Theor Appl Genet (2006) 114:31–39
DOI 10.1007/s00122-006-0407-y

ORIGINAL PAPER

Comparative sequence and genetic analyses of asparagus BACs reveal no microsynteny with onion or rice

Jernej Jakše · Alexa Telgmann · Christian Jung · Anil Khar · Sergio Melgar · Foo Cheung · Christopher D. Town · Michael J. Havey

Received: 23 March 2006 / Accepted: 31 August 2006 / Published online: 22 September 2006

Abstract The Poales (includes the grasses) and Asparagales [includes onion (*Allium cepa* L.) and asparagus (*Asparagus officinalis* L.)] are the two most economically important monocot orders. The Poales are a member of the commelinoid monocots, a group of orders sister to the Asparagales. Comparative genomic analyses have revealed a high degree of synteny among the grasses; however, it is not known if this synteny extends to other major monocot groups such as the Asparagales. Although we previously reported no evidence for synteny at the recombinational level between onion and rice, microsynteny may exist across shorter genomic regions in the grasses and Asparagales. We sequenced nine asparagus BACs to reveal physically linked genic-like sequences and determined their most similar positions in the onion and rice genomes. Four of the asparagus BACs were selected using molecular markers tightly linked to the sex-determining *M* locus on chromosome 5 of asparagus. These BACs possessed only two putative coding regions and had long tracts of degenerated retroviral elements and transposons. Five asparagus BACs were selected after hybridization of three onion cDNAs that mapped to three different onion chromosomes. Genic-like sequences that were physically linked on the cDNA-selected BACs or genetically linked on the *M*-linked BACs showed significant similarities ($e < -20$) to expressed sequences on different rice chromosomes, revealing no evidence for microsynteny between asparagus and rice across these regions. Genic-like sequences that were linked in asparagus were used to identify highly similar ($e < -20$) expressed sequence tags (ESTs) of onion. These onion ESTs mapped to different onion chromosomes and no relationship was observed between physical or genetic linkages in asparagus and genetic linkages in onion. These results further indicate that synteny among grass genomes does not extend to a sister order in the monocots and that asparagus may not be an appropriate smaller genome model for plants in the Asparagales with enormous nuclear genomes.

Communicated by J. S. Heslop-Harrison.

J. Jakše · A. Khar · S. Melgar
Department of Horticulture, University of Wisconsin,
1575 Linden Drive, Madison, WI 53706, USA

J. Jakše
Agronomy Department, Biotechnical Faculty,
University of Ljubljana, Jamnikarjeva 101,
Ljubljana 1000, Slovenia

A. Telgmann · C. Jung
Plant Breeding Institute,
Christian-Albrechts-University of Kiel,
Am Botanischen Garten 1-9, 24118 Kiel, Germany

F. Cheung · C. D. Town
The Institute for Genomic Research,
9712 Medical Center Dr., Rockville,
MD 20850, USA

M. J. Havey (✉)
Agricultural Research Service, USDA,
Department of Horticulture, University of Wisconsin,
1575 Linden Drive, Madison, WI 53706, USA
e-mail: mjhavey@wisc.edu

Introduction

The Commelinanae and Asparagales are sister groups within the monocots (APG II 2003). The commelinoid monocots include the orders Arecales (palms),

Commelinales, Zingiberales (bananas and ginger), and Poales (grasses). The grasses are the most economically significant monocots and include such important plants as barley (*Hordeum vulgare* L.), maize (*Zea mays* L.), rice (*Oryza sativa* L.), sorghum (*Sorghum bicolor* L.), or wheat (*Triticum aestivum* L.). The order Asparagales is the second most economically valuable group of monocots and carries significant plants such as aloe (*Aloe vera* L.), agave (*Agave tequilana* Web.), asparagus (*Asparagus officinalis* L.), garlic (*Allium sativum* L.), leek (*Allium ampeloprasum* L.), and onion (*Allium cepa* L.), as well as ornamental amarylids, irises, and orchids (Chase et al. 1995). Onion is the most economically significant member of the Asparagales and the second most valuable vegetable in the world, following only tomato (FAO 2003). Sequencing of chloroplast genes has revealed that the Alliaceae (garlic and onion), Agavacaeae (agave and yucca), and Asparagaceae (asparagus) are members of a monophyletic group within the Asparagales, termed the core Asparagales (Chase et al. 2000).

In spite of their economic importance, few genomic resources have been developed for the Asparagales primarily due to their enormous nuclear genomes. Onion is a diploid ($2n = 2x = 16$) with a genome size of 16,415 Mbp/1C, similar in size to that of hexaploid wheat and 107, 36, and 6 times greater than *Arabidopsis*, rice, and maize, respectively (Arumuganathan and Earle 1991). C_0t reassociation kinetics revealed that the onion genome consists of middle-repetitive sequences occurring in short-period interspersions among low-copy regions (Stack and Comings 1979). Do et al. (2003) sequenced a 35-kb genomic region of onion and revealed the target alliinase gene, one hypothetical gene, and copious retroviruses and transposons. These results suggest that much of the onion genome is composed of repetitive elements; however, it is not known if gene-rich regions exist in the onion genome similar to those in the grasses (Keller and Feuillet 2000).

Asparagus is a diploid plant ($2n = 2x = 20$) with one of the smallest nuclear genomes (1,308 Mbp/1C) among all plants in the core Asparagales, over 12 times smaller than onion and about half that of maize (Arumuganathan and Earle 1991). Leitch et al. (2005) observed that the Asparagaceae may have undergone a significant reduction in genome size relative to other families in the core Asparagales. Although European *Asparagus* species have the same chromosome number and twice the nuclear DNA as South African species (Stajner et al. 2002), there is no evidence of a recent genome doubling of European relative to South African *Asparagus* species (Kuhl et al. 2005). As a result, the relatively small asparagus genome may be a useful genomic model for plants in the core Asparagales with enormous nuclear genomes, such as the Alliums.

