Theor Appl Genet (2012) 124:249–259
DOI 10.1007/s00122-011-1701-x

ORIGINAL PAPER

Construction of a high density integrated genetic map for cucumber (*Cucumis sativus* L.)

Wei-Wei Zhang · Jun-Song Pan · Huan-Le He · Chi Zhang · Zheng Li · Jun-Long Zhao · Xiao-Jun Yuan · Li-Huang Zhu · San-Wen Huang · Run Cai

Received: 17 March 2011 / Accepted: 7 September 2011 / Published online: 5 October 2011

Abstract The high-density consensus map was constructed based on the GY14 × PI 183967 map from an inter-subspecific cross and the extended S94 × S06 map from an intra-subspecific cross. The consensus map was composed of 1,369 loci, including 1,152 SSR loci, 192 SRAP loci, 21 SCAR loci and one STS locus as well as three gene loci of fruit external quality traits in seven chromosomes, and spanned 700.5 cM, of which 682.7 cM (97.5%) were covered by SSR markers. The average genetic distance and physical interval between loci were 0.51 cM and ~268 kbp, respectively. Additionally, the physical position of the sequence-associated markers aligned along the assembled cucumber genome sequence established a relationship between genetic maps and cucumber genome sequence and to a great extent validated the order of markers in individual maps and consensus map. This consensus map with a high marker density and well-ordered markers is a saturated and reliable linkage map for genetic analysis of cucumber or the Cucurbitaceae family of plants.

Communicated by M. Havey.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1701-x) contains supplementary material, which is available to authorized users.

W.-W. Zhang · J.-S. Pan · H.-L. He · C. Zhang · Z. Li · J.-L. Zhao · X.-J. Yuan · R. Cai (✉)
School of Agriculture and Biology,
Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: cairun@sjtu.edu.cn

W.-W. Zhang
Shanghai Chenshan Plant Science Research Center,
Chinese Academy of Sciences, Chenshan Botanical Garden,
Shanghai 201602, China

L.-H. Zhu
The State Key Laboratory of Plant Genomics and National Plant Gene Research Centre (Beijing), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences,
Beijing 100101, China

S.-W. Huang (✉)
Key Laboratory of Horticultural Crops Genetic Improvement of Ministry of Agriculture, Sino-Dutch Joint Lab of Horticultural Genomics Technology, Institute of Vegetables and Flowers,
Chinese Academy of Agricultural Sciences,
Beijing 100081, China
e-mail: huangsanwen@caas.net.cn

Abbreviations

SSR	Simple sequence repeat
SRAP	Sequence-related amplified polymorphism
SCAR	Sequence characterized amplified region
STS	Sequence tagged site
RIL	Recombinant inbred lines

Background

Cucumber (*Cucumis sativus* L.), which belongs to the family Cucurbitaceae, is one of the most important vegetable plants grown worldwide. Due to having fewer chromosome numbers (seven chromosomes), hundreds of known functional genes (Xie and Wehner 2001), and a smaller-sized genome (367 Mb) than other economically important Cucurbitaceae crops such as melon (480 Mb), watermelon (430 Mb), squash and pumpkin (539 Mb) (Ren et al. 2009), cucumber has been a model plant for genetic research of the Cucurbitaceae crops. Recently, Huang et al. (2009) selected the 'Chinese long' inbred line 9,930 and first reported the draft genome sequence of cucumber (Cucumber Genome DataBase http://cucumber.genomics.org.cn/page/cucumber/index.jsp) with the total length of 243.5 Mb, about 70% of

the genome size estimated by Arumuganathan (1991) (367 Mb). Cucumber genome sequence published promotes the gene map-based cloning of important agronomic traits in cucumber. For example, Li et al. (2009) used the cucumber genome sequence provided by Huang et al. (2009) to successfully clone and identify the *M/m* gene controlling cucumber flower sexuality.

Although the cucumber genome sequence has been published, it is essential to construct a high-density and representative cucumber genetic map for facilitating the use of the cucumber genome sequence data in map-based gene isolation. Recently, Ren et al. (2009) published a genetic map of cucumber including 995 SSR markers spanning 572.9 cM, which was the most saturated linkage map published in cucumber or even in the Cucurbitaceae family. The map contained a great deal of sequence-based molecular markers and anchored linkage groups on the cucumber chromosomes by fluorescence in situ hybridization (FISH) technology. However, the low population numbers (77 RILs of Gy14 × PI 183967) would supply insufficient genetic recombination information, thus leading to low reliability of position markers mapped on the genetic map. In fact, this is also one of the reasons that the published cucumber genome sequence cannot have high ratio coverage of the entire genome.

So far, by using all types of markers including morphologic, isozyme, and various DNA molecular markers, over ten genetic maps have been constructed in cucumber (Kennard et al. 1994; Serquen et al. 1997; Park et al. 2000; Bradeen et al. 2001; Fazio et al. 2003; Li et al. 2005; Wang et al. 2005; Sun et al. 2006;Yuan et al. 2008a, b; Ren et al. 2009), of which two maps other than the Gy14 × PI 183967 map published by Ren et al. (2009) are relatively saturated, and they approach the total genomic distance (750–1,000 cM) estimated by Staub and Meglic (1993). The map was developed by Fazio et al. (2003) from the 171 individuals of G421 × H-19, consisted of 131 marker loci (27AFLPs, 62 RAPDs, 14 SSRs, 24 SCARs, one SNP, and three MTM) and spaned 706 cM. The other map from 224 RILs of S94 × S06 (Yuan et al. 2008b) included 257 molecular markers (206 SRAPs, 22 SSRs, 25 SCARs, one STS and three MTM) and spaned 1,005.8 cM. Many important agronomic traits, involving qualitative loci (Yuan et al. 2008b; Zhang et al. 2010) and quantitative trait loci (QTL) of flower-related and fruit-related traits (Yuan et al. 2008b) and lateral branch-related traits (Jiang et al. 2008), were mapped in S94 × S06 map. Compared with the Gy14 × PI 183967 map, most previously published genetic maps were constructed by using larger mapping groups (>100 individuals) with the various agronomic traits loci. However, these maps included fewer marker numbers and lower saturation levels, fewer sequence-based markers, and did not correspond to the linkage groups of the genetic map to the cucumber chromosomes. Therefore, it was difficult to compare the genetic maps constructed by different laboratories, which limits the application of these maps in the molecular breeding of cucumber.

