

SHORT REPORT

Open Access

Construction and characterization of two BAC libraries representing a deep-coverage of the genome of chicory (*Cichorium intybus* L., Asteraceae)

Lucy Gonthier^{1†}, Arnaud Bellec^{2†}, Christelle Blassiau¹, Elisa Prat², Nicolas Helmstetter², Caroline Rambaud¹, Brigitte Huss¹, Theo Hendriks¹, H el ene Berg es^{2*}, Marie-Christine Quillet^{1*}

Abstract

Background: The Asteraceae represents an important plant family with respect to the numbers of species present in the wild and used by man. Nonetheless, genomic resources for Asteraceae species are relatively underdeveloped, hampering within species genetic studies as well as comparative genomics studies at the family level. So far, six BAC libraries have been described for the main crops of the family, *i.e.* lettuce and sunflower. Here we present the characterization of BAC libraries of chicory (*Cichorium intybus* L.) constructed from two genotypes differing in traits related to sexual and vegetative reproduction. Resolving the molecular mechanisms underlying traits controlling the reproductive system of chicory is a key determinant for hybrid development, and more generally will provide new insights into these traits, which are poorly investigated so far at the molecular level in Asteraceae.

Findings: Two bacterial artificial chromosome (BAC) libraries, CinS2S2 and CinS1S4, were constructed from *Hind*III-digested high molecular weight DNA of the contrasting genotypes C15 and C30.01, respectively. C15 was hermaphrodite, non-embryogenic, and S_2S_2 for the *S*-locus implicated in self-incompatibility, whereas C30.01 was male sterile, embryogenic, and S_1S_4 . The CinS2S2 and CinS1S4 libraries contain 89,088 and 81,408 clones. Mean insert sizes of the CinS2S2 and CinS1S4 clones are 90 and 120 kb, respectively, and provide together a coverage of 12.3 haploid genome equivalents. Contamination with mitochondrial and chloroplast DNA sequences was evaluated with four mitochondrial and four chloroplast specific probes, and was estimated to be 0.024% and 1.00% for the CinS2S2 library, and 0.028% and 2.35% for the CinS1S4 library. Using two single copy genes putatively implicated in somatic embryogenesis, screening of both libraries resulted in detection of 12 and 13 positive clones for each gene, in accordance with expected numbers.

Conclusions: This indicated that both BAC libraries are valuable tools for molecular studies in chicory, one goal being the positional cloning of the *S*-locus in this Asteraceae species.

* Correspondence: helene.berges@toulouse.inra.fr; marie-christine.quillet@univ-lille1.fr

† Contributed equally

¹Univ Lille Nord de France, F-59000 Lille, France, Stress Abiotiques et Diff erenciati on des V eg etaux Cultiv es (SADV), UMR INRA-USTL 1281, B at. SN2, F-59655 Villeneuve d'Ascq, France

²Centre National des Ressources G enomiques V eg etales (CNRGV), INRA, Chemin de Borde Rouge, BP 52627, F-31326 Castanet Tolosan, France
Full list of author information is available at the end of the article

Introduction

Chicory (*Cichorium intybus* L.) is an Asteraceae (Compositae) species belonging to the Lactuceae tribe of the Lactuoideae subfamily [1,2], as are lettuce and other vegetable crops (e.g. endive and salsify), and widespread weeds (e.g. dandelion). The genus *Cichorium* includes 6 diploid species ($2n = 18$) native from the Old World [3]. The closely related species chicory (*C. intybus* L.) and endive (*C. endivia* L.) have been domesticated and are mainly cultivated in Europe. Wild chicory is perennial [3] but the crop has been selected to be cultivated as a biennial species. Four main cultivar groups are distinguished, depending on the use of their roots in transformed products (root or industrial cultivar group) or the consumption of leaves as fresh or cooked vegetables (witloof, pain de sucre, and radicchio cultivar groups) [4]. Though these crops are not highly ranked on the scale of world's agriculture, they are important at the regional level in Belgium, the North of France, and The Netherlands, where the production of root and witloof chicory is concentrated. Similarly, about 85% of the Italian radicchio production is performed in the north eastern part of that country [4].

Wild forms of *C. intybus* have been reported to be strictly allogamous [3], and diallel cross analyses performed in witloof and radicchio cultivar groups showed that selfing is avoided by a strong sporophytic self-incompatibility system controlled by a single multiallelic *S*-locus [5,6]. As for many allogamous species, breeder's work in the early 1960's showed heterosis effects in progenies obtained from crosses between relatively distant selected genotypes, indicating that F_1 hybrids would be the cultivar type to develop for chicory [7]. Variability for the level of self-fertility was found in different cultivar groups [6,8,9], and in the cultivar group witloof, partly self-compatible genotypes were selected to produce inbred lines used as F_1 hybrids progenitors [10,11]. Selection pressure for self-fertility was not strongly applied in other cultivar groups.

Resolving the molecular mechanisms underlying traits controlling the reproductive system of chicory is a key determinant for chicory hybrid development, and will more generally provide new insights into these traits, which are poorly understood so far at the molecular level in Asteraceae [12-14]. To achieve these goals, both molecular genetic tools and genomic resources are necessary, and a consensus genetic map based on 472 transferable codominant SSR and SNP markers covering the 9 chromosomes of the *C. intybus* haploid genome was recently established [15]. Furthermore, over 80,000 chicory and endive EST sequences are now publicly available on the NCBI dbEST site [16,17] and the CGPDB site (Compositae Genome Project Database, http://cgpdb.ucdavis.edu/asteraceae_assembly/).