Extensive genetic-linkage conservation (synteny) among related species aids in the identification, mapping, and cloning of economically important qualitative and quantitative trait loci (Paterson et al. 1996; Maughan et al. 1996). Synteny among the cultivated grasses is widely recognized (Binelli et al. 1992; Ahn et al. 1993; Ahn and Tanksley 1993; Devos et al. 1994; Dunford et al. 1995; Dubcovsky et al. 2001), although exceptions exist at the recombinational and sequence levels (Chen et al. 1997; Tikhonov et al. 1999; Tarchini et al. 2000; Bennetzen 2000; Bennetzen and Ramakrishna 2002; Ma and Bennetzen 2004). Genomic comparisons are necessary to establish if the grass genomes are representative of other major monocots. We analyzed expressed regions in the asparagus, onion, and rice genomes and revealed significant differences for GC contents (Kuhl et al. 2004, 2005), as well as little to no synteny on the recombinational level between onion and rice (Martin et al. 2005). These results suggest that the grass genomes (family Poaceae in order Poales of the commeliniod monocots) may not be representative of the Asparagales. In this study, we selected and sequenced asparagus BACs to assess synteny with onion and rice and to establish whether the smaller genome of asparagus could serve as a model for larger genome plants in the Asparagales.

Materials and methods

The BACs were selected by screening an asparagus library (Jamsari et al. 2004) with P^{32}-labeled onion cDNAs (Bark et al. 1994) for sucrose transporter (Genbank accession BE205593), a sulfite reductase (AF403294), and AOB272 [(AA451592) which reveals a restriction fragment length polymorphism (RFLP) tightly linked to the *Ms* locus of onion (Gokce et al. 2002)]. These three cDNAs revealed simple hybridization patterns in onion and mapped to onion chromosomes 5, 3, and 2, respectively (Martin et al. 2005). The BAC library was also screened with molecular markers tightly linked at 0.25 (Asp4-SP6, STS3156, and STS4150.3) and 0.37 (EM3646) cM to the *M* locus on asparagus chromosome 5 (Telgmann, unpublished), which controls male versus female plants (Löptien 1979). BAC ends of positive clones were sequenced, cloned, and hybridized to DNA-gel blots of asparagus to identify low-copy regions, which were then used to select adjacent BACs. Putative BAC contigs were constructed after fingerprinting using FPC (Soderlund

et al. 2000). A total of nine asparagus BACs were selected, corresponding to four different regions of the asparagus genome.

The cDNA-selected BACs were sequenced to GenBank HTGS phase 2 quality sequence revealing structure, relative orientation, and position of all genes on the BACs. Random small (2–3 kb) and large (10–12 kb) insert libraries were constructed from hydrodynamically sheared, size-selected BAC DNA in a medium copy vector. These libraries were sequenced up to a total coverage of at least 8×. Base calling of the cDNA-selected BACs was performed by Paracel TraceTuner and sequence quality trimming and elimination of vector and *E. coli* sequences were conducted using TIGR software (Chou and Holmes 2001). Sequences were assembled, ordered, and oriented using Bambus software (Pop et al. 2004). The *M-l*inked BACs were sequenced after shearing and subcloning into a high-copy plasmid vector. For each BAC, 182 subclones with an average insert size of 1.1 kb were sequenced in both directions to yield coverage between 2.6 and 4.0×, corresponding to GenBank HTGS phase 1 in which ordering and the orientation of the contigs were not complete. Base calling, quality trimming, and elimination of vector sequence was performed using PHRED (Ewing et al. 1998; Ewing and Green 1998) and contigs were assembled using the Staden Package (Staden 1996). The GC contents of BAC sequences were calculated and GC distributions revealed using the Swaap program (Pride 2000) with a window size of 200 and step size of 100 bp. Repetitive DNAs among the asparagus BACs were revealed using multi-percent identity plots (PIP) masking against the PANICOID database with repeat masker (Schwartz et al. 2003).

Asparagus BAC sequences were compared using translated searches to identify similar sequences in the EST and genomic databases at TIGR. Hypothetical gene-like regions were identified when greater than 100 amino acids by Genemark.hmm using the *Arabidopsis* matrix, excluding all matches with even low similarities to transposons or transposon-related sequences, and were used to calculate the percentage of putative genic-like regions on the asparagus BACs. Hypothetical proteins greater than 100 amino acids and having significant ($e < -20$) hits in the rice databases were compared to their most similar positions in the rice genome. *E*-values between asparagus genomic DNA and the onion and rice ESTs were determined by the Blast 2 sequences program available on the NCBI site (Tatusova and Madden 1999). Copy numbers of highly similar ($e < -20$) onion ESTs were estimated after hybridizations (Bark et al. 1994) to DNA-gel blots of onion doubled haploid (DH) populations RNG3616, RNG4088, REG654, or REG652 (gift of Seminis Seed Company, Woodland, CA, USA), singly digested with at least three restriction enzymes. Onion ESTs revealing 1–3 fragments on DH DNAs were mapped in the BYG15-23 × AC43 mapping population (King et al. 1998). ACAEV16 (GeneBank CF452679) and ACAHW85 (GeneBank CF446239) were mapped as RFLPs using *Eco*RI and *Hin*dIII, respectively. ACADN70 (GeneBank CF451233) was amplified as 1,042-bp fragment from AC43 and BYG15-23 using nested primers TGAGTATCCCTT TGCTTTCATGT (external forward), GGACTGAA ATCCCACCAATATG (external reverse), CTTGTG GTGGCTATTTGCTTACT (nested forward), and AGCTTTGTTTTCTGCACCAACT (nested reverse) and was mapped as an A/T single nucleotide polymorphism (SNP) at position 249. ACAER17 (GeneBank CF452530) was amplified using nested primers [GGACCCACAATGGGAGCTT (external forward, CCCAAACACCTCATGAAACC (external reverse), GACGAAGCCCTAGGTTGCT (nested forward), and CCAACATCAAGAACAGCATCA (nested reverse)] and mapped by scoring a 4-bp indel. The PCR conditions were optimized as previously described (Martin et al. 2005). Mapping of all polymorphisms was completed using Map Manager (Manly et al. 2001).

