. (Smith et al., 1990), *Sorghum bicolor* L. (Cui et al., 1995; Vierling et al., 1994), *Lycopersicon esculentum*

L. (Miller and Tanksley, 1990), and *Brassica oleracea* L. (dos Santos et al., 1994). In addition to confirming previously accepted relationships, RFLPs have revealed differences among closely related germplasms, such as restorer and maintainer lines of cytoplasmic male sterility in sunflower, *Helianthus annuus* L. (Gentzbittel et al., 1994). Cultivar identification and quality assurance of commercial seed lots are growing areas for application of molecular markers.

The identification and mapping of RFLPs have been reported for many economically important plants (O'Brien, 1993). For some crops, interspecific crosses were used to maximize the number of polymorphisms segregating in a single cross (e.g., Chittenden et al., 1994; Helentjaris et al., 1986; Kennard et al., 1994; Reinisch et al., 1994). An obvious disadvantage of using interspecific crosses is that only a limited number of the identified markers will exist in elite germplasm, restricting their use in plant improvement. We identified random RFLPs among open-pollinated (OP) onion populations and observed relatively few (mean of 1.9) polymorphic fragments per polymorphic probe-enzyme combination, due to either a restricted genetic background or high heterozygosity within OP onion populations that masked variability (Bark and Havey, 1995). A low-density genetic map of bulb onion (*Allium cepa* L.) composed of primarily RFLPs was constructed using a segregating family from a cross of two commercial inbreds (King et al., 1998). In the present study, we assess variability for genetically characterized and uncharacterized RFLPs among elite onion inbreds to estimate allelic diversity and establish the efficacy of this marker class to distinguish among commercially important germplasms.

Materials and Methods

Fourteen elite inbreds were acquired from commercial entities (Table 1). Each inbred had in its pedigree at least one generation of

Received for publication 17 Feb. 1998. Accepted for publication 16 June 1998. Names are necessary to report factually on available data; however, the USDA neither guarantees nor warrants the standard of the product, and the use of the name by USDA implies no approval of the product to the exclusion of others that may also be suitable. We gratefully acknowledge the technical help of Mark Petrashek and the financial support of USDA-NRICGP 92-37300-7546, Rio Colorado, Sakata, Shippo, SunSeeds, and S&G/Sandoz. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked *advertisement* solely to indicate this fact.
¹Research scientist, Seminis Seed Co., Woodland, CA.
²Postdoctoral research geneticist.
³Research geneticist and associate professor of horticulture; to whom reprint requests should be addressed.

self pollination, guaranteeing a maximum of two alleles per locus. Two additional inbreds (AC43 and BYG15-23) were previously described (Bradeen and Havey, 1995), were parents of our mapping family (King et al., 1998), and were included to identify alleles at RFLP loci. DNA was extracted from lyophilized leaf tissue of ≈ 50 seedlings and purified through cesium chloride gradients, and concentrations were determined spectrophotometrically (Havey, 1991). Digestion and blotting of DNAs were described previously (Bark and Havey, 1995). Labeling of probes, hybridization, washing, and autoradiography were as reported by King et al. (1998). Normal male-fertile (N) and male-sterile (S) cytoplasms were distinguished as described by Havey (1993). In total, 241 cDNA clones (Bark and Havey, 1995; King et al., 1998) were hybridized to digests with one restriction enzyme that revealed polymorphisms between AC43 and BYG15-23. Alleles at 75 RFLP loci segregating in the BYG15-23 x AC43 cross (King et al., 1998) were scored among the elite inbreds to produce the genetically characterized data set. Polymorphic fragments revealed by 69 anonymous cDNA clones and a cDNA clone of alliinase (van Damme et al., 1992), none of which segregated in the BYG15-23 x AC43 cross, were scored to generate the genetically uncharacterized data set (probe-enzyme combinations and fragment sizes are available from M.J. Havey).