BAC libraries are essential tools for detailed characterisation of genomic regions containing genes of interest by the construction of physical maps and positional cloning. Despite their importance in terms of numbers of species present in the wild, and used by man [2], Asteraceae genomic resources are relatively underdeveloped, hampering the development of within species genetic studies, and comparative genomics studies at the family and clade levels [e.g. [18-20]]. So far, six BAC libraries have been described for Asteraceae species, and they are restricted to the main crops of the family, i.e. lettuce [21] and sunflower [22-25]. Here we present the characterization of BAC libraries of *C. intybus* constructed from two genotypes differing in traits related to sexual reproduction, i.e. self-incompatibility and male sterility, and vegetative reproduction, i.e. somatic embryogenesis. To demonstrate the utility of these genomic libraries, we used two cDNA probes corresponding to single copy genes putatively involved in somatic embryogenesis to screen the libraries in order to have access to the gene sequences.

Methods

Plant materials

Two different *C. intybus* genotypes (C15 and C30.01), selected from a Hungarian landrace population (Koospol), were used for the construction of BAC libraries. C15 was obtained after selfing of the K59 genotype, and is hermaphrodite, non-embryogenic [16], and homozygous S_2S_2 for the *S*-locus. C30.01 was obtained from an open pollination of C30, descendant from K59 upon selfing. C30.01 is male sterile, embryogenic, and heterozygous S_1S_4 at the *S*-locus. *S*-locus genotypes were determined by test crosses; dominance relationships between the *S*-alleles being $S_1 = S_4 > S_2$ for the pistil, and $S_1 > S_4 > S_2$ for the pollen. Both genotypes were *in vitro* propagated, either by organogenesis (C15) [26], or by somatic embryogenesis (C30.01) [27]. Ten clones of each genotype were acclimated in a glasshouse at 20°C, with a 16/8 h light/dark cycle, and grown for about three months. After transferring the plants 5 to 7 days in a dark room, young leaves were collected, frozen in liquid nitrogen and stored at -80°C before extracting high molecular weight (HMW) genomic DNA.

BAC libraries construction

HMW DNA of C15 and C30.01 genotypes was prepared as described by Peterson et al. [28] with the following modifications: (1) PVP was lowered to 0.25% w/v in sucrose-based extraction buffer (0.01 M Tris, 0.1 M KCl, 0.01 M EDTA pH 9.4, 500 mM sucrose, 4 mM spermidine, 1 mM spermine tetrahydrochloride, 0.1% w/v ascorbic acid, 0.25% w/v PVP-40, and 0.13% w/v sodium

diethyldithiocarbamate) as described in Chalhoub et al. [29], (2) lysis buffer was 1% w/v sodium lauryl sarcosine, 0.3 mg/ml proteinase K and 0.13% w/v sodium diethyldithiocarbamate dissolved in 0.5 M EDTA pH 9.1, (3) after lysis of the nuclei, agarose plugs were pre-washed 1 h in 0.5 M EDTA pH 9.1 at 50°C, 1 h in 0.05 M EDTA pH 8.0 at 4°C, and then stored at 4°C [30]. Embedded HMW DNA was partially digested with *Hind*III (New England Biolabs, Ipswich, Massachusetts), subjected to two size selection steps by pulsed-field electrophoresis, using a BioRad CHEF Mapper system (BioRad Laboratories, Hercules, California), and ligated to pIndigoBAC-5 *Hind*III-Cloning Ready vector (Epicentre Biotechnologies, Madison, Wisconsin). Pulsed-field migration programs, electrophoresis buffer and ligation desalting conditions were performed for the construction of CinS2S2 library as described in Peterson et al. [28]. These steps of the procedure were modified for the construction of CinS1S4 library, according to Chalhoub et al. [29], in order to increase the insert size (see Results and discussion).

After electroporation of T1 resistant DH10B electrocompetent cells (Invitrogen, Carlsbad, California), transformants were incubated in SOC medium for one hour, and 100 µL of transformants were plated on LB agarose medium containing 12.5 µg/mL chloramphenicol, 80 mg/mL X-gal and 100 mg/mL IPTG for colony counting and insert size estimation. Glycerol stocks (6%, v/v) of the remaining transformants (2.9 mL) were stored at -80°C, for later plating and colony picking. BAC DNA from randomly selected plated colonies of the two libraries was isolated using NucleoSpin fast purification kit (Macherey-Nagel, Düren, Germany). DNA from each BAC was dissolved in 30 µL, and half of this amount was digested with 10 units of *Not*I (New England Biolabs) to excise inserts. The digested BAC DNA was separated in 1% agarose gel by pulsed-field electrophoresis at 12°C in 0.25 × TBE buffer at 6 V/cm with a linear ramp pulse time of 5-15 s for 16 h. Insert sizes were estimated with reference to a Lambda ladder (New England Biolabs).

Colony picking was carried out using a robotic workstation QPix2 XT (Genetix, New Milton, Hampshire, UK) using a white/blue selection. White colonies were arranged in 384-well microtiter plates containing LB medium with chloramphenicol (12.5 µg/mL) supplemented with 6% (v/v) glycerol.

BAC libraries screening

High-density colony filters were prepared from the complete CinS2S2 and CinS1S4 libraries using a robotic workstation QPix2 XT (Genetix). BAC clones were spotted using a 5 × 5 pattern onto 22 × 22 cm Immobilon-Ny+ filters (Millipore Corporate, Billerica,

Massachusetts). On each filter, 27,648 unique clones were spotted in duplicate, and clones were grown at 37°C for 17 h. Filters were then processed as follows: (1) denaturation on Whatman paper soaked with a solution of 0.5 M NaOH and 1.5 M NaCl for 4 min at room temperature, and for 10 min at 100°C, (2) neutralization on Whatman paper soaked with 1 M Tris-HCl pH 7.4, and 1.5 M NaCl for 10 min, incubation in a solution of 0.25 mg/mL proteinase K (Sigma Aldrich, St. Louis, Missouri) for 45 min at 37°C, baking for 45 min at 80°C, and (3) fixation by UV on a Biolink 254 nm cross-linker (Thermo Fischer Scientific, Waltham, Massachusetts) with an energy of 120,000 µJoules. Radiolabelling of probes and hybridization of the filters were performed as described in Sambrook and Russel [31]. Hybridized filters were imaged with a Storm 860 PhosphorImager (GE Healthcare, Little Chalfont, UK), and analyses were performed using the HDFR software (Incogen, Williamsburg, Virginia).