Results

Sequencing of asparagus BACs

We selected asparagus BACs showing strong hybridization signals with independently inherited onion cDNAs for sucrose transporter (BAC B4-L7-94), sulfite reductase (BAC B2-A14-8), and AOB272 (overlapping BACs B1-K23-67, B6-A19-72, and C5-O1-101). We also selected four asparagus BACs carrying markers flanking the *M* locus at 0.25 cM (Ki-55B1, Ko-106C13, and Ko-31L4) and 0.37 (Ko-109H1) cM on asparagus chromosome 5. Sequencing of these nine BACs yielded 705,279 bp (Table 1) and revealed a plethora of repetitive sequences showing significant similarities to retroviruses such as Copia-Ty1 and Gypsy-Ty3-like elements (Fig. 1). The average GC content across all BACs was 39.8% (Fig. 1). Multi-PIP analyses revealed similarities among repetitive sequences on different asparagus BACs (Fig. 1), which were primarily retroelements or AT-rich regions. We observed only weak similarities among repetitive sequences in the asparagus and rice genomes (data not shown).

Table 1 Sizes and Genbank accessions for bacterial artificial chromosomes (BACs) of asparagus with estimated percentage of putative coding regions in asparagus and rice

Target gene[a]	Asparagus BACs	Genbank accessions	Size in basepairs	Percent genic-like regions[b]	
				Asparagus	Rice
AOB272	B1-K23-67 and B6-A19-72 C5-O1-101	DQ273273 and DQ273275	213,317	16	24
Sucrose transporter	B4-L4-94	DQ273271	112,052	16	25
Sulfite reductase	B2-A14-8	DQ273274	67,545	11	16
STS4150.3	Ki-55B1	AC183433	89,413	1	10
Asp4-SP6	Ko-106C13	AC183434	71,281	2	8
EM3646	Ko-109H1	AC183435	52,700	0	–
STS3156	Ko-31L4	AC183436	98,971	0	–

[a] Asp4-SP6, EM3646, STS4150.3, and STS3156 are genetic markers tightly linked to *M* locus on asparagus chromosome 5 (Jamsari et al. 2004; Telgmann unpublished)

[b] Percent genic-like regions across equally sized, putatively orthologous locations in the asparagus and rice genomes were estimated by counting all hypothetical proteins greater than 100 amino acids with similarities of at least $e-20$ in databases. All matches to transposon-like sequences were excluded

Physically linked sequences in asparagus show no synteny in rice

We previously demonstrated that expressed regions in the onion genome show no evidence for synteny on the recombinational level with rice (Martin et al. 2005). Because microsynteny may exist between the Asparagales and grass genomes, we compared the physical locations of highly similar sequences in asparagus and rice. The cDNA-selected asparagus BACs had 33% fewer genic-like regions than corresponding, similarly sized locations in the rice genome (Table 1), which would be expected given the larger asparagus genome. Asparagus BAC B2-A14-8 was 67,545 bp in size and carried coding regions with high similarities to the target sulfite reductase, a calcium-dependent kinase (CDPK), and a MYB-related DNA-binding protein. Comparisons with rice revealed that the most highly similar region to sulfite reductase was on rice chromosome 5, and a sequence highly similar to CDPK was on rice chromosome 2 (Table 2). The contig of BACs B1-K23-67, B6-A19-72, and C5-O1-101 was 213,317 bp in size and carried target sequences highly similar to an Isp-4-like oligopeptide transporter protein (corresponding to the onion cDNA AOB272 used for selection), a zinc transporter, and a eukaryotic initiation factor 3E subunit; the corresponding most similar regions in rice were located on chromosomes 3 and 7 (Table 2). The two highly similar regions on rice chromosome 3 corresponding to the zinc transporter and Isp4-like protein were separated by 14 megabases. We also compared the positions in rice of physically linked asparagus sequences of lower similarities (between $e-20$ and the most significant hit) and observed no evidence for synteny among these regions (data not shown). Asparagus BAC B4-L7-94 was 112,052 bp in size (Table 1) and possessed only the target sucrose transporter and no other putative coding regions as defined by >100 amino acids with hits to databases below $e-20$.

Sequencing of the four *M*-linked BACs generated 312,365 bp (Table 1). Asparagus BAC Ki-55B1 was 89,431 bp in size and carried two putative coding regions with high similarities to one uncharacterized protein in rice. BAC Ko-106C13 was 71,281 bp in size and possessed a putative coding region similar to a photosystem I protein. These regions cosegregate in asparagus and showed significant similarities to sequences on rice chromosomes 2 and 12 (Table 2). No regions with significant similarities to rice genes were found on asparagus BACs Ko-31L4 and Ko-109H1. The *M*-linked BACs had fewer genic-like regions than rice or the cDNA-selected BACs, and carried 31 regions with high similarities to retroelements or transposons.