The array of gametic genotypes was computer generated for each inbred and combined to produce single-cross nonreciprocal virtual hybrids. For cultivar protection or seed-purity estimates, we assumed that DNA would be pooled from numerous hybrid plants. Therefore, heterozygous loci in the inbred parents would appear monomorphic in pooled DNA of the hybrid, even though individual plants would be polymorphic for these RFLPs.

Genetic distances and relationships among inbreds and virtual hybrids were estimated by NTSYS version 1.7 (NTSYS-pc 1.70, Exeter Software, Setauket, N.Y.) using unweighted pair group method (UPGMA) of cluster analyses of the simple-matching and Jaccard coefficients. For simple matching, the presence and absence of fragments are scored as matches; the Jaccard coefficient scores only the shared presence of fragments. For the genetically characterized RFLP data set, two alleles at a locus were known and simple-

matching is an appropriate measure of similarity. For the genetically uncharacterized RFLP data set, the Jaccard is a more conservative coefficient because we cannot be sure that the absence of a fragment in two populations arose from the same event. We randomly generated 10 independent data sets of RFLPs revealed with 10, 20, 30, 40, 50, and 60 genetically characterized and uncharacterized polymorphic probe-enzyme combinations. We calculated simple-matching (genetically characterized) and Jaccard (genetically uncharacterized) distances, noting when marker sets were unable to distinguish any two inbred or hybrid populations.

Results and Discussion

VARIABILITY FOR RFLPS AMONG THE ELITE INBRED AND VIRTUAL HYBRID ONION POPULATIONS. We chose commercial inbred lines that had at least one generation of self-pollination, theoretically ensuring that no more than two alleles would be present at each locus. Segregation analyses identified alleles at RFLP loci (King et al., 1998). For 20 of 43 (46%) codominant RFLPs, both known alleles (King et al., 1998) were absent and at least one other differently sized fragment was present (Fig. 1). Therefore among these commercial inbreds, there appeared to be more than two alleles at about half of the RFLP loci. We previously reported that cDNAs hybridized to DNA isolated from OP populations revealed an average of 1.9 polymorphic fragments per polymorphic probe-enzyme combination (Bark and Havey, 1995), unexpectedly few given the phenotypic diversity, outcrossing nature, and severe inbreeding depression shown by onion. However this study using inbred materials revealed that the paucity of RFLPs observed among OP populations was likely due to the high level of heterozygosity of the OP population.

The heterozygosity of each inbred was estimated by scoring the number of homozygous versus heterozygous loci for two genetically characterized alleles, as identified in the BYG15-23 x AC43 family (King et al., 1998), at the 23 loci with no evidence of greater than two alleles per locus (as described in the previous paragraph). These estimates (Table 1) were in general agreement with proprietary pedigree information (as provided by the seed companies) and

Table 1. Elite commercial bulb onion inbreds evaluated for nuclear restriction fragment length polymorphisms.

No. ^c	Co.	Day ^y	Cytoplasm ^x	Heterozygosity ^w (%)	Origin ^v
R1	Rio Colorado, Yuma, Ariz.	SD	N	13	Yellow Bermuda
R2	Rio Colorado	SD	N	17	Texas Early Grano
R3	Rio Colorado	SD	N	30	Japanese Overwinter
R4	Rio Colorado	ID	N	13	Spanish x Grano
R5	Rio Colorado	ID	N	26	Spanish x Grano
SG1	S&G/Sandoz, Holland	LD	N	61	Rijnsburger
SG2	S&G/Sandoz	LD	S	13	Sapporo-Ki
SG3	S&G/Sandoz	LD	N	9	Stuttgarter
SH1	Shippo, Kagawa, Japan	LD	S	35	Sapporo-Ki
SH2	Shippo	SD	N	57	Japanese Overwinter
BYG15-23	SunSeeds, Brooks, Ore.	LD	N	13	Storage
AC43	SunSeeds	LD	N	39	United Kingdom
SUN1	SunSeeds	LD	N	30	Spanish
SUN2	SunSeeds	LD	S	48	Spanish
SUN3	SunSeeds	LD	N	43	Storage
SUN4	SunSeeds	LD	N	26	Spanish x storage

^yBYG15-23 and AC43 were previously described (Bradeen and Havey, 1995).