To evaluate the level of contamination by organelle DNA in the libraries, filters were hybridized with four mitochondrial and four chloroplast specific probes generated by PCR using total (*i.e.* nuclear, mitochondrial and chloroplast) DNA from *C. intybus* (genotype K59) as template, and primers described in Table 1. PCR products were purified with the NucleoSpin Extract II kit (Macherey-Nagel). Labelling of the probes with [³²P] dCTP was performed by random priming using the Ready-To-Go DNA Labelling Beads kit (GE Healthcare), and unincorporated nucleotides were then removed using Illustra ProbeQuant G-50 Micro Columns (GE Healthcare).

Some positive BAC clones detected with mitochondrial specific probes were subjected to a fingerprint analysis. BAC DNA was isolated as described in the previous section, and was digested with 100 units of *Eco*RI (New England Biolabs) in a final volume of 60 µL. Thirty five µL of restriction product were separated in 0.7% agarose gel in 1 × TAE buffer at 80 V for 22 h, and fragments were visualised under UV after ethidium bromide staining. Fragment sizes were estimated with reference to the Raoul™ molecular weight marker (MP Biomedicals, Santa Ana, California).

DNA probes corresponding to chicory genes *CiAGP* (arabinogalactan protein), and *CiSTM* (*SHOOT MERISTEMLESS*, a *KNOTTED1-LIKE* transcription factor) were used to ascertain the utility of the libraries to obtain clones containing single copy genes. The partial cDNA sequence of *CiAGP* was obtained from subtractive EST libraries, constructed from leaf explants of two chicory genotypes differing for their embryogenic capacity when cultivated under somatic embryogenesis-inducing conditions [16]. The full length *CiSTM* cDNA was obtained after amplification with degenerated primers,

Table 1 Mitochondrial, chloroplast, and nuclear genes derived probes, and associated primer sequences used for the characterization of the CinS2S2 and CinS1S4 libraries

Probe ^a	Description	Accession numbers ^b	Primer sequences (5' → 3')	Tm (°C)	Size (bp)
<i>atpA</i>	F1-F0 ATP synthase, subunit alpha	X80469, X51422, AF034118, X52838	L: gatcttgtaagcgactg R: agtaatgcctgagtcgagc	66 64	1000
<i>atp9</i>	F1-F0 ATP synthase, subunit 9	X51895, DQ539624	L: aataggggcccggagctgc R: gaaagccatcattggggc	69 69	184
<i>cob</i>	Apocytochrome b	X98362, EF674047, EF674014, DQ916732	L: gagttatagcagtcctaggg R: ctagtagtaagcaatccgcc	55 58	697
<i>cox2</i>	Cytochrome c oxidase, subunit 2	AJ414385, EF547230, DQ004553, EF488904	L: attcaagacgcagcaacacc R: gtactacctcgctcattgag	65 56	264
<i>matK</i>	Maturase K	AJ633132	L: tggttcaggctcttcgctattgg R: cgccccctttgaagcaagaattg	70 68	398
<i>ndhF</i>	NADH dehydrogenase	AY504736	L: tacttgattgattctattctttg R: caacaagattaaagattaaaaaa	55 55	581
<i>rbcl</i>	Ribulose 1,5-biphosphate carboxylase/oxygenase, large subunit	L13652	L: ttgccgagataatggcctac R: ccaagatctcggtcagagc	64 64	336
<i>trnL-trnF</i>	Intergenic spacer tRNA-Leu (trnL)-tRNA-Phe (trnF) genes	FJ490769	L: gggccaagtcctctatcccc R: ctaccagctgagctatcccc	65 64	397
<i>CiAGP</i>	Arabinogalactan protein	DT212458	L: aaaccaaccaagactttgaccacg R: ccccttaagttccacaattac	58 64	418
<i>CiSTM</i>	<i>SHOOT MERSISTEMLESS</i> transcription factor	GU189066	L: gtaggatcatcatgtttgatggggtttggag R: ccctaccttttggaattca	73 64	344

^a*atpA*, *atp9*, *cob* and *cox2* correspond to mitochondrial sequences, *matK*, *ndhF*, *rbcl* and *trnL-trnF* to chloroplast sequences, and *CiAGP* and *CiSTM* to cDNA sequences.^b Multiple accession numbers for a probe indicate that primer pairs were designed from the consensus sequence of these accessions.

and RACE PCR of cDNA obtained from leaf explants of the embryogenic interspecific (*C. intybus* × *C. endivia*) genotype 474 [32] cultivated under somatic embryogenesis-inducing conditions (unpublished). *CiAGP* and *CiSTM* probes were generated by PCR from cDNA clones by using specific primers (Table 1). Positive clones detected by hybridization were validated individually by PCR amplification using the primer pairs used for probes synthesis, and visualisation of PCR products after agarose gel electrophoresis.

Results and discussion

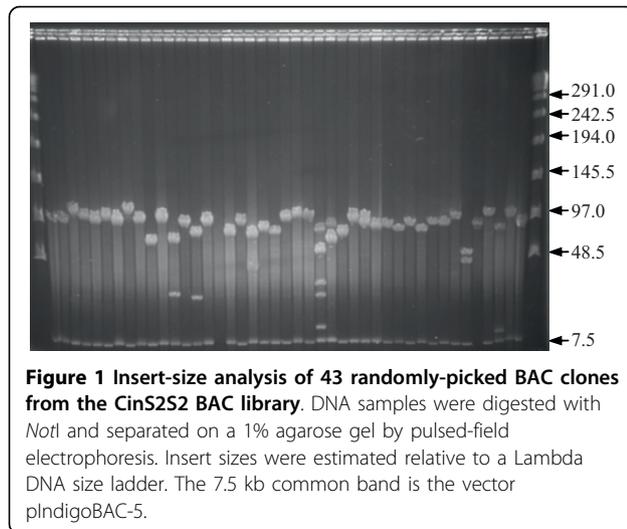
BAC libraries construction, and general features