Genetic mapping in onion of physically linked sequences on the asparagus BACs

Synteny may exist among plants within the monophyletic core Asparagales. We assessed synteny between the smaller genome of asparagus and the enormous genome of onion by genetic mapping of onion ESTs showing highly significant ($e < -20$) similarities to physically linked sequences on the asparagus BACs. The asparagus BACs B1-K23-67 and B6-A19-72 carried sequences highly similar to Isp4-like protein (AOB272) that reveals a single-copy RFLP mapping to

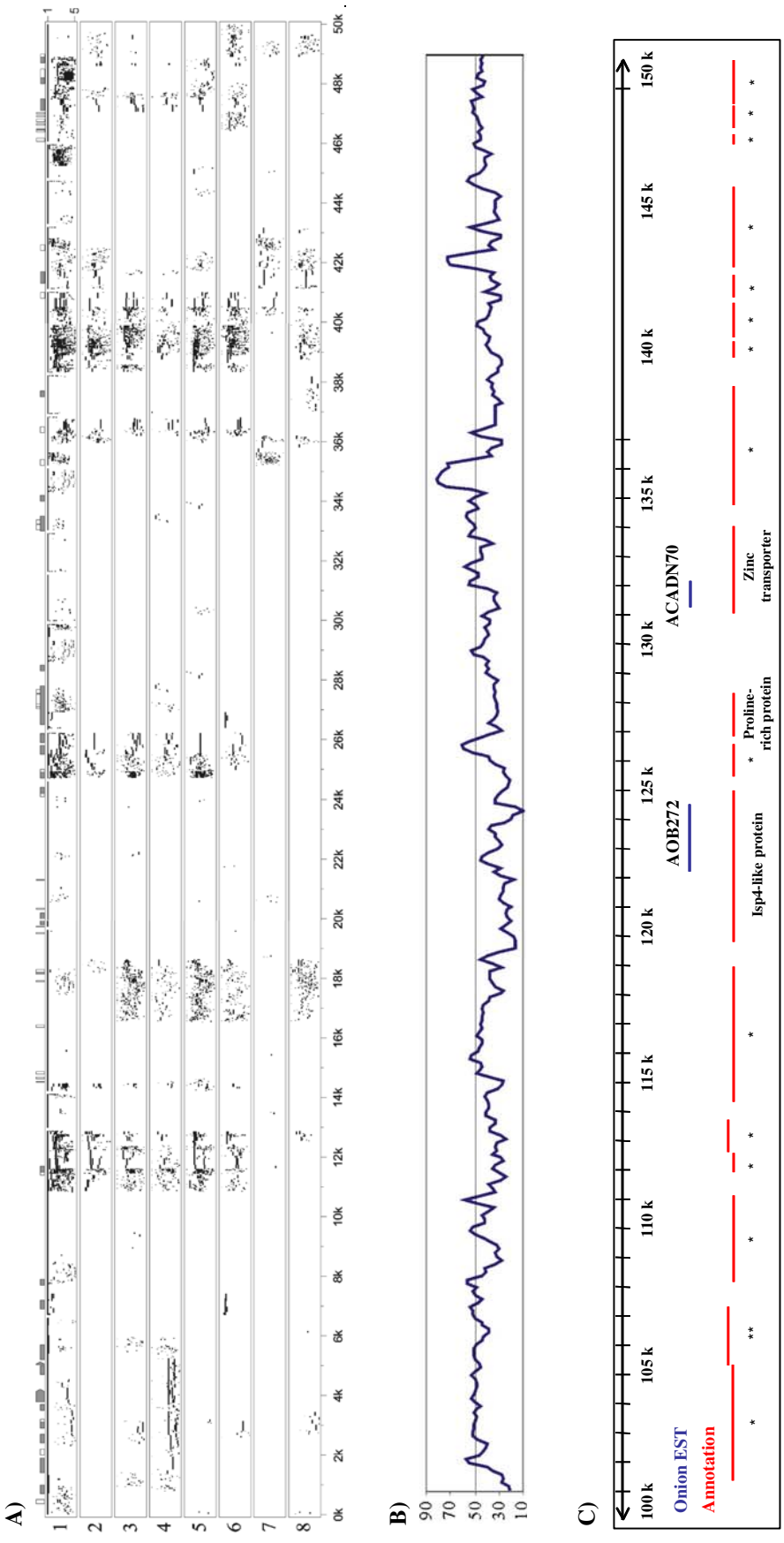

Fig. 1 **a** Multi-percent identity plots (multi-PIP) from 100 to 150 kb on the contig of asparagus BACs B1-K23-67 and B6-A19-72 as compared to the whole contig (1) and individual BACs C5-O1-101 (2), B4-L4-94 (3), B2-A14-8 (4), Ki-55B1 (5), Ko-106C13 (6), Ko-109H1 (7), or Ko-31L4 (8). Positions in kilobases are shown at *bottom* and sequence similarities from 50 to 100% are shown on *right*. *Bars above plots* show CpG islands (*low gray and white bars* represent >0.75 and 0.60, respectively), simple repeats (*clear tall bars*), and long terminal repeats (*pointed bars*) (Schwartz et al. 2003). **b** GC sliding window analysis with a window size of 200 and step size of 100 bp for the genomic region from 100 to 150 kb on the asparagus BAC contig B1-K23-67 and B6-A19-72. **c** Graphical representation of sequences from 100 to 150 kb on BAC contig B1-K23-67 and B6-A19-72 with putative annotations shown in red (* designates retroviral or transposon-like sequences) and significant similarities to onion ESTs shown in *blue*

Table 2 Physically linked sequences on asparagus BACs and their corresponding most similar positions in the rice genome and genetic map of onion

Asparagus		Rice				Onion		
BAC[a]	Putative gene	Marker	Chromosome	5′ Position (bp)	Score	EST	Chromosome	cM[b]
DQ273273	Zinc transporter	OJ1212_C05.1	3	16,975,302	$e-73$	ACADN70	5	54.9
DQ273273	Eukaryotic initiation factor 3E subunit	Os07g12110	7	6,765,633	$e-73$	ACAEV16	1	134.2
DQ273273	Isp4-like protein	OSJNBa0047E24.17	3	30,916,235	$e-150$	AOB272	2	98.0
DQ273274	Calcium-dependent protein kinase	Os02g58520	2	35,720,142	$e-109$	ACAHW85	1	149.2
DQ273274	Sulfite reductase	Os05g42350	5	24,594,944	$e-90$	ACAAJ79	3	229.8
AC183433	Uncharacterized protein family	Os02g46150	2	28,143,483	$e-40$	ACAER17	8	71.7
AC183434	Photosystem I subunit XI chloroplast precursor	Os12g23200	12	13,115,162	$e-33$	API56	6	[c]