^ySD, ID, and LD are short, intermediate, and long day-length response, respectively.

^xN and S are male fertile and sterile cytoplasms, respectively (Havey, 1993; Jones and Clarke, 1943).

^wPercent heterozygosity estimated by known alleles at 23 RFLP loci (King et al., 1998).

^vSpanish is a LD germplasm grown primarily in Idaho, Oregon, Washington, and Utah. Grano and Yellow Bermuda are SD germplasms grown from Southern California to Georgia. Rijnsburger and Stuttgarter are LD storage populations grown primarily in Germany and Holland. Sapporo-ki is a Japanese LD storage population. Japanese overwintering is a SD bolting-resistant population.

Fig. 1. RFLPs among elite onion inbreds as revealed by probe API21 and *EcoRV*. Approximate fragment sizes in kilobases (kb) are listed on right. Segregation analysis established alleles (4.1 and 4.5 kb) at this RFLP locus (King et al., 1988). The absence of both known alleles and presence of another fragment at 3.0 kb in SG2 was scored as a putative third allele. Abbreviations listed in Table 1.

revealed that onion can be highly inbred (e.g., R1 or SG3 in Table 1). Eleven of the commercial inbreds possessed N cytoplasm; three possessed S (Table 1).

The complete data sets of genetically characterized (King et al., 1998) and uncharacterized RFLPs distinguished all inbreds (Fig. 2). We simulated hybrid combinations among inbred populations for specific market classes by combining gametic genotypes at RFLP loci. All hybrid combinations were distinguished by the complete genetically characterized and uncharacterized data sets (Fig. 3); no two virtual hybrids were identical. These analyses indicate that RFLPs are an effective marker to distinguish among elite onion inbred and hybrid populations.

Random sampling of increasingly smaller data sets demonstrated that 10 genetically characterized or uncharacterized probe-enzyme combinations were adequate to distinguish among all inbred lines; sets of 20 genetically characterized or 10 genetically uncharacterized cDNA clones distinguished among all virtual hybrids. These numbers indicated that relatively few polymorphic probe-enzyme combinations are required to detect mixtures or mistakes during stock-seed maintenance. However, it is obviously much more difficult to estimate whether two inbred or hybrid populations are the same and many more probe-enzyme combinations would have to be evaluated.

RELATIONSHIPS AMONG THE ELITE INBREDS AND VIRTUAL HYBRID BULB ONIONS. UPGMA clustering of the simple-matching coefficients for the genetically characterized and uncharacterized data sets clearly separated BYG15-23 and AC43 (Fig. 2), as expected because RFLPs were identified initially between these two inbreds. There was a clear tendency for populations from individual companies to cluster

together, likely because similar germplasm was used to develop hybrids for specific market niches. For the elite inbred populations, UPGMA clustering generally agreed with known pedigree and phylogenies of major bulb onion populations (Bark and Havey, 1995). European inbred populations were derived from Rijnsburger (SG1) and Stuttgarter (SG3) and placed together (Fig. 2). In the USA, Spanish germplasm traces back to an introduction in 1927 of a population from Valencia, Spain (Havey and Bark, 1994). Inbreds derived from this source of Spanish germplasm clustered loosely together (R1, R2, R4, and SUN1). R1 was derived from Yellow Bermuda, another SD population, and was placed with the Spanish-derived lines. Storage populations (SH1 and SUN3) and inbreds developed from storage $\times$ Spanish crosses (SUN2 and SUN4) were placed loosely together.

The Japanese overwintering onions are early-maturing, bolting-resistant selections that may be fall or spring seeded in the southern or northern U.S., respectively. Two inbred populations selected from this germplasm (R3 and SH2) were placed together for the genetically characterized and uncharacterized data sets, but still differed for at least 30% of the markers (Fig. 2). These inbred populations are more closely related to short-day Spanish germplasm than the European or North American storage populations.