Map integration is an effective way to increase the marker numbers and the saturation of the map. An integrated map synthesizes the information from multiple segregating populations and various marker types of diverse genetic backgrounds, thus offering the opportunity to map larger number of markers than in most single crosses, and providing greater coverage of the genome that would be availability for molecular breeding purposes, as well as allowing comparison of locations of genes of interest across maps (Truco et al. 2007; Mace et al. 2009). Recently, integrated linkage maps have been constructed in many crop species involving lettuce (Truco et al. 2007), soybean (Hwang et al. 2009), cowpea (Muchero et al. 2009) and Sorghum (Mace et al. 2009) etc. In fact, cucumber map integration has also been reported (Bradeen et al. 2001), but the integrated maps are quite unsaturated (<200 loci). With the development of the cucumber genome sequencing projects, the ability to construct a high-density and highly representative cucumber "reference" genetic map has become possible.

In this study, we select two cucumber genetic maps of good quality, the S94 × S06 map (Yuan et al. 2008b) and the Gy14 × PI 183967 map (Ren et al. 2009) to construct a consensus map. The S94 × S06 map was constructed from larger segregating population of 224 recombinant inbred lines (RILs), while the Gy14 × PI 183967 map consisted of a great deal of sequence-based markers despite of a smaller population (77 RILs). In the study, we extend the previous S94 × S06 map using SSR markers developed from the cucumber genome sequencing project (CGSP) to generate potential anchor markers for map integration. Two individual linkage maps, the extended S94 × S06 map and the Gy14 × PI 183967 map, are then used to construct a consensus map, which will provide for a better coverage of all genomic regions. It is also more representative of *C. sativus* var. *hardwickii*, a feral form of *C. sativus* var. *sativus* and its inclusion allows for more universal utility of markers across different genetic backgrounds.

Materials and methods

Plant materials

Two mapping populations were used to develop an integrated map of cucumber (Table 1). The first one consisted of 77 F_6–F_8 recombinant inbred lines (RILs) derived from an inter-subspecific cross between GY14 and PI183967 (Ren et al. 2009). GY14 is a North American processing

Table 1 Summary of two mapping populations used to construct the consensus map of *Cucumis sativus* L

Mapping population	Cross types	Population type	No.of lines	No. of markers	Predominant markers type	No. of SSR markers	References
GY14 × PI 183967	Inter-subspecific cross	F_8RIL	77	995	SSR	995	Ren et al. (2009)
S94 × S06	Intra-subspecific cross	F_7RIL	224	257	SRAP	22	Yuan et al. (2008b)

RIL recombinant inbred lines, *SSR* simple sequence repeat, *SRAP* sequence-related amplified polymorphism

Table 2 SSR markers on the integrated consensus map

SSRs code	Source of markers	No.of loci	References
C_	Cucumber genomic library	2	Our laboratory
CM_	Melon ESTs	5	Kong et al. (2007)
CMBR_	Melon genomic library	6	Ritschel et al. (2004)
CMCT_	Melon genomic library	1	Chiba et al. (2003)
CMGA_	Melon genomic library	1	Danin-Poleg et al. (2001)
CMMS_	Melon genomic library	1	Danin-Poleg et al. (2001)
CMTC_	Melon genomic library	3	Danin-Poleg et al. (2001)
CS_	Cucumber ESTs	7	Kong et al. (2007)
CSAT_	Cucumber ESTs	2	Danin-Poleg et al. (2001)
CSTCC_	Cucumber cDNA library	1	Danin-Poleg et al. (2001)
CSEPGN_	Cucumber fruit ESTs	1	Our laboratory
CSFR_	Cucumber fruit ESTs	1	Our laboratory
CSJCT_	Cucumber genomic library	12	Watcharawongpaiboon and Chunwongse (2007)
CSWAC_	GeneTrapper kit to select plasmids harboring microsatellites in cucumber	1	Fazio et al. ((2002)
CSWCT_	GeneTrapper kit	10	Fazio et al. (2002)
CSWGAAT_	GeneTrapper kit	1	Fazio et al. (2002)
CSWGATT_	GeneTrapper kit	2	Fazio et al. (2002)
CSWTA_	GeneTrapper kit	3	Fazio et al. (2002)
CSWTAAA_	GeneTrapper kit	1	Fazio et al. (2002)
SSR_	Cucumber genome sequencing project	1,091	Ren et al. (2009)

EST expressed sequence tags

market type cucumber cultivar and PI183967 is an accession of *C. sativus* var. *hardwickii* originated from India. Another population derived from an intra-subspecific cross between a Northern China type line S94 and a Northern European type line S06 (*C. sativus* var. *sativus* × *C. sativus* var. *sativus*) consisted of 224 F_7 RILs (Yuan et al. 2008b). DNA extraction was performed as previously described (Yuan et al. 2008a).

SSR markers and segregation data

Several sources of markers, including SRAPs, SCARs, STS, SSRs and MTM, mapped in previous cucumber individual maps, were used to prepare the cucumber consensus map. Markers designated with the prefixes 'ME_', 'PM_' and 'e_' indicate SRAP markers, while a number and postfix 'STS' represent STS marker. 'S_', 'SCZ_', 'B_', 'F' and postfix 'SCAR' designate SCAR markers. Nomenclature and source of SSR markers loci integrated into the consensus map are listed in Table 2. All primer sequences for the SRAP markers and primer sequences of SCARs and STS are available in Supplement material S1 and Supplementary material S2, respectively. The segregation data of SCARs, SRAPs, SSRs (22 loci), STS and MTM were from the S94 × S06 map (SS map) (Yuan et al. 2008b). The segregation data of 995 SSR loci mapped in GY14 × PI 183967 map (GP map) were obtained from Ren et al. (2009), and of all primer sequences are listed in Supplementary material S3.