Two BAC libraries were constructed in pIndigoBAC-5 with *HindIII* from DNA of two *C. intybus* genotypes differing in traits related to sexual and vegetative reproduction. Names of the libraries, CinS2S2 and CinS1S4, refer to the *S*-locus genotypes of the plants used to prepare HMW DNA, being *S*₂*S*₂, for C15, and *S*₁*S*₄ for C30.01. The main characteristics of the BAC libraries are summarized in Table 2. The CinS2S2 and CinS1S4 libraries consist of 89,088 and 81,408 clones, respectively, arrayed in a total of 444 384-well microtiter plates. A random selection of 178 clones for CinS2S2 and 303 clones for CinS1S4 was analysed after *NotI* digestion, followed by pulsed-field gel electrophoresis (Figure 1). From the clones sampled, the number of clones without insert was 1 (0.56%) and 8 (2.64%) for the CinS2S2 and CinS1S4 libraries, respectively. The average insert sizes of the CinS2S2 and

CinS1S4 libraries were calculated to be about 90 kb and 120 kb, respectively (Figure 2). More than 70% of clones from CinS2S2 and 90% for CinS1S4 libraries showed an insert size greater than 75 kb. The higher mean insert size obtained for CinS1S4 library could be explained by modifications of the first sizing step applied for the construction of CinS2S2 library. During the construction of this library, partial digested DNA was separated using the following parameters: 6 V/cm, 18 h, 120°, 1-40 s linear ramping in 0.5 × TBE. For the CinS1S4 library construction performed later on, two programs were aggregated: 6 V/cm, 18 h, 120°, 1-40 s linear ramping in 0.25 × TBE directly followed, in the same gel without removing or handling the DNA, by 5 V/cm, 8 h, 120°, 4-5 s ramping linear in the same buffer as described by Chalhoub et al. [29].

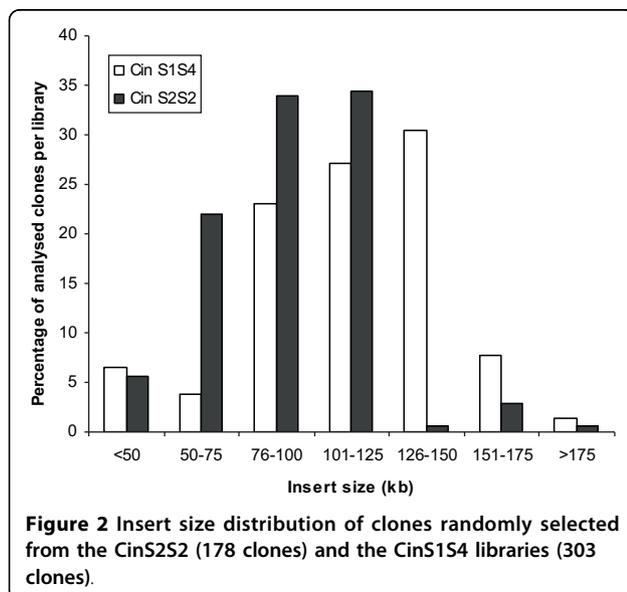
Table 2 Characteristics of CinS2S2 and CinS1S4 BAC libraries

Library	CinS2S2	CinS1S4	Total
No. of clones	89,088	81,408	170,496
No. of 384-well plates	232	212	444
Average insert size (kb)	91.78	118.20	104.40
Empty clones (%)	0.56	2.64	1.55
mtDNA clones (%)	0.02	0.03	0.03
cpDNA clones (%)	1.00	2.35	1.65
Nuclear haploid genome equivalents	5.75	6.53	12.28



Screening with mitochondrial and chloroplast genome probes

The mitochondrial (mt) genome of chicory has been estimated to be 300-400 kb [33], but genome structure and gene order are unknown. Similarly, in other Asteraceae species, mt genomes are poorly documented, and no complete sequences are publicly available. The probe mixture of four chicory mitochondrial genes (*atpA*, *atp9*, *cob*, and *cox2*) led to the detection of 43 and 29 positive clones in the CinS2S2 and CinS1S4 libraries, respectively. A set of 10 positive clones were randomly selected from each library, and subjected to PCR amplification with primers designed for the mt probes production. Five clones for library CinS2S2 and 8 clones for library CinS1S4 gave PCR products of the expected sizes for at least one of the targeted genes (Table 3). Three



clones gave a PCR product only with *cob* primers, and four clones with *atpA*, *atp9* and *cox2* primers. The remaining six clones could be amplified with *atpA* primers (1 clone), *atpA* and *cox2* primers (2 clones), or *atpA* and *atp9* primers (3 clones). To investigate whether the BAC clones containing the *cob* gene were possibly contiguous with those containing other mt genes, the 13 PCR validated BAC clones were subjected to *EcoRI* digestion for fingerprint analysis. Assuming that bands of the same size estimated after electrophoresis represent identical DNA fragments, pair-wise comparisons of fingerprints revealed that BAC clones of which the inserts contained the *cob* gene shared at least 2 fragments (Table 4). Similarly, at least 3 fragments of the same size were found present in fingerprints of BAC clones with inserts containing two or three other mt genes. The smallest insert (clone 09A06) containing the 3 genes *atpA*, *atp9*, and *cox2*, had an estimated size of 80 kb, delimiting the maximum interval containing these 3 genes (Table 4). However, the fingerprint of this clone did not show any fragment in common with fingerprints of *cob*-containing clones. In contrast, fingerprints of 3 other BAC clones with larger inserts containing the genes *atpA*, *atp9*, and *cox2* (92M03 and 31K22), or the genes *atpA* and *atp9* (96A09), shared 2 fragments (9.4 and 11.2 kb, respectively) with fingerprints of *cob*-containing clones (06I01, 26G13, and 25G18) (Table 4), suggesting that the inserts of these BAC clones partly overlapped. Considering all informative fingerprint fragments, the construction of a single contig encompassing the 4 mt genes was only possible when the BAC clone inserts are assumed to originate from a circular mt genome. As only 2 fragments were found in common between the *cob*-containing clones and the other clones, the *cob* gene may be located in a more distant part of the mt genome. Further experiments are necessary to determine the physical distances between these 4 genes. In sunflower, *atpA* and *cob* genes were reported to be only 10 kb apart, in a region associated with cytoplasmic male sterility [34,35]. However, because of frequent genome rearrangements, the organization of mitochondrial genes can vary between related species, and even within a single species [36].