Variability for alleles at RFLP loci or genetically uncharacterized RFLPs clearly separated the virtual hybrids into market classes of storage, Spanish, European, and Japanese overwintering (Fig. 3). Combination of gametic arrays from inbreds generated virtual hybrid data sets that were monomorphic for many fragments and more (as compared to inbreds) allelic or unmapped RFLPs (10 and 20, respectively) must be evaluated to distinguish all virtual hybrids.

In summary, our results indicate that significant allelic diversity exists at RFLP loci within elite germplasm and that onion can tolerate significant levels of inbreeding (Table 1). The entire genetically characterized and uncharacterized data sets distinguished among all elite inbred (Fig. 2) and hybrid (Fig. 3) populations of onion. Random sampling generated smaller data sets and indicated that as few as 10 genetically characterized or uncharacterized probe-enzyme combinations were required to distinguish among the elite inbred populations. As few as 20 genetically characterized or 10 uncharacterized

Fig. 2. Relationships among elite onion inbreds as estimated by unweighted pair group (UPGMA) cluster analyses of the simple-matching [alleles at 75 RFLP (A)] or Jaccard [RFLPs revealed by 69 anonymous cDNAs and alliinase (B)] coefficients.

Fig. 3. Relationships among virtual hybrids as estimated by unweighted pair group (UPGMA) cluster analyses of the simple-matching [alleles at 71 RFLP loci (A)] or Jaccard [RFLPs revealed by 69 anonymous cDNAs and alliinase (B)] coefficients. RFLP profiles for single-cross nonreciprocal virtual hybrids were computer-generated by combining gametic arrays from inbreds of similar market classes. Virtual hybrids were European long-day (LD) storage Eur1 (SG1 x SG2), Eur2 (SG1 x SG3), and Eur3 (SG2 x SG3); US/Japanese LD storage Str1 (SH1 x SUN2), Str2 (SH1 x SUN3), Str3 (SH1 x SUN4), Str4 (SUN2 x SUN3), Str5 (SUN2 x SUN4), and Str6 (SUN3 x SUN4); LD, intermediate-day (ID), and short-day (SD) Spanish Spn1 (R1 x R2), Spn2 (R1 x R4), Spn3 (R1 x R5), Spn4 (R1 x SUN1), Spn5 (R2 x R4), Spn6 (R2 x R5), Spn7 (R2 x SUN1), Spn8 (R4 x R5), Spn9 (R4 x SUN1), and Spn10 (R5 x SUN1); and Japanese overwintering Jap1 (R3 x SH2).

probe-enzyme combinations distinguished the virtual hybrid populations. These results should be useful for cultivar identification and quality control of inbred or hybrid seed lots of onion. However, this study did not address the numbers of polymorphic probe-enzyme combinations required to confidently establish identity among onion inbreds or hybrids.