A total of 2,010 SSR markers, including 1,920 markers from the CGSP (Cucumber Genome-Sequencing Project) and 90 markers designated with the prefixes 'CM_', 'CMMS_' and 'CSJCT_'from previous reports (see Table 2), were test in the polymorphism analysis using the

S94 × S06 mapping population to generate bridging markers for the consensus map. The PCR for SSR was carried out on a 10 μl reaction volume with the following conditions: 94°C for 5 min, followed by 32 cycles at 94°C for 30 s, 50°C for 30 s, and 72°C for 30 s and a final extension at 72°C for 5 min. The amplification products were separated on 6% denatured polyacrylamide gels with 1× TBE buffer at a constant power of 50 W for 1.5 h. After electrophoresis, the gel was silver-stained (Bassam et al. 1991) and photographed with a digital camera (Olympus). The primers were synthesized by Sangon Biological Engineering Technology & Service Co. (Shanghai).

Construction of the individual linkage map

JoinMap program version 3 was used to analyze the segregation data from each mapping population (Van Ooijen and Voorrips 2001). The χ^2 test was used to assess goodness-of-fit to the expected 1:1 segregation ratio for each marker. The stable co-dominant SSR markers were used to construct the framework map in SS map. Markers were assigned to linkage groups (LGs) by the LOD score in the range from 3 to 10 for grouping. A maximum LOD score was selected to group by approaching the number of cucumber chromosomes. Linkage analysis was carried out using the following thresholds for JoinMap 3.0: REC smaller than 0.4, LOD larger than 6.0, RIPPLE of 1 and JUMP in goodness-of-fit of 5.0. The recombination percentage was converted to genetic distance by the Kosambi mapping function (Kosambi 1944). The other marker types including SRAP, SCAR, STS and MTM on previous cucumber maps (Yuan et al. 2008b) were added to the SSR framework map using the 'Fixed Order module' in JoinMap. Parameter was the same with that of the framework map expected for the LOD 3.0. Markers which had been excluded at higher LOD scores were reevaluated by decreasing the LOD threshold step by step until it reached the minimum LOD of 1.0. Markers that did not affect the order of the framework map and noticeably increased the distances of the map were included in the final extended SS map. The GP map was constructed using the same parameters described by Ren et al. (2009).

Construction of the consensus map

The two individual maps were integrated into a consensus map using JoinMap program version 3 by bridging markers. Bridging markers were identified as having an identical name and a similar map position in the different mapping populations. Markers with the same name that had inconsistent positions (>10 cM) in different populations were not considered as bridging markers and were not used for map integration in the paper. The position of bridging markers on each chromosome between the individual maps was graphically detected by Mapchart 2.2 software (Voorrips 2002). The homologous chromosomes between the individual maps were integrated by applying the 'combine groups for map integration' module with the following thresholds for Kosambi mapping function: REC smaller than 0.40, LOD larger than 3.0, RIPPLE of 1 and JUMP in goodness-of-fit of 5.0. The colinearity of an integrated map and the individual maps was graphically evaluated using Mapchart.

Map validation through the cucumber genome sequence

The draft genome sequence of *C. sativus* var. *sativus* L. covering 243.5 Mbp was used for the map validation. The sequences of the sequence-associated mapped markers were aligned along the cucumber assembled genome sequence to validate their order by the *Blastn* tool at the cucumber genome database (CGD) website (http://cucumber.genomics.org.cn/page/cucumber/blast.jsp).

Results

SSR marker analysis

In order to increase saturation of the SS map and to generate bridging markers for the integrated map, a total of 2,010 SSR primers, including 1,920 markers from the CGSP and 90 markers from the previous reports, were used to screen the polymorphism from the parents (S94 and S06) of the SS mapping population. Of 2,010 SSR primers, 1,567 (81.6%) and 89 (98.9%) were informative and could yield products in both parents or in only a single parent. Polymorphism between the parents was observed with 364 (364/1567 × 100% = 23.2%) and 17 (17/89 × 100% = 19.1%) of primers, respectively. In fact, the 364 SSR primers amplified 366 polymorphic bands. Each SSR primers (SSR00153 and SSR14392) amplified two polymorphic loci. A total of 383 (383/1657 × 100% = 23.1%) polymorphic marker loci were observed and used for mapping. Moreover, the polymorphism ratio (23.1%) obtained from the parents of the SS mapping population (intra-subspecific cross) in the study is approximately one-third of that (68.1%) from the parents (GY14 and PI 183967) of the GP mapping population (inter-subspecific cross, 1,322 polymorphisms among the 1,940 SSR markers used for analysis) according to Ren et al. (2009).

The individual maps

New individual map for the S94 × S06 population was constructed by integrating the old and new marker data. After removing 19 unlinked markers and 11 unsuccessfully

Table 3 Description of the two individual component maps and the integrated consensus map

Chromosome	S94 × S06 extension map			GY14 × PI 183967 map			The consensus map							
	Genome coverage (cM)	No. of loci (SSR)[a]	Maker density (cM)	Genome coverage (cM)	No. of loci (SSR)	Maker density (cM)	Bridging marker (SSR)	Genome coverage (cM)	No. of loci					Maker density (cM)
									SSR	SRAP	SCAR	STS	MTM	Total	
chr. 1	114.5	83 (50)	1.38	96.2	118 (118)	0.82	19	114.3	137	28	0	1	–	166	0.69
chr. 2	100.8	75 (42)	1.34	100.2	126 (126)	0.80	22	109.4	136	31	2	–	–	169	0.65
chr. 3	144.5	132 (82)	1.10	112.7	187 (187)	0.60	35	139.2	223	46	2	–	–	271	0.51
chr. 4	98.5	69 (46)	1.43	37.3	114 (114)	0.33	16	76.3	139	21	2	–	–	162	0.47
chr. 5	106.1	79 (52)	1.34	59.9	160 (160)	0.37	20	75.9	188	19	5	–	3	215	0.35
chr. 6	106.3	124 (75)	0.86	106.5	203 (203)	0.53	33	108.6	225	38	8	–	–	271	0.40
chr. 7	78.5	48 (37)	1.64	60.1	87 (87)	0.69	15	76.8	104	9	2	–	–	115	0.67
Total	749.2	610 (384)	1.2 (ave)	572.9	995 (995)	0.58 (ave)	160	700.5	1,152	192	21	1	3	1,369	0.51 (ave)

ave average marker interval, *mtm* morphological trait marker

[a] Number of total SSR markers on each chromosome

Table 4 Distribution of distorted markers cross the chromosomes in two mapping populations in the study