Complete sequencing of chloroplast (cp) genomes in *Helianthus annuus* and *Lactuca sativa* indicated that genome size (151.1-152.7 kb), and gene content and order are the same in these two species, representing distantly related subfamilies in Asteraceae [37]. The four sequences (*matK*, *ndhF*, *rbcL* and *trnL-trnF*), chosen to evaluate cp contaminations, are distributed over a region of approximately 120 kb of the consensus cp genome of Asteraceae, the most distant genes (*ndhF* and *rbcL*) being separated by a distance of about 70 kb. The mixture of four probes led to the detection of 994 and 1919

Table 3 PCR amplifications of mitochondrial and chloroplast genome specific sequences on ten randomly selected clones of the CinS2S2 and CinS1S4 libraries detected after filter hybridization

	Mitochondrial genome specific sequences				Chloroplast genome specific sequences					
	Clone	<i>cox2</i>	<i>atpA</i>	<i>atp9</i>	<i>cob</i>	Clone	<i>matK</i> 2 kb*	<i>trnI-trnf</i> 47 kb*	<i>rbcL</i> 54 kb*	<i>ndhF</i> 124 kb*
CinS2S2 (90 kb)	06I01	-	-	-	+	02L02	+	-	-	-
	09A06	+	+	+	-	04M20	-	-	-	-
	09E24	-	-	-	-	06I22	+	-	-	+
	09I04	-	-	-	-	12E23	-	+	-	-
	13F07	-	-	-	-	26G24	+	+	+	-
	13M05	-	+	-	-	32H03	-	+	+	+
	26G13	-	-	-	+	37C04	+	+	-	-
	27J01	-	-	-	-	41L23	-	+	+	+
	41H20	+	+	+	-	51H20	-	-	+	-
	54B09	-	-	-	-	61O16	-	+	+	-
CinS1S4 (120 kb)	01H06	+	+	-	-	02N01	-	+	+	+
	16P11	+	+	-	-	21F01	+	-	-	+
	22B14	-	-	-	-	45D02	+	+	+	+
	25G18	-	-	-	+	51D03	-	-	+	+
	27O16	-	+	+	-	55M03	+	+	+	+
	31K22	+	+	+	-	62G03	+	+	+	+
	91G03	-	+	+	-	72G02	-	+	+	+
	92M03	+	+	+	-	84D02	+	-	-	+
	96A09	-	+	+	-	91N03	+	+	+	+
	108A12	-	-	-	-	106P01	+	+	+	+

+ indicates the presence of a PCR product of the expected size after agarose gel electrophoresis;
 - indicates no amplification; * Sequence coordinates on *H. annuus* and *L. sativa* chloroplast genome according to [37].

positive clones in the CinS2S2 and CinS1S4 libraries, respectively. PCR validation with specific primers performed on 10 positive clones of each library gave PCR products of expected size for at least one primer pair for 9 clones of the CinS2S2 library, and the 10 clones selected from CinS1S4 library (Table 3). A range of 1-3 (CinS2S2), and 2-4 (CinS1S4) PCR products per clone were detected, and the PCR data confirm the gene order described for the cp genome of *H. annuus* and *L. sativa*

[37]. In addition, five of the 10 clones sampled in the CinS1S4 library produced amplification products with the four primer pairs, suggesting that a large continuous fragment of the cp genome is present in these clones.

Based on the proportion of clones validated by PCR, we estimated that 0.024% and 0.028% of the BAC clones would result from mtDNA contamination, and that 1.00% and 2.35% of the clones contain putative cpDNA in CinS2S2 and CinS1S4 libraries, respectively. These

Table 4 Fingerprint analysis of BAC clones containing mt genes

Library	Clone	mt gene present	Insert size (kb)	Informative fingerprint fragment (kb)						
				4.8 4.9	11.2	5.6	11.7 9.0 8.8	7.0	9.4	3.6 4.4
CinS2S2	09A06	<i>cox2, atpA, atp9</i>	80	-	-	+	+	-	-	-
CinS1S4	96A09	<i>atpA, atp9</i>	150	-	-	-	+	+	+	-
CinS1S4	92M03	<i>cox2, atpA, atp9</i>	150	-	-	+	+	+	+	-
CinS1S4	31K22	<i>cox2, atpA, atp9</i>	95	-	+	+	+	+	-	-
CinS2S2	06I01	<i>cob</i>	55	+	+	-	-	-	+	+
CinS2S2	26G13	<i>cob</i>	60	+	+	-	-	-	-	+
CinS1S4	25G18	<i>cob</i>	25	-	-	-	-	-	+	+

DNA of BAC clones was digested with *NotI* for insert size determination, and with *EcoRI* for fingerprint analysis. Fingerprint fragments ranging from 3 to 12 kb were analysed, and fragments were considered as informative when present in at least 2 clones. Only the results for clones containing at least *atpA* and *atp9*, and sharing a fragment with *cob*-containing clones are shown. Clone 09A06 was included as it contained the smallest insert with the 3 genes *atpA*, *atp9*, and *cox2*. + indicates presence of a fingerprint fragment; - indicates absence of a fingerprint fragment.

results are consistent with: (1) the generally low frequency of mtDNA contamination (0-0.3%) reported in recent plant BAC libraries characterizations [38,39], and (2) the range of commonly observed frequencies (0.3-2.5%) attributed to cpDNA contamination [40,41].