Literature Cited

- Bark, O.H. and M.J. Havey. 1995. Similarities and relationships among open-pollinated populations of the bulb onion as estimated by nuclear RFLPs. *Theor. Appl. Genet.* 90:607-614.
- Bradeen, J.M. and M.J. Havey. 1995. Randomly amplified polymorphic DNA in bulb onion and its use to assess inbred integrity. *J. Amer. Soc. Hort. Sci.* 120:752-758.
- Chittenden, L.M., K.F. Schertz, Y.R. Lin, R.A. Wing, and A.H. Paterson. 1994. A detailed RFLP map of *Sorghum bicolor* *S. propinquum*, suitable for high-density mapping, suggests ancestral duplication of *Sorghum* chromosomes or chromosomal segments. *Theor. Appl. Genet.* 87:925-933.
- Cui, Y.X., G.W. Xu, C.W. Magill, K.F. Schertz, and G.E. Hart. 1995. RFLP-based assay of *Sorghum bicolor* (L.) Moench genetic diversity. *Theor. Appl. Genet.* 90:787-796.
- dos Santos, J.G., J. Nienhuis, P. Skroch, J. Tivang, and J.K. Slocum. 1994. Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. *Theor. Appl. Genet.* 87:909-915.
- Dudley, J.W. 1994. Comparison of genetic distance estimators using molecular marker data, p. 3-7. In: *Proc. Symp. Analysis of Molecular Marker Data*. Amer. Soc. Hort. Sci.-Crop Sci. Soc. Amer. Joint Plant Breeding Symp. Ser.
- Dudley, J.W. 1995. Analysis of molecular marker data: How to do it; what does it mean? p. 29-51. In: R.J. Lambert (ed.). *Proc. 31st Ill. Corn Breeders School*, Champaign.
- Gentzittel, L., Y.-X. Zhang, F. Vear, B. Griveau, and P. Nicolas. 1994. RFLP studies of genetic relationships among inbred lines of the cultivated sunflower, *Helianthus annuus* L.: Evidence for distinct restorer and maintainer germplasm pools. *Theor. Appl. Genet.* 89:419-425.
- Havey, M.J. 1991. Phylogenetic relationships between cultivated *Allium* species from restriction enzyme analysis of the chloroplast genome. *Theor. Appl. Genet.* 81:752-757.
- Havey, M.J. 1993. A putative donor of S-cytoplasm and its distribution among open-pollinated populations of onion. *Theor. Appl. Genet.* 86:128-134.
- Havey, M.J., and O.H. Bark. 1994. Molecular confirmation that sterile cytoplasm has been introduced into open-pollinated populations of Grano-type onion. *J. Amer. Soc. Hort. Sci.* 119:90-93.
- Helentjaris, T., M. Slocum, S. Wright, A. Schaefer, and J. Nienhuis. 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theor. Appl. Genet.* 72:761-769.
- Jones, H. and A. Clarke. 1943. Inheritance of male sterility in the onion and the production of hybrid seed. *Proc. Amer. Soc. Hort. Sci.* 43:189-194.
- Kennard, W.C., K. Poetter, A. Dijkhuizen, V. Meglic, J.E. Staub, and M.J. Havey. 1994. Linkages among RFLP, RAPD, isozyme, disease resistance, and morphological markers in narrow and wide crosses of cucumber. *Theor. Appl. Genet.* 89:42-48.
- King, J.J., J.M. Bradeen, O. Bark, J.A. McCallum, and M.J. Havey. 1998. A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome. *Theor. Appl. Genet.* 92:52-62.
- Miller, J.D. and S.D. Tanksley. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theor. Appl. Genet.* 80:437-448.
- O'Brien, S.J. 1993. *Genetic maps: Locus maps of complex genomes*. Cold Spring Harbor Lab. Press., Cold Spring Harbor, N.Y.
- Reinisch, A.J., J.M. Dong, C.L. Brubaker, D.M. Stelly, J.F. Wendel, and A.H. Paterson. 1994. A detailed RFLP map of cotton, *Gossypium hirsutum* *Gossypium barbadense*: Chromosome organization and evolution in a disomic polyploid genome. *Genetics* 138:829-847.
- Smith, O.S., J.S.C. Smith, S.L. Bowen, R.A. Tenborg, and S.J. Wall. 1990. Similarities among a group of elite maize inbreds as measured by pedigree, F_1 grain yield, heterosis and RFLPs. *Theor. Appl. Genet.* 80:833-840.
- van Damme, E.J., K. Smeets, S. Torrekens, F. van Leuven, and W.J. Peumans. 1992. Isolation and characterization of alliinase cDNA clones from garlic (*Allium sativum* L.) and related species. *Eur. J. Biochem.* 209:751-757.
- Vierling, R.A., Z. Xiang, C.P. Joshi, M.L. Gilbert, H.T. Nguyen. 1994. Genetic diversity among elite *Sorghum* lines revealed by restriction fragment length polymorphisms and random amplified polymorphic DNAs. *Theor. Appl. Genet.* 87:816-820.