INT	Chr.	1	2	3	4	5	6	7
SS map	S/T	17/83	37/75	17/132	16/69	16/79	60/124	12/48
var.sativus × *var.sativus*	G	1	2	1	1	–	3	1
28.7% distorted markers	L	6.6	26.6	7.1	19.4	–	35.9	8.2
	D	a_s	b_s	a_s	a_s	–	a_s	a_s
GP map	S/T	42/118	15/126	25/187	68/114	3/160	49/203	0/87
C. sativus var. *sativus* × *C. sativus* var. *hardwickii*	G	1	1	1	1	–	1	–
20.3% distorted markers	L	41.4	3.4	13.7	34.4	–	18.3	–
	D	b_g	b_g	b_g	b_g	–	b_g	–

SS S94 × S06, *GP* GY14 × PI 183967, *S/T S* number of distorted markers, *T* total number of markers on chromosome, *G* a group of distorted markers defined as a group of four or more closely linked distorted markers with gaps of non-distorted markers of three or less, *L* interval in cM of the largest distorted group, *D* direction of the distortion, a_s towards the S94 allele, b_s towards the S06 allele, b_g toward the PI183967 allele

positioned markers, the extended S94 × S06 map including 610 markers (384 SSRs, 199 SRAPs, 23 SCARs, one STS and three morphologic markers) was constructed and consisted of seven linkage groups corresponding to the seven cucumber chromosomes (Table 3; Supplementary material S2). Of 384 mapped SSR loci, 346 are from CGSP and 271 of 346 loci are first mapped on the genetic map. The map covered 749.2 cM at an average marker density of 1.2 cM, and there was only a large genetic interval with a gap of >10 cM at the distal end of chromosome 4. The GY14 × PI 183967 map contained seven chromosomes with 995 SSR markers and covered 572.9 cM with the average marker interval of 0.6 cM (Table 3; Supplementary material S3). Both populations had some markers with segregation ratios that deviated from Mendelian expectation (expected 1:1 for RIL population). Distorted markers were widely distributed through the genome present in all the chromosomes in SS and GP populations, except for chr. 7 in GP map (Table 4; Supplementary material S2 and S3). The segregation distorted ratio of each chromosome varied, and the higher distorted chromosomes were chr. 2, chr. 4 and chr. 6 on SS map, and chr. 1, chr. 4 and chr. 6 on GP map (Table 4). Most of distorted markers were associated in groups on the individual maps, indicating that the distortion was due to selection favoring one parental allele rather than sampling error.

The consensus map

Common SSR markers between homologous chromosomes were used as bridges to integrate the individual maps into a single consensus map. A total of 175 markers (all SSR markers) at two individual maps were in common. Of the 175 markers, 15 (8.6%) were discarded for showing contradictory genetic positions (>10 cM) and 160 bridging markers were used for map integration (Table 3). The order of bridging marker between two individual maps was consistent, except for a few marker inversions at the end of chr. 2,

chr. 3, chr. 6 and chr. 7. JoinMap 3.0 was used to integrate two individual maps into a consensus map based on 160 bridging markers in seven chromosomes. The joint segregation analysis with a total of 1,605 loci and 301 RILs from the two populations resulted in an integrated map with 1,369 loci, including 1,152 SSR loci, 192 SRAP loci, 21 SCAR loci and one STS locus, as well as three gene loci of fruit external quality traits (Table 3; Fig. 1). Detailed information of all marker loci in consensus map and their features, including the genetic distance and physical position of each locus in chromosomes, source of each marker locus, marker types, bridging markers and multi-copy marker, are listed in Supplementary material S4. The consensus map spanned 700.5 cM, of which 682.7 cM (97.5%) were covered by SSR markers. The length of seven chromosomes ranged from 75.9 cM (chr. 5) to 139.2 cM (chr. 3) and 94.7% (chr. 5) to 100% (chr. 2 and chr. 4) were covered by SSR markers. The average genetic distance between loci was 0.51 cM and ranged from 0.35 cM (chr. 5) to 0.69 cM (chr. 1). This was an average physical interval of ~268 kbp per marker considering the cucumber genome size to be approximately 367 Mbp (Arumuganathan and Earle 1991). There was no large gap (>10 cM) detected on the integrated map (Fig. 1). Most marker loci were relatively evenly distributed along the chromosomes on the integrated map except for four marker cluster regions detected on chromosomes 4, 5 and 7 (Figs. 1, 2). Markers of three chromosome regions showed a tendency to cluster around centromeres, and markers of one chromosome region clustered on a distal end region of chromosome 4. The clustering markers mainly originated from that of the GP map.

Alignment of the maps

Compared to the individual maps, the marker order of the integrated map was generally consistent between homologous chromosomes. Well-ordered collinear markers were found on chromosomes 1, 2, 3 and 6 (Fig. 2a). There are a few markers inconsistent and the inconsistency was usually within a small interval (<5 cM on the integrated map). Large differences between the integrated map and the individual maps were detected on the marker cluster regions of chromosomes 4, 5 and 7 (Fig. 2a). Two SSR markers (SSR23549 and SSR22231) from the top end on chr. 4 and a 20 cM length fragment on the top end of chr. 5 on S94 × S06 map were separately integrated into the marker cluster regions on the corresponding chromosomes of the integrated map, making the marker order in these regions quite different between the integrated map and the individual maps. This may be due to wrong location with a weak linkage (>10 cM) or chromosome structure differences among parents of different mapping populations.

Map validation

The sequences of the sequence-associated markers, including 316 markers (206 SSRs, nine SCARs and one STS) in SS map, 892 SSRs in GP map and 995 markers (985 SSRs, SCARs and one STS) in consensus map, were used to align along the assembled cucumber genome sequence to validate their order by using the Blastn tool at CGD, respectively. The result showed that 80% of the sequence-associated markers in consensus map were ordered along the chromosomes of the assembled cucumber genome sequence, except for some markers mapped in smaller regions (<5 cM) at the top end of chromosomes 2, 3 and 6 and larger chromosome regions (>20 cM) at the top end of chromosomes. 4, 5 and 7 (Fig. 2b; Supplementary material S4). Differences of markers order existed in individual maps explain that of the consensus map (Fig. 2b; Supplementary materials S2, S3). In addition, the consensus map spanned 700.5 cM in this study and covered 70–93.4% of the cucumber genome according to the estimated size of the entire cucumber genome of 750–1,000 cM (Staub and Meglic 1993). This observation was in good agreement with the physical covering size based upon these sequence-associated markers, which covered about 172.5 Mbp (70.8%) out of 243.5 Mbp of the assembled genome sequence (Huang et al. 2009). The sequences of the sequence-associated markers were listed in Supplementary material S5.