The detected BAC clones containing mtDNA and cpDNA sequences could represent valuable tools as a starting point to investigate the structure of organelle genomes in chicory, particularly the mt genome involved in cytoplasmic male sterility [33,42]. However, the possibility that some of these organelle DNA contaminants actually are the result of horizontal transfer of DNA from mitochondria and/or plastid to the nucleus, as has been observed in most of the plant species studied [43-45], cannot be excluded. Whether large organelle DNA fragments exist in Asteraceae nuclear genomes remains to be determined; chromosome fluorescent *in situ* hybridization experiments [46] using chicory mtDNA- and cpDNA-containing BAC clones would be an approach to investigate this question.

Theoretical and empirical genome coverage

C. intybus has a haploid genome size estimated by flow cytometry to be between 1300 Mb (P. Devaux, personal communication) and 1400 Mb [47]. After subtraction of non-recombinant and putative organelle DNA contaminant clones, considering a haploid genome size of 1400 Mb, and mean insert sizes of 90 kb for CinS2S2 and 120 kb for CinS1S4, the two libraries represent approximately 5.7 and 6.5 haploid genome equivalents, respectively. Together these two libraries provide a coverage of the *C. intybus* genome of about 12.3-fold (Table 2), resulting in a theoretical probability superior to 0.9999 of recovering any targeted genomic sequence [48]. In comparison, the lettuce BAC library described by Frijters et al. [21] covers about 2 genome equivalents, and the 5 combined sunflower BAC libraries published represent about 20 genome equivalents [22-25]. This indicates that the chicory libraries represent a significant genomic resource to investigate Asteraceae genomes.

To validate empirically the genome coverage calculated for both libraries, BAC clones were screened with two cDNA probes (*CiAGP* and *CiSTM*). Mapping data for each of the corresponding genes [49] revealed a unique locus position in the chicory map of Cadalen et al. [15], indicating that they represent single copy sequences. Hybridization allowed the identification of 13 positive clones for *CiAGP*: 2 clones in the CinS2S2 library and 11 clones in the CinS1S4 library. The *CiSTM* probe revealed 5 positive clones in the CinS2S2 library and 7 positive clones in the CinS1S4 library, *i.e.* a total of 12 clones in the combined libraries. The clones detected by hybridization were subjected to PCR amplification with the primer pairs used for the

synthesis of the hybridization probes, and agarose gel electrophoresis of PCR products showed amplicons of the expected size for all of them. As for the *CiAGP* gene, an under- and over-representation was found in the CinS2S2 and CinS1S4 libraries, respectively. Further screening with a larger set of single copy sequence probes are necessary to determine if significant differences between libraries exist as for their content in single copy sequences. Nonetheless, experimental results are in accordance with the theoretical genome coverage estimated for the combined libraries, indicating that they constitute an efficient tool to recover single copy genes or genomic sequences for chicory.

Perspective

One future use of the *C. intybus* BAC libraries will be the positional cloning of the *S*-locus in this species. As in some other plant families, Asteraceae species exhibit a sporophytic self-incompatibility (SSI) system where the pollen phenotype depends on the genotype of the plant that produced the pollen [50]. The pollen and stigma determinants of SSI have been characterised in *Brassica* [51,52], and molecular and genetic studies showed that these genes are different from those involved in the SSI response in Convolvulaceae and Asteraceae, suggesting that different molecular mechanisms of SSI have evolved independently [12,53]. More recently, candidate genes with unknown functions have been identified in *Ipomoeae trifida* (Convolvulaceae) through a positional cloning approach [54]. Similarly, we have chosen for genetic and subsequent physical mapping of the *S*-locus in chicory to identify pollen and stigma determinants of SSI in this Asteraceae species.

Library availability

The libraries described are accessible for scientific research based on collaboration with the laboratories represented by the corresponding authors.

Acknowledgements

The authors thank Bruno Desprez (Florimond Desprez, Cappelle-en-Pévèle, France) for providing seeds of the Koospol population. Part of this work was financed by a 'Bonus Qualité Recherche' of USTL to the UMR 1281. Lucy Gonthier was supported by a doctoral fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche, France. We also like to thank the three anonymous reviewers for their valuable comments on a previous version of the paper.

Author details

¹Univ Lille Nord de France, F-59000 Lille, France, Stress Abiotiques et Différenciation des Végétaux Cultivés (SADV), UMR INRA-USTL 1281, Bât. SN2, F-59655 Villeneuve d'Ascq, France. ²Centre National des Ressources Génétiques Végétales (CNRGV), INRA, Chemin de Borde Rouge, BP 52627, F-31326 Castanet Tolosan, France.

Authors' contributions

The project was conceived and supervised by MCQ and TH. LG, AB, CB, and NH realized the construction of the two BAC libraries under the supervision

of HB in the CNR GV - INRA, Toulouse. EP performed the hybridizations using probes provided by CR and BH, and NH realized the fingerprint experiment. All authors contributed to the analyses and interpretations of the results, took part in the redaction of the manuscript, and approve its final version.

Competing interests

The authors declare that they have no competing interests.