[a] Genbank accession number for asparagus BACs

[b] Map positions described by Martin et al. (2005)

[c] Onion cDNA API56 was assigned to onion chromosome 6 by the former Native Plants, Inc., as described by King et al. (1998). We were unable to map this cDNA because it was not polymorphic in the BYG15-23 and AC43 segregating family

the center of onion chromosome 2 (Gokce et al. 2002; Martin et al. 2005). Annotation of these asparagus BACs revealed physically linked sequences highly similar to onion ESTs ACADN70 and ACAEV16 (Fig. 1, Table 2). We established that these onion ESTs are low copy in the onion genome by hybridizations to DNA-gel blots carrying onion doubled-haploid DNAs (Fig. 2). The three onion ESTs that showed high similarities to physically linked sequences on asparagus BAC contig B1-K23-67 and B6-A19-72 mapped to onion chromosomes 1, 2, and 5 (Fig. 3). Two onion ESTs showing significant similarities to putative coding regions on two *M*-linked BACs (Ki-55B1 and Ko-106C13) were assigned to onion chromosomes 6 and 8 (Table 2). Overall, no genetic linkages were detected in onion among ESTs highly similar to physically linked sequences on the cDNA-selected BACs or genetically linked sequences on the *M*-linked BACs from asparagus (Table 2).

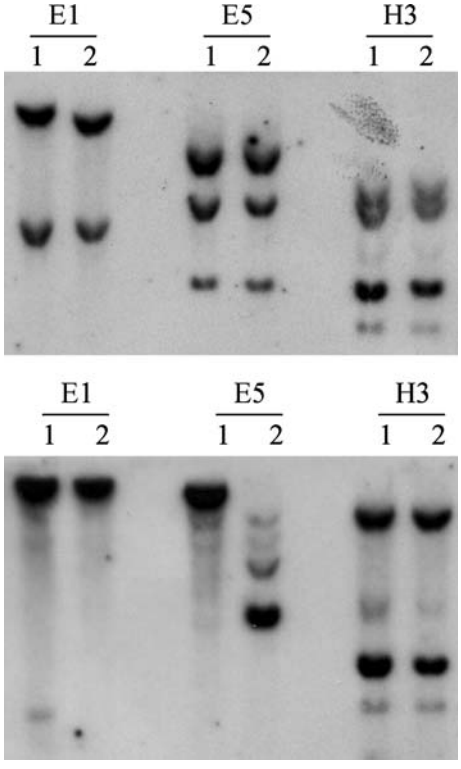

Fig. 2 Autoradiograms revealing low-copy numbers of onion expressed sequence tags (ESTs) ACAEV16 (*top*) and ACADN70 (*bottom*) after hybridization to doubled haploid onion DNAs [RNG3616 (1) and REG652 (2)] digested with *Eco*RI (E1), *Eco*RV (E5), or *Hin*dIII (H3). These ESTs show highly significant similarities to physically linked sequences on the B1-K23-67 and B6-A19-72 BAC contig

Discussion

Sequence characteristics of the asparagus genome

Sequencing of the cDNA-selected asparagus BACs revealed putative coding regions showing significant similarities to other plants. These regions were surrounded by pockets of repetitive DNAs, most of which showed similarities to retroviral elements and transposons (Fig. 1). This structure is similar to the maize genome, in which isolated coding regions are flanked by long stretches of retroelements and transposons (San-Miguel et al. 1996). Sequencing of the *M*-linked BACs revealed that 312,365 bp carried only three putative coding regions showing significant ($e < -20$) similarities