Discussion

Map integration is an effective way to increase marker numbers and map saturation. Recently, abundant sequence-associated markers (SSRs) have been developed from the cucumber genome-sequencing project, which provides the opportunity for integrating cucumber genetic maps constructed by different laboratories. In the study, we extended the previous S94 × S06 map (Yuan et al. 2008b) using SSR markers developed from the CGSP to generate potential anchor markers for map integration. Two individual linkage maps, the extended S94 × S06 map and the Gy14 × PI 183967 map, were integrated into a consensus map. The stable and saturated map serves as a reference map for cucumber genetic analysis involving map-based gene isolation, quantitative trait loci (QTL) fine analysis, construction of physical maps and marker assisted selection (MAS), even for comparative analysis in other cucurbit genomes.

The individual map

Previously, a genetic map consisting of 257 markers was developed by a RIL population (224 individual lines) from

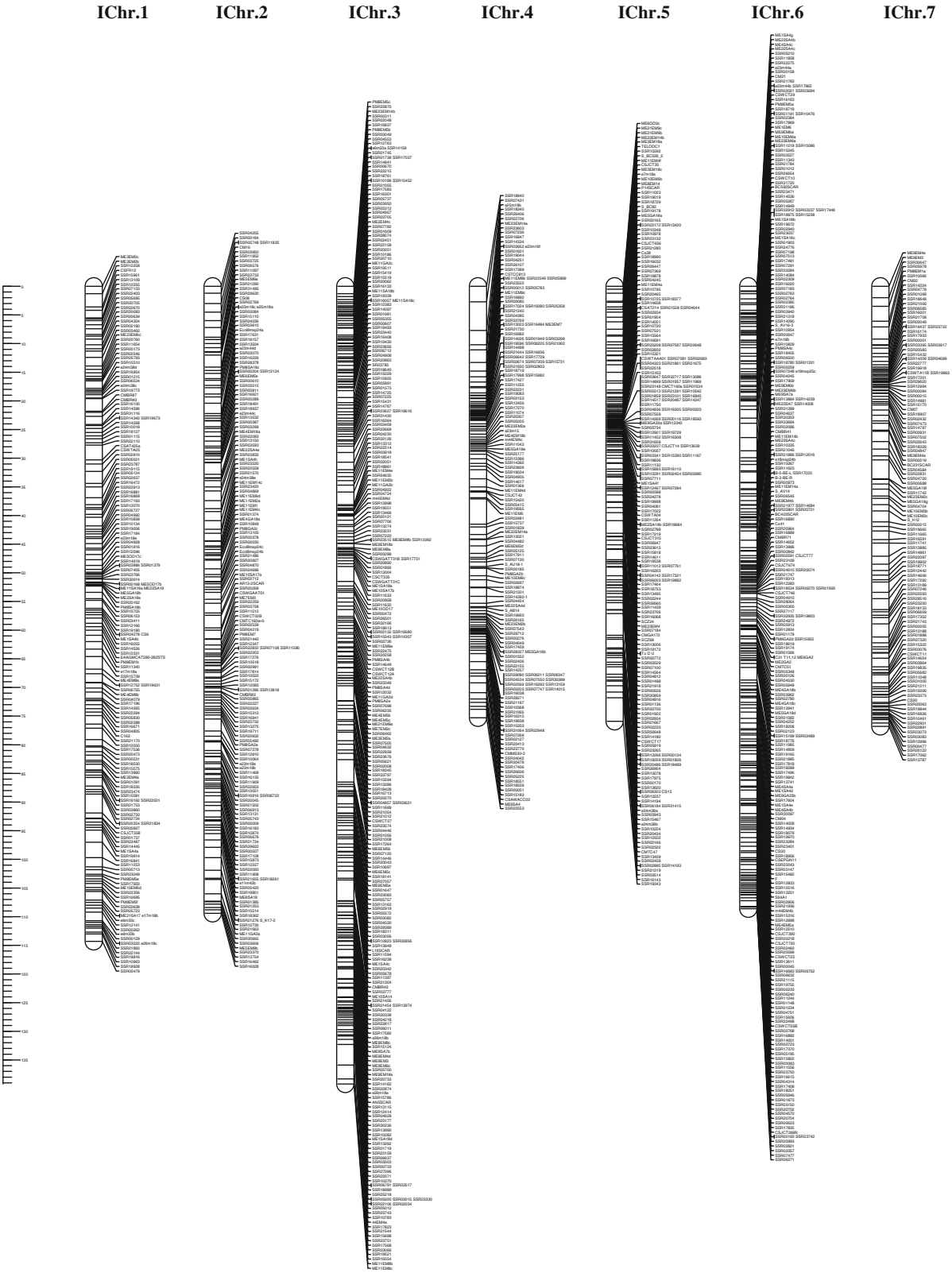

Fig. 1 The consensus map of cucumber from the two independent genetic maps. *Ruler* on the right side shows the distance in centiMorgan from the top of each chromosome. Detailed information of each marker loci in chromosome, including the genetic and physical position of each locus in chromosomes, source of each marker locus, marker types, bridging markers and multi-copy marker are available in Supplementary material S4