Received: 8 June 2010 Accepted: 11 August 2010

Published: 11 August 2010

References

- Bremer K: *Asteraceae: Cladistics and classification*. Portland, Oregon: Timber Press 1994.
- Funk VA, Bayer RJ, Keeley S, Chan R, Watson L, Gemeinholzer B, Schilling E, Panero JL, Baldwin BG, Garcia-Jacas N, et al: **Everywhere but Antarctica: using a supertree to understand the diversity and distribution of the Compositae**. *Biologiske Skrifter* 2005, **55**:343-374.
- Kiers AM, Mes THM, Van der Meijden R, Bachmann K: **Morphologically defined *Cichorium* (Asteraceae) species reflect lineages based on chloroplast and nuclear (ITS) DNA data**. *Syst Bot* 1999, **24**(4):645-659.
- Lucchin M, Varotto S, Barcaccia G, Parrini P: **Chicory and endive**. In *Vegetables I*. Edited by: Prohen J, Nuez F. New-York: Springer; 2008:3-48.
- Eenink AH: **Compatibility and incompatibility in witloof-chicory (*Cichorium intybus* L.)**. 2. The incompatibility system. *Euphytica* 1981, **30**(1):77-85.
- Varotto S, Pizzoli L, Lucchin M, Parrini P: **The incompatibility system in Italian red chicory (*Cichorium intybus* L.)**. *Plant Breed* 1995, **114**(6):535-538.
- Bannerot H, Deconinck B: **Travaux d'amélioration de la chicorée witloof à l'INRA**. *Annales de Gembloux* 1965, **71**:143-158.
- Coppens d'Eeckenbrugge G, Evrard B, Louant B-P: **Aspects quantitatifs de l'auto-incompatibilité chez la chicorée de witloof (*Cichorium intybus* L.)**. Relation entre fertilité et autofertilité. *Revue de l'Agriculture* 1987, **40**(1):35-40.
- Castano CI, Demeulemeester MAC, De Proft MP: **Incompatibility reactions and genotypic identity status of five commercial chicory (*Cichorium intybus* L.) hybrids**. *Sci Hortic* 1997, **72**(1):1-9.
- Bannerot H, Deconinck B: **L'utilisation des hybrides F1: nouvelle méthode d'amélioration de la chicorée de bruxelles**. *Symposium interne section horticole d'Eucarpia* Gembloux, Belgium 1970, 99-118.
- Doré C, Varoquaux F: **La chicorée endive ou witloof**. *Histoire et amélioration de cinquante plantes cultivées* Quae 2006, 207-217.
- Hiscock SJ, McInnis SM, Tabah DA, Henderson CA, Brennan AC: **Sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae) - the search for S**. *J Exp Bot* 2003, **54**(380):169-174.
- McInnis SM, Costa LM, Gutierrez-Marcos JF, Henderson CA, Hiscock SJ: **Isolation and characterization of a polymorphic stigma-specific class III peroxidase gene from *Senecio squalidus* L. (Asteraceae)**. *Plant Mol Biol* 2005, **57**(5):659-677.
- Tabah DA, McInnis SM, Hiscock SJ: **Members of the S-receptor kinase multigene family in *Senecio squalidus* L. (Asteraceae), a species with sporophytic self-incompatibility**. *Sex Plant Reprod* 2004, **17**(3):131-140.
- Cadalen T, Mörchen M, Blassiau C, Clabaut A, Scheer I, Hilbert J-L, Hendriks T, Quillet M-C: **Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.)**. *Mol Breed* 2010, **25**(4):699-722.
- Legrand S, Hendriks T, Hilbert JL, Quillet MC: **Characterization of expressed sequence tags obtained by SSH during somatic embryogenesis in *Cichorium intybus* L.** *BMC Plant Biol* 2007, **7**:27.
- Dauchot N, Mingeot D, Purnelle B, Muys C, Watillon B, Boutry M, Van Cutsem P: **Construction of 12 EST libraries and characterization of a 12,226 EST dataset for chicory (*Cichorium intybus*) root, leaves and nodules in the context of carbohydrate metabolism investigation**. *BMC Plant Biol* 2009, 9:14.
- Feng Q, Huang T, Zhao Q, Zhu JJ, Lin ZX, Han B: **Analysis of collinear regions of *Oryza* AA and CC genomes**. *J Genet Genomics* 2009, **36**(11):667-677.
- Yan HH, Mudge J, Kim DJ, Shoemaker RC, Cook DR, Young ND: **Comparative physical mapping reveals features of microsynteny between *Glycine max*, *Medicago truncatula*, and *Arabidopsis thaliana***. *Genome* 2004, **47**(1):141-155.
- Cenci A, Combes MC, Lashermes P: **Comparative sequence analyses indicate that *Coffea* (Asterids) and *Vitis* (Rosids) derive from the same paleo-hexaploid ancestral genome**. *Mol Gen Genet* 2010, **283**(5):493-501.
- Frijters ACJ, Zhang Z, vanDamme M, Wang GL, Ronald PC, Michelmore RW: **Construction of a bacterial artificial chromosome library containing large *EcoRI* and *HindIII* genomic fragments of lettuce**. *Theor Appl Genet* 1997, **94**(3-4):390-399.
- Gentzbittel L, Abbott A, Galaud JP, Georgi L, Fabre F, Liboz T, Alibert G: **A bacterial artificial chromosome (BAC) library for sunflower, and identification of clones containing genes for putative transmembrane receptors**. *Mol Gen Genet* 2002, **266**(6):979-987.
- Ozdemir NR, Horn R, Friedt W: **Construction and characterization of a BAC library for sunflower (*Helianthus annuus* L.)**. *Euphytica* 2004, **138**(2):177-183.
- Bouzidi MF, Franchel J, Tao QZ, Stormo K, Mraz A, Nicolas P, Mouzeyar S: **A sunflower BAC library suitable for PCR screening and physical mapping of targeted genomic regions**. *Theor Appl Genet* 2006, **113**(1):81-89.
- Feng JH, Vick BA, Lee MK, Zhang HB, Jan CC: **Construction of BAC and BIBAC libraries from sunflower and identification of linkage group-specific clones by overgo hybridization**. *Theor Appl Genet* 2006, **113**(1):23-32.
- Decout E, Dubois T, Guedira M, Dubois J, Audran JC, Vasseur J: **Role of temperature as a triggering signal for organogenesis or somatic embryogenesis in wounded leaves of chicory cultured *in vitro***. *J Exp Bot* 1994, **45**(281):1859-1865.
- Bellettre A, Couillerot JP, Blervacq AS, Aubert S, Gout E, Hilbert JL, Vasseur J: **Glycerol effects both carbohydrate metabolism and cytoskeletal rearrangements during the induction of somatic embryogenesis in chicory leaf tissues**. *Plant Physiol Biochem* 2001, **39**(6):503-511.
- Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH: **Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide**. *Journal of Agricultural Genomics* 2000, 5.
- Chalhoub B, Belcram H, Caboche M: **Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size**. *Plant Biotechnol J* 2004, **2**(3):181-188.
- Construction and manipulation of large-insert bacterial clone libraries-manual. [http://hbz7.tamu.edu/homelinks/tool/protocol.htm].
- Sambrook J, Russel DW: **Hybridization of bacterial DNA on filters**. *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor Laboratory Press NY, Third 2001, 138-142.
- Dubois T, Guedira M, Dubois J, Vasseur J: **Direct somatic embryogenesis in leaves of *Cichorium* - a histological and SEM study of early stages**. *Protoplasma* 1991, **162**(2-3):120-127.
- Rambaud C, Dubois J, Vasseur J: **Male-sterile chicory cybrids obtained by intergeneric protoplast fusion**. *Theor Appl Genet* 1993, **87**(3):347-352.
- Köhler RH, Horn R, Lossl A, Zetsche K: **Cytoplasmic male-sterility in sunflower is correlated with the co-transcription of a new open reading frame with the ATPa gene**. *Mol Gen Genet* 1991, **227**(3):369-376.
- Laver HK, Reynolds SJ, Moneger F, Leaver CJ: **Mitochondrial genome organization and expression associated with cytoplasmic male-sterility in sunflower (*Helianthus annuus*)**. *Plant J* 1991, **1**(2):185-193.
- Kubo T, Newton KJ: **Angiosperm mitochondrial genomes and mutations**. *Mitochondrion* 2008, **8**(1):5-14.
- Timme RE, Kuehl JV, Boore JL, Jansen RK: **A comparative analysis of the *Lactuca* and *Helianthus* (Asteraceae) plastid genomes: Identification of divergent regions and categorization of shared repeats**. *Am J Bot* 2007, **94**(3):302-312.
- Gutman W, Pawelkowicz M, Woycicki R, Piszczek E, Przybecki Z: **The construction and characteristics of a BAC library for *Cucumis sativus* L. 'B10'**. *Cell Mol Biol Lett* 2008, **13**(1):74-91.
- Baig MNR, Yu A, Guo WW, Deng XX: **Construction and characterization of two *Citrus* BAC libraries and identification of clones containing the phytoene synthase gene**. *Genome* 2009, **52**(5):484-489.
- Ruffel S, Caranta C, Palloix A, Lefebvre V, Caboche M, Bendahmane A: **Structural analysis of the eukaryotic initiation factor 4E gene controlling potyvirus resistance in pepper: exploitation of a BAC library**. *Gene* 2004, **338**(2):209-216.
- Yasui Y, Mori M, Matsumoto D, Ohnishi O, Campbell CG, Ota T: **Construction of a BAC library for buckwheat genome research - an application to positional cloning of agriculturally valuable traits**. *Genes Genet Syst* 2008, **83**(5):393-401.