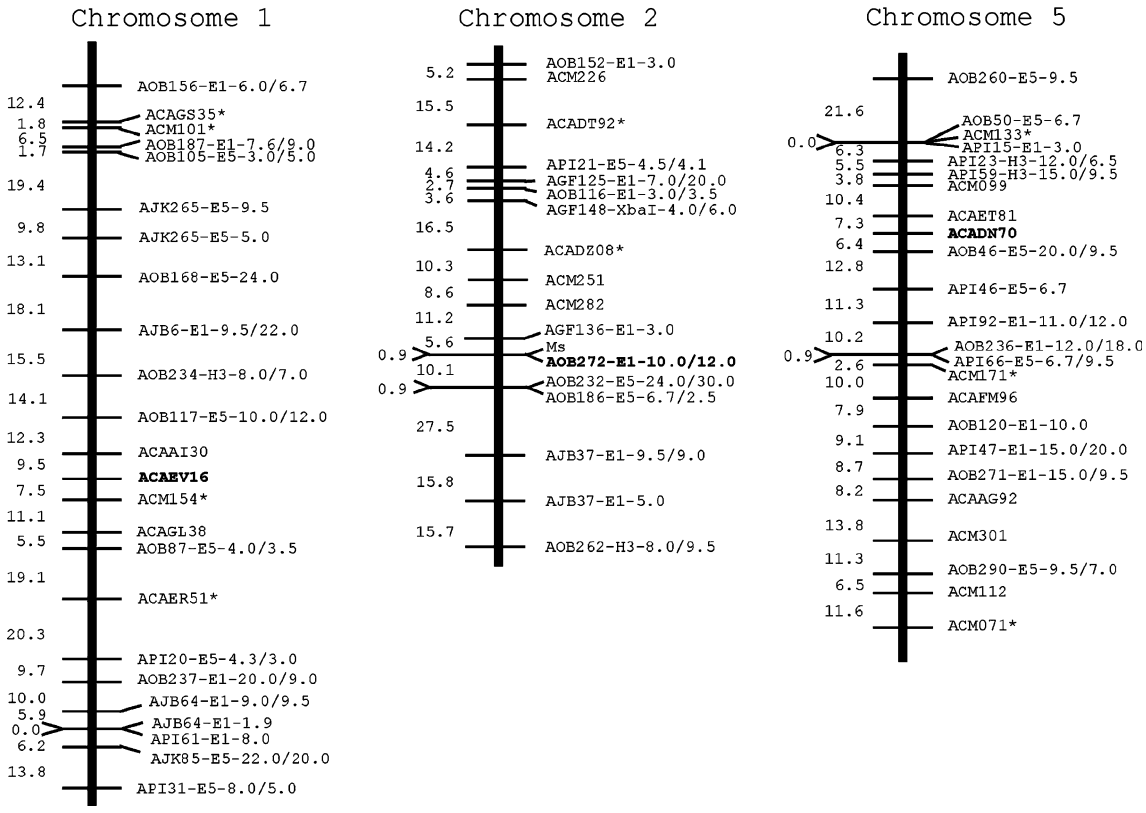

Fig. 3 Onion expressed sequence tags (ESTs) AOB272, ACAEV16, and ACADN70 are highly similar to physically linked sequences on asparagus BAC contig B1-K23-67 and B6-A19-72 and map to different onion chromosomes (positions shown with *bold characters*). Genetic markers in onion were described by King et al. (1998) and Martin et al. (2005)

in rice and gene densities on the *M*-linked BACs were lower than the cDNA-selected BACs (Table 1). The lower gene densities (Table 1) and plethora of repetitive DNAs on the *M*-linked BACs are consistent with high frequencies of repetitive elements in chromosome regions carrying sex-determination loci. Dioecy in plants is conditioned by homomorphic [papaya (*Carica papaya*)] or heteromorphic sex chromosomes [*Silene latifolia*, sorrel (*Rumex acetosa*), and hemp (*Cannabis sativa*)]. Accumulation of repetitive sequences has been documented on the Y chromosomes of *S. latifolia* (Lengerova et al. 2004), sorrel (Shibata et al. 1999), liverwort (*Marchantia polymorpha*) (Okada et al. 2001; Ishizaki et al. 2002), and hemp (Sakamoto et al. 2000, 2005). Liu et al. (2004) demonstrated that the male-specific chromosome region of papaya accumulated retroelements due to suppression of recombination. In a previous study, five AFLP markers tightly linked to the *M* locus of asparagus were characterized by FISH and Southern analyses and found to be highly repetitive and not specific to chromosome 5 (Reamon-Büttner et al. 1999). The accumulation of repetitive DNAs near the *M* locus indicates that map-based cloning of this locus may be difficult.

Asparagales genomics

The larger grass genomes exist as gene-rich islands in a sea of repetitive elements (SanMiguel et al. 1996). Orthologous gene islands show a high degree of synteny among the grasses (e.g. Chen et al. 1997; Devos and Gale 2000; Paterson et al. 2000, 2003), with tandem duplications and small inversions common among these gene islands (Bowers et al. 2003; Dubcovsky et al. 2001; Fu and Dooner 2002; Li and Gill 2002; Song et al. 2002). We did not reveal synteny among rice, asparagus, and onion on the recombinational (Martin et al. 2005) and sequence levels (Table 2, Fig. 3). Although physically linked sequences on the asparagus BACs occasionally showed high similarities to sequences on the same rice chromosomes, these regions were separated by megabases and provided little evidence for microsynteny between asparagus and rice (Table 2). This lack of synteny may be expected because the lineages giving rise to the grasses and Asparagales split over 130 million years ago (MYA), and the Asparagaceae and Alliaceae split approximately 87 MYA (Jansson and Bremer 2005). The Asparagaceae, Alliaceae, and Poaceae diversified over the last approximately 87, 85, and 83 MYA,

respectively. Therefore, the older split of the Asparagaceae from the Alliaceae, relative to the diversification of the grasses, may be responsible for the deterioration of synteny among these Asparagales genomes. However, it is possible that highly similar sequences on the asparagus BACs and the onion ESTs were not orthologous. Paterson et al. (2004) proposed that the progenitor of the grasses may have experienced a whole-genome duplication approximately 70 MYA, after the split of the Alliaceae and Asparagaceae. Such duplication events, followed by targeted or random gene loss (Ilic et al. 2003; Paterson et al. 2004), may have produced paralogs that retain significant sequence similarities, but show no synteny. Within the core Asparagales, we observed no synteny between physically linked coding regions in the smaller genome of asparagus and genetic linkages in onion. These results demonstrate that the grass genomes are not necessarily representative of the genomes of other monophyletic groups of monocots, at least across the sequenced regions, and that genomic resources may have to be independently developed for the Asparagales and possibly for important families such as the Alliaceae and Asparagaceae.

Acknowledgments This work was completed in compliance with the current laws governing genetic experimentation in the United States and was supported by the Initiative for Future Agriculture and Food Systems Grant 2001-52100-11344 from the USDA Cooperative State Research, Education, and Extension Service; German Research Foundation DFG grant Ju 205/2-3; a Fulbright-Hayes Post-doctoral Fellowship to JJ; and a Boyscast Fellowship to AK.

Disclaimer: Names are necessary to report factually on available data; however, the US Department of Agriculture (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.