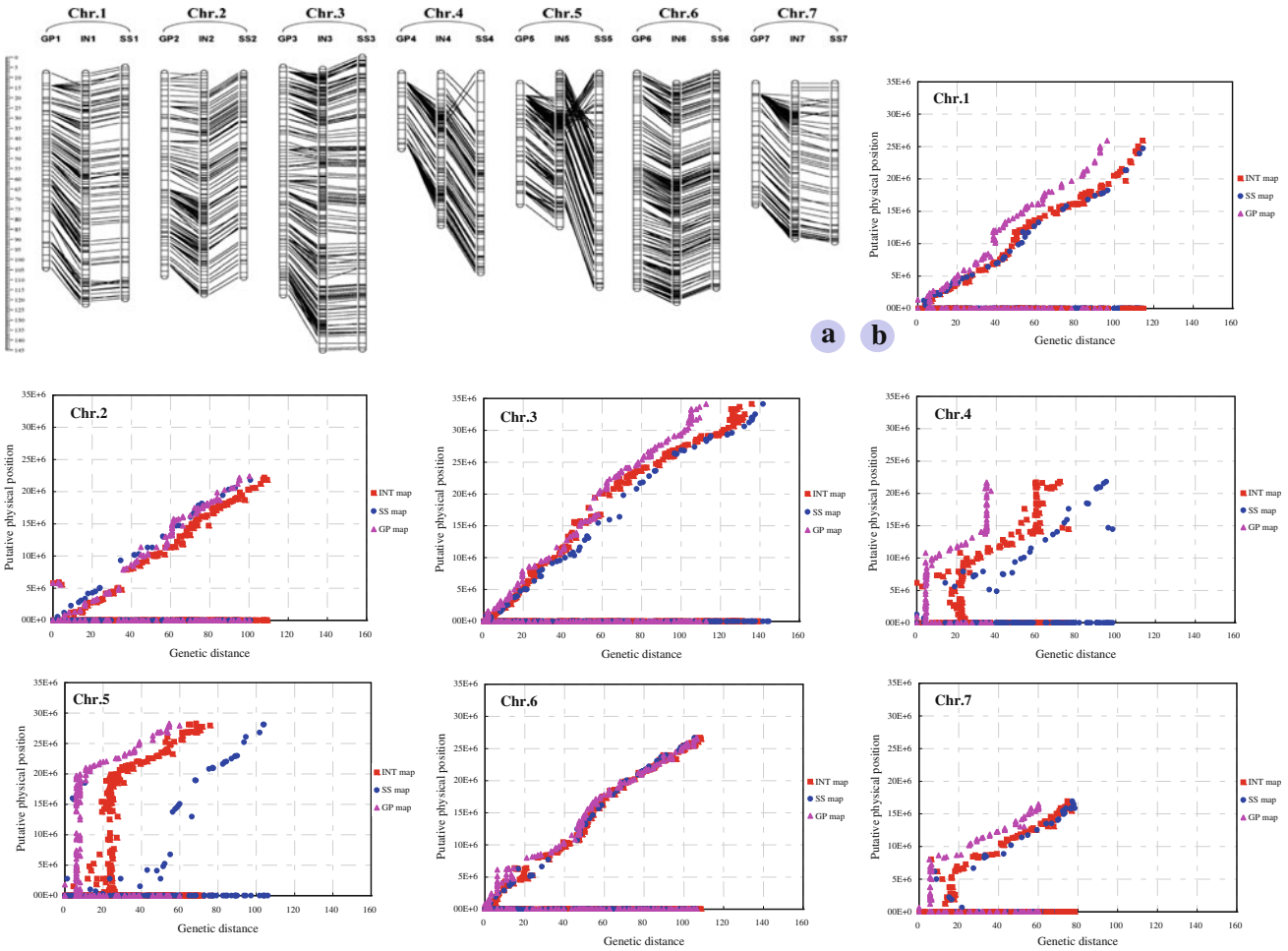

Fig. 2 **a** Collinear analysis of the consensus map and two individual maps for seven cucumber chromosomes. *GP* GY14 × PI 183967 map, *IN* integrated consensus map, *SS* S94 × S06 map, *Line* indicates the same markers between the integrated consensus map and GY14 × PI 183967 map or S94 × S06 map. *Black region* on chr. 4, chr. 5 and chr. 7 of the consensus map indicates the markers cluster regions. **b** Genetic versus physical mapping of marker on the 7 cucumber chromosomes of the two individual maps and the consensus map. *Purple triangle*, *blue circle* and *red box* on the vertical and horizontal axis indicates the genetic mapping of markers (cM) and the physical position of markers (Mb) on the GP map, SS map and INT map, respectively (color figure online)

a cross of S94 × S06 (Yuan et al. 2008b). In the present study, using the same RIL population, a S94 × S06 map were extended by increasing the number of SSR markers from CGSP. The extended map contained 610 marker loci (384 SSRs, 199 SRAPs, 23 SCARs, one STS and three morphological traits) and covered 749.2 cM with an average marker interval of 1.2 cM (Table 3; Supplementary material S2). As compared to the original S94 × S06 map, the new map contained more DNA markers (607). In particular, the sequence-specific markers (SSR and SCAR) increased from the original 48 to 408, and the number of gaps (>10 cM) decreased from the original 27 to only one gap, making the new S94 × S06 map constructed in the study the most saturated linkage map among the published genetic maps from intra-subspecific crosses (narrow cross type) to date.

The segregation distortion ratio of each chromosome varied widely, and the most segregation distortion regions (SDRs) differed in each individual map, indicating that segregation distortion can occur anywhere in the cucumber genome (Supplementary material S2, S3). All marker loci within these SDRs in SS map were associated with *C. sativus* var. *sativus* parent S94 except for markers on chr. 2 with *C. sativus* var. *sativus* parent S06, while all markers loci in GP map were associated with *C. sativus* var. *hardwickii* parent PI 183967 (Table 4). *C. sativus* var. *sativus* parent S94 shows most phenotypic characteristics of wild cucumber and *C. sativus* var. *hardwickii* parent PI183967 is a wild form of cucumber, indicating that the wild alleles confer stronger viability than the domesticated ones.

The consensus map

In the study, bridging markers between homologous chromosomes were used to integrate the individual maps (the extended SS map and the GP map) into a single consensus map by using Joinmap3.0 software. The consensus map contained 1,369 loci, including 1,152 SSR loci, 192 SRAP loci, 21 SCAR loci and one STS locus as well as three gene loci of fruit external quality traits, covered densely all seven cucumber chromosomes and spanned 700.5 cM at the average marker density of 0.51 cM. The consensus map with high marker density is the most saturated genetic map in cucumber or even in the Cucurbitaceae family to date. Moreover, it is also more representative of *C. sativus* var. *hardwickii*, a feral form of *C. sativus* var. *sativus* and its inclusion allows for more universal utility of markers across different genetic backgrounds.