42. Dubreucq A, Berthe B, Asset JF, Boulidard L, Budar F, Vasseur J, Rambaud C: **Analyses of mitochondrial DNA structure and expression in three cytoplasmic male-sterile chicories originating from somatic hybridisation between fertile chicory and CMS sunflower protoplasts.** *Theor Appl Genet* 1999, **99**(7-8):1094-1105.
43. Richly E, Leister D: **NUMTs in sequenced eukaryotic genomes.** *Mol Biol Evol* 2004, **21**(6):1081-1084.
44. Richly E, Leister D: **NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs.** *Mol Biol Evol* 2004, **21**(10):1972-1980.
45. Leister D: **Origin, evolution and genetic effects of nuclear insertions of organelle DNA.** *Trends Genet* 2005, **21**(12):655-663.
46. Stupar RM, Lilly JW, Town CD, Cheng Z, Kaul S, Buell CR, Jiang JM: **Complex mtDNA constitutes an approximate 620-kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors caused by large-unit repeats.** *Proc Natl Acad Sci USA* 2001, **98**(9):5099-5103.
47. Koopman WJM: **Zooming in on the lettuce genome: species relationships in *Lactuca* s.l., inferred from chromosomal and molecular characters.** *PhD Thesis Wageningen University, Nederlands* 2002.
48. Clarke L, Carbon J: **A colony bank containing synthetic Col EI hybrid plasmids representative of the entire *E. coli* genome.** *Cell* 1976, **9**(1):91-99.
49. Clabaut A: **Analyse génétique de l'embryogenèse somatique chez la chicorée (*Cichorium intybus* L.): cartographie des QTL et gènes candidats.** *PhD Thesis Lille, France: University of Lille* 2009, **1**.
50. de Nettancourt D: **Incompatibility in angiosperms.** New York, USA: Springer-Verlag 1977.
51. Stein JC, Howlett B, Boyes DC, Nasrallah ME, Nasrallah JB: **Molecular cloning of a putative receptor protein-kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*.** *Proc Natl Acad Sci USA* 1991, **88**(19):8816-8820.
52. Schopfer CR, Nasrallah ME, Nasrallah JB: **The male determinant of self-incompatibility in *Brassica*.** *Science* 1999, **286**(5445):1697-1700.
53. Kowiyama Y, Tsuchiya T, Kakeda K: **Sporophytic self-incompatibility in *Ipomoea trifida*, a close relative of sweet potato.** *Ann Bot* 2000, **85**:191-196.
54. Rahman MH, Uchiyama M, Kuno M, Hirashima N, Suwabe K, Tsuchiya T, Kagaya Y, Kobayashi I, Kakeda K, Kowiyama Y: **Expression of stigma- and anther-specific genes located in the S locus region of *Ipomoea trifida*.** *Sex Plant Reprod* 2007, **20**(2):73-85.

doi:10.1186/1756-0500-3-225

Cite this article as: Gonthier et al.: Construction and characterization of two BAC libraries representing a deep-coverage of the genome of chicory (*Cichorium intybus* L., Asteraceae). *BMC Research Notes* 2010 **3**:225.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