References

Ahn S, Tanksley S (1993) Comparative linkage maps of the rice and maize genomes. Proc Natl Acad Sci USA 90:7980–7984

Ahn S, Anderson JA, Sorrells ME, Tanksley SD (1993) Homoeologous relationships of rice, wheat and maize chromosomes. Mol Gen Genet 241:483–490

APG II [Bremer B, Bremer K, Chase MW, Reveal JL, Soltis DE, Soltis PS, Stevens PF, Anderberg AA, Fay MF, Goldblatt P, Judd WS, Kallersjo M, Karehed J, Kron KA, Lundberg J, Nickrent DL, Olmstead RG, Oxelman B, Pires JC, Rodman JE, Rudall PJ, Savolainen V, Sytsma KJ, van der Bank M, Wurdack K, Xiang JQY, Zmarzty S] (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot J Linn Soc 141:399–436

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Reporter 9:208–218

Bark OH, Corgan J, Havey MJ (1994) RFLP analysis of progeny from an interspecific hybrid between *Allium fistulosum* and *Allium cepa*. J Am Soc Hortic Sci 119:1046–1049

Bennetzen JL (2000) Comparative sequence analysis of plant nuclear genes: microcolinearity and its many exceptions. Plant Cell 12:1021–1029

Bennetzen JL, Ramakrishna W (2002) Numerous small rearrangements of gene content, order and orientation differentiate grass genomes. Plant Mol Biol 48:821–827

Binelli G, Ginafranceschi L, Pe M, Taramino G, Busso C, Stenhouse J, Ottaviano E (1992) Similarity of maize and sorghum genomes as revealed by maize RFLP probes. Theor Appl Genet 84:10–16

Bowers JE, Abbey C, Anderson S, Chang C, Draye X, Hoppe AH, Jessup R, Lemke C, Lennington J, Li Z, Lin YR, Liu SC, Luo L, Marler BS, Ming R, Mitchell SE, Qiang D, Reischmann K, Schulze SR, Skinner DN, Wang YW, Kresovich S, Schertz KF, Paterson AH (2003) A high-density genetic recombination map of sequence-tagged sites for sorghum, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. Genetics 165:367–386

Chase M, Duvall M, Hills H, Conran J, Cox A, Eguiarte L, Hatwell K, Fay M, Caddick R, Cameron K, Hoot S (1995) Molecular phylogenetics of Lilianae. In: Rudall P, Crib P, Culter D, Humphries C (eds) Monocotyledons: systematics and evolution. Royal Botanic Gardens, Kew, pp 109–137

Chase MW, Soltis DE, Soltis PS, Rudall PJ, Fay MF, Hahn WH, Sullivan S, Joseph J, Molvray M, Kores PJ, Givnish TJ, Sytsma KJ, Pires JC (2000) Higher-level systematics of the monocotyledons: an assessment of current knowledge and a new classification. In: Wilson KL, Morrison S (eds) Monocots: systematics and evolution. CSIRO, Melbourne, pp 3–16

Chen M, SanMiquel P, DeOliviera A, Woo S, Zhang H, Wing R, Bennetzen J (1997) Microcolinearity in sh2-homologous regions of maize, rice, and sorghum genomes. Proc Natl Acad Sci USA 94:3431–3435

Chou HH, Holmes MH (2001) DNA sequence quality trimming and vector removal. Bioinformatics 17:1093–1104

Devos KM, Gale MD (2000) Genome relationships: the grass model in current research. Plant Cell 12:637–646

Devos KM, Chao S, Li QY (1994) Relationship between chromosome 9 of maize and wheat homeologous group 7 chromosomes. Genetics 138:1287–1292

Do S, Suzuki G, Mukai Y (2003) Genomic organization of a novel root alliinase gene, *ALL1*, in onion. Gene 325:17–24

Dubcovsky J, Ramakrishna W, SanMiguel PJ, Busso CS, Yan L, Shiloff BA, Bennetzen JL (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. Plant Physiol 125:1342–1353

Dunford RP, Kurata N, Laurie DA, Money TA, Minobe Y, Moore G (1995) Conservation of fine-scale DNA marker order in the genomes of rice and the Triticeae. Nucleic Acids Res 23:2724–2728

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8:186–194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8:175–185

FAO (2003) World production and trade statistics. http://www.apps.fao.org/

Fu H, Dooner HK (2002) Intraspecific violation of genetic colinearity and its implications in maize. Proc Natl Acad Sci USA 99:9573–9578

Gokce AF, McCallum J, Sato Y, Havey MJ (2002) Molecular tagging of the *Ms* locus in onion. J Am Soc Hortic Sci 127:576–582

Ilic K, SanMiguel PJ, Bennetzen JL (2003) A complex history of rearrangement in an orthologous region of the maize, sorghum, and rice genomes. Proc Natl Acad Sci USA 100:12265–12270

Ishizaki K, Shimizu-Ueda Y, Okada S, Yamamoto M, Fujisawa M, Yamato KT, Fukuzawa H, Ohyama K (2002) Multicopy genes uniquely amplified in the Y chromosome-specific repeats of the liverwort *Marchantia polymorpha*. Nucleic Acids Res 30:4675–4681

Jamsari A, Nitz I, Reamon-Büttner SM, Jung C (2004) BAC-derived diagnostic markers for sex determination in asparagus. Theor Appl Genet 108:1140–1146

Jansson T, Bremer K (2005) The age of major monocot groups inferred from 800 + *rbcL* sequences. Bot J Linn Soc 146:395–398

Keller B, Feuillet C (2000) Colinearity and gene density in grass genomes. Trends Plant Sci 5:246–251

King JJ, Bradeen JM, Bark O, McCallum JA, Havey MJ (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 96:52–62

Kuhl JC, Cheung F, Yuan Q, Martin W, Zewdie Y, McCallum J, Catanach A, Rutherford P, Sink KC, Jenderek M, Prince JP, Town CD, Havey MJ (2004) A unique set of 11,008 onion (*Allium cepa*) ESTs reveals expressed sequence and genomic differences between monocot orders Asparagales and Poales. Plant Cell 16:114–125

Kuhl JC, Havey MJ, Cheung F, Yuan Q, Leebens-Mack J, Town CD, Sink KC (2005) Comparative genomic analyses of the genus *Asparagus*. Genome 48:1052–1060