One of the challenges of map integration was to deal with multi-loci markers. Even for the same primers, alleles mapped in one population could be the same or different from those mapped in other populations since multiple alleles or loci can be detected and mapped from some SSR markers (Gustafson Perry et al. 2009). In the study, two SSR markers SSR00153 and SSR14392 behaved as multi-loci markers in SS extension map and each primer amplified two polymorphic loci. For SSR00153, two polymorphic loci mapped to chr. 1 and chr. 4, while two SSR14392 loci both mapped to chr. 4 at the interval of more than 20 cM (Supplementary material S4). Furthermore, visual observation of the common SSR markers between the SS extended map and the GP map by Mapchart software found that the polymorphic loci were amplified by four same SSR primers located on the different chromosomes in different mapping populations, for example, SSR05830 on chr. 1, SSR20859 on chr. 5, as well as SSR04530 and SSR04245 on chr. 6 in SS extended map were mapped on chr. 3, chr. 6, as well as chr. 3 and chr. 5 in GP map, respectively (Supplementary material S4). The five SSR markers may be potential multi-loci markers, or may be mapped inaccurately by small mapping populations and missing or poor quality data. Due to the existence of multi-loci markers, common markers showing inconsistent genetic mapping (>10 cM) and significant heterogeneity (i.e., potentially different loci) between two mapping populations were eliminated to ensure the accuracy of the consensus map. Of 175 common markers between the two individual maps in the study, 15 (8.6%) were discarded.

The consensus map constructed in the study is a more informative genetic resource. It consists of 1,369 loci and spans 700.5 cM with the average interval of 0.51 cM (Fig. 1). The consensus map has a higher marker density and smaller average marker interval than the SS map (749.2 cM, 1.2 cM) and GP map (572.9 cM, 0.6 cM). However, the total length of the consensus map is 48.7 cM shorter than that of SS map, and 127.6 cM larger than that of GP map (Table 3). This mainly is the reason that there were four marker clustering regions detected on chr. 4 (two cluster regions), chr. 5 (one cluster region), and chr. 7 (one cluster region) of the GP map (Fig. 2a), and the genetic intervals among markers in GP map were enlarged on the consensus map, similar to results obtained after the integration of markers in a reported by Gustafson Perry et al. (2009). Compared to the SS map (749.2 cM), the total length of the consensus map (700.5 cM) was shortened, while actual genome coverage was increased. An extra 9 cM at the top end of chr. 2 and 10 cM at the bottom of chr. 5 in consensus map lacked in SS map, and 4–6 cM at the top end of chr. 1 and chr. 7 as well as 12–13 cM at the top end of chr. 4 and chr. 5 lacked in GP map, indicating that the consensus map had larger coverage of the whole cucumber genome than the two individual maps. Furthermore, analysis of physical position of the sequence-associated markers showed that the consensus covered about 172.5 Mbp (70.8%) out of the 243.5 Mbp of assembled genome sequence (Huang et al. 2009). Some regions in consensus map, such as the top end of chr. 5 and chr. 6, contained many SRAP markers which had no information of the corresponding sequence (Supplementary material S4), thus the genome coverage of the consensus map was underestimated.

Recombination suppression caused by chromosomal rearrangement may be the cause of the marker clustering (Rieseberg and Livingstone 2003). In the study, these recombination suppression regions were detected by common markers between the GP map and SS map (Supplementary material S2, S3). For example, the cluster on chr. 5 of the GP map spanned 1.9 cM, but it increased dramatically to 60.0 cM on the SS map. A chromosome inversion between GY14 and PI 183967 by FISH analysis with SSR markers from one cluster region on the top end of chr. 5 was observed (Ren et al. 2009). Therefore, The chromosomal structure rearrangement between chromosomes of GY14 and PI 183967 resulted in suppression of meiotic recombination of chr. 5. Physical position of the sequence-asssociated markers on chr. 5 in GP map showed that markers in recombination suppression region randomly arranged at the cucumber genomic sequence (Fig. 2b: chr. 5, Supplementary material S3), while the same result on the top end of chr. 5 in SS map was observed. Physical positions of all markers were varied from 0–20Mbp on chr. 5 in individuals map and the consensus map (Fig. 2b: chr. 5, Supplementary material S2). The SS map was from the intra-subspecific cross and both parents belong to *C. sativus* var. *sativus*, so physical positions of the sequence-associated should theoretically align better along the cucumber genomic sequence from *C. sativus* var. *sativus* inbred line 9,930. However,

random arrangement at 50 cM intervals was observed, indicating that different cucumber parents may have the chromosomal structure differences on chr. 5. Similarly, the two clusters on chr. 4 of the GP map had covered almost two-thirds of the total length of chr. 4 of the SS map (Fig. 2b: chr. 4, Supplementary material S2, S3), which may be explained by chromosomal structure rearrangement. The differences in locus order on the genetic map caused by chromosomal rearrangement have been reported in some plants (Hayashi et al. 2001; Isobe et al. 2009). These regions of recombination suppression are useful for studying cucumber evolution during domestication. The marker cluster of chr. 7 in GP map may be caused by less recombination events around centromeres (Huang et al. 2009). Random arrangement markers were varied under 0–8 Mbp on chr. 7 in individual maps and the consensus map (Fig. 2b: chr. 7, Supplementary material S2). In addition, putative errors of the assembly of the 9,930 genomic sequences would also account for disagreement between genetic and physical linkages observed in the study. Well-ordered collinear markers were detected on chromosomes 1, 2, 3 and 6 of the consensus map (Fig. 2), where the order of only a few markers were inconsistent within a small interval (<5 cM on the integrated map), indicating that sequence structure on most chromosomes are highly conserved in cucumber. Local inconsistencies in marker order may be partially attributed to a sampling bias mainly in the small mapping population (GY14 × PI 183967, 77 RIL individuals). Well ordered markers were highlighted in consensus map (Supplementary material S4) and regions of likely structural rearrangement were identified and served as a resource to map in additional population of better genetic resolution.

Accessibility of the consensus map

Previous research showed that if the genetic map was constructed for the purpose of map-based gene cloning in cucumber, the average interval between the markers would be less than 1 cM. Actually, the average marker interval on all seven chromosomes of the consensus map was less than 0.7 cM (Table 3). Therefore, this consensus map should be used for map-cloning genes of key agronomic traits in cucumber. In the study, three valuable external quality traits related to the market values of cucumber, tuberculate fruit (*Tu*), dull skin (*D*) and uniform immature fruit color (*u*), were mapped between SCZ69 (2.5 cM) and SSR04142 (4.4 cM) on chr. 5 of the S94 × S06 extended map (Supplementary material S2). Two new SSR markers in consensus map, SSR19172 and SSR00772, narrowed the flanking markers interval to 0.2 cM and 0.9 cM, respectively (Fig. 1; Supplementary material S4). Obviously, the saturated and high-density consensus map will further promote identification and map-based cloning of the three quality trait genes.