Leitch IJ, Soltis DE, Soltis PS, Bennett MD (2005) Evolution of DNA amounts across land plants (Embryophyta). Ann Bot 95:207–217

Lengerova M, Kejnovsky E, Hobza R, Macas J, Grant SR, Vyskot B (2004) Multicolor FISH mapping of the dioecious model plant, *Silene latifolia*. Theor Appl Genet 108:1193–1199

Li W, Gill BS (2002) The colinearity of the Sh2/A1 orthologous region in rice, sorghum and maize is interrupted and accompanied by genome expansion in the triticeae. Genetics 160:1153–1162

Liu Z, Moore PH, Ma H, Ackerman CM, Ragiba M, Yu Q, Pearl HM, Kim MS, Charlton JW, Stiles JI, Zee FT, Paterson AH, Ming R (2004) A primitive Y chromosome in papaya marks incipient sex chromosome evolution. Nature 427:348–352

Löptien H (1979) Identification of the sex chromosome pair in asparagus (*Asparagus officinalis* L.). Z Pflanzen 82:162–173

Ma J, Bennetzen JL (2004) Rapid recent growth and divergence of rice nuclear genomes. Proc Natl Acad Sci USA 101:12404–12410

Manly KF, Cudmore RH Jr, Meer JM (2001) Map Manager QTX, cross-platform software for genetic mapping. Mamm Genome 12:930–932

Martin W, McCallum J, Shigyo M, Jakse J, Kuhl JC, Yamane N, Sink KC, Town CD, Havey MJ (2005) Genetic mapping of expressed sequences in onion reveal a collinear region with garlic, but in silico comparisons show scant collinearity with rice. Mol Genet Genomics 274:197–204

Maughan P, Saghai-Maroof M, Buss G (1996) Molecular-marker analysis of seed weight: genomic location, gene action, and evidence for orthologous evolution among three legume species. Theor Appl Genet 93:574–579

Okada S, Sone T, Fujisawa M, Nakayama S, Takenaka M, Ishizaki K, Kono K, Shimizu-Ueda Y, Hanajiri T, Yamato K T, Fukuzawa H, Brennicke A, Ohyama K (2001) The Y chromosome in the liverwort *Marchantia polymorpha* has accumulated unique repeat sequences harboring a male-specific gene. Proc Natl Acad Sci USA 98:9454–9459

Paterson A, Lan T, Reischmann K (1996) Toward a unified genetic map of higher plants, transcending the monocot-dicot divergence. Nat Genet 14:380–382

Paterson A, Bowers JE, Burow MD, Draye X, Eisik CG, Jiang CX, Katsar CS, Land TH, Lin YR, Ing R, Wright RJ (2000) Comparative genomics of plant chromosomes. Plant Cell 12:1523–1539

Paterson AH, Bowers JE, Peterson DG, Estill JC, Chapman BA (2003) Structure and evolution of cereal genomes. Curr Opin Genet Dev 13:644–650

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci USA 101:9903–9908

Pop M, Kosack D, Salzberg S (2004) Hierarchical scaffolding with *Bambus*. Genome Res 14:149–159

Pride DT (2000) SWAAP Version 1.0.0—sliding windows alignment analysis program: a tool for analyzing patterns of substitutions and similarity in multiple alignments. Distributed by the author at http://www.bacteriamuseum.org/SWAAP/SwaapPage.htm

Reamon-Büttner S, Schmidt T, Jung C (1999) AFLPs represent highly repetitive sequences in *Asparagus officinalis* L. Chromosome Res 7:297–304

Sakamoto K, Ohmido N, Fukui K, Kamada H, Satoh S (2000) Site-specific accumulation of a LINE-like retrotransposon in a sex chromosome of the dioecious plant *Cannabis sativa*. Plant Mol Biol 44:723–732

Sakamoto K, Abe T, Matsuyama T, Yoshida S, Ohmido N, Fukui K, Satoh S (2005) RAPD markers encoding retrotransposable elements are linked to the male sex in *Cannabis sativa* L. Genome 48:931–936

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274:765–768

Schwartz S, Elnitski L, Li M, Weirauch M, Riemer C, Smit N (2003) MultiPipMaker and supporting tools: alignments and analysis of multiple genomic DNA sequences. Nucleic Acids Res 31:3518–3524

Shibata F, Hizume M, Kuroki Y (1999) Chromosome painting of Y chromosomes and isolation of a Y chromosome-specific repetitive sequence in the dioecious plant *Rumex acetosa*. Chromosoma 108:266–270

Soderlund C, Humphrey S, Dunham A, French L (2000) Contigs built with fingerprints, markers, and FPC v4.7. Genome Res 10:1772–1787

Song R, Llaca V, Messing J (2002) Mosaic organization of orthologous sequences in grass genomes. Genome Res 12:1549–1555

Stack SM, Comings DE (1979) The chromosomes and DNA of *Allium cepa.* Chromosoma 70:161–181

Staden R (1996) The Staden sequence analysis package. Mol Biotech 5:233–241

Stajner N, Bohanec B, Javornik B (2002) Genetic variability of economically important *Asparagus* species as revealed by genome size analysis and rDNA ITS polymorphisms. Plant Sci 162:931–937

Tarchini R, Biddle P, Wineland R, Tingey S, Rafalski A (2000) The complete sequence of 340 kb of DNA around the rice Adh1–Adh2 region reveals interrupted colinearity with maize chromosome 4. Plant Cell 12:381–392

Tatusova TA, Madden TL (1999) Blast 2 sequences—a new tool for comparing protein and nucleotide sequences. FEMS Microbiol Lett 174:247–250

Tikhonov A, SanMiquel P, Nakajima Y, Gorenstein N, Bennetzen J, Avramova Z (1999) Colinearity and its exceptions in orthologous adh regions of maize and sorghum. Proc Natl Acad Sci USA 96:7409–7441