The consensus map contains 1,369 loci, of which 1,174 (86%) are special markers that can be amplified by low cost and non-labor intensive assays. Thus, these markers will greatly enhance the efficiency of cucumber breeding programs via marker-assisted selection (MAS). Furthermore, the physical location of the 995 sequence-associated markers on the assembled cucumber genome sequence established the relationship between the genetic map and the whole genome sequence of cucumber, which greatly validated the order of the markers on the consensus map and will provide valuable information for genomic, structural and evolutionary studies in cucumber and other species of Cucurbitaceae.

Acknowledgments This work was supported by grants from the National Natural Science Foundation of China (No. 30671111, 30871707), the Shanghai Leading Academic Discipline Project (No. B209), the Chinese Ministry of Agriculture (948 Program: 2008-Z42), the Shanghai Municipal Scientific and Technological Commission Project (No. 09JC1407300) and the National Basic Research Program of China (973 Program, Grant No. 2012CB113902).

References

Arumuganathan K, Earle E (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208

Bassam BJ, Caetana-Anolles G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 196:80–83

Bradeen JM, Staub JE, Wye C, Antonise R, Peleman J (2001) Towards an expanded and integrated linkage map of cucumber (*Cucumis sativus* L.). Genome 44:111–119

Chiba N, Suwabe K, Nunome T, Hirai M (2003) Development of microsatellite markers in melon (*Cucumis melo* L.) and their application to major cucurbit crops. Breed Sci 53:21–27

Danin-Poleg Y, Reis N, Tzuri G, Katzir N (2001) Development and characterisation of microsatellite markers in *Cucumis*. Theor Appl Genet 102:61–72

Fazio G, Staub JE, Chung SM (2002) Development and characterization of PCR markers in cucumber. J Amer Soc Hort Sci 127(4):545–557

Fazio G, Staub JE, Stevens MR (2003) Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Theor Appl Genet 107:864–874

Gustafson JP, Ma XF, Korzum V, Snape JW (2009) A consensus map of rye integrated mapping data from five mapping populations. Theor Appl Genet 118:793–800

Hayashi M, Miyahara A, Sato S et al (2001) Construction of a genetic linkage map of the model legume Lotus japonicus using an intraspecific F2 population. DNA Res 8:301–310

Huang SW, Li R, Zhang Z et al (2009) The genome of the cucumber, *Cucumis sativus* L. Nat Genet 41:1275–1281

Hwang TY, Sayama T, Takahashi M et al (2009) High-density integrated linkage map based on SSR markers in soybean. DNA Res 16:213–225

Isobe S, Kolliker R, Hisano H, Sasamoto S, Wada T, Klimenko I, Okumura K, Tabata S (2009) Construction of a consensus linkage map for red clover (*Trifolium pratense* L.). BMC Plant Biol 9:57

Jiang S, Yuan XJ, Pan JS, He HL, Cai R (2008) Quantitative trait locus analysis of lateral branch-related traits in cucumber (*Cucumis sat-*

ivus L.) using recombinant inbred lines. Sci China C Life Sci 51:833–841

Kennard WK, Poetter K, Dijkhuizen A et al (1994) Linkages among RFLP, RAPD, isozyme, disease-resistance, and morphological markers in narrow and wide crosses of cucumber. Theor Appl Genet 89:42–48

Kong Q, Xiang C, Yu Z, Zhang C, Liu F, Peng C, Peng X (2007) Mining and charactering microsatellites in *Cucumis melo* expressed sequence tags from sequence database. Mol Ecol Notes 7:281–283

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Li XZ, Pan JS, Wang G, Tian LB, Si LT, Wu AZ, Cai R (2005) Localization of genes for lateral branch and female sex expression and construction of a molecular linkage map in cucumber (*Cucumis sativus* L.) with RAPD markers. Prog Nat Sci 15:143–148

Li Z, Huang SW, Liu SQ et al (2009) Molecular isolation of the M gene suggests that a conserved-residue conversion induces the formation of bisexual flowers in cucumber plants. Genetics 182:1381–1385

Mace ES, Rami JF, Bouchet S et al (2009) A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. BMC Plant Biology 9:1–13

Muchero W, Diop NN, Bhat PR et al (2009) A consensus genetic map of cowpea [*Vigna unguiculata* (L) Walp.] and synteny based on EST-derived SNPs. PNAS 106:18159–18164

Park YH, Sensoy S, Wye C, Antonise R, Peleman J, Havey MJ (2000) A genetic map of cucumber composed of RAPDs, RFLPs, AFLPs, and loci conditioning resistance to papaya ringspot and zucchini yellow mosaic viruses. Genome 43:1003–1010

Ren Y, Zhang ZH, Liu JH, Staub JE, Han YH et al (2009) An integrated genetic and cytogenentic map of the cucumber genome. PloS ONE 4:e5795

Rieseberg LH, Livingstone K (2003) Evolution. Chromosomal speciation in primates. Science 300:267–268

Ritschel PS, deLins TCL, Tristan RL, Buso GSC, Buso JA, Ferreira ME (2004) Development of microsatellite markers from an enriched genomic library for genetic analysis of melon (*Cucumis melo* L.). BMC Plant Biol 4:9

Serquen FC, Bacher J, Staub JE (1997) Mapping and QTL analysis of horticultural traits in a narrow cross in cucumber (*Cucumis sativus* L.) using random-amplified polymorphic DNA markers. Mol Breed 3:257–268

Staub JE, Meglic V (1993) Molecular genetic markers and their legal relevance for cultivar discrimination, a case study in cucumber. Hort Technol 3:291–300

Sun Z, Staub JE, Chung SM, Lower RL (2006) Identification and comparative analysis of quantitative trait loci associated with parthenocarpy in processing cucumber. Plant Breed 125:281–287

Truco JM, Antonise R, Lavelle D et al (2007) A high-density integrated genetic linkage map of lettuce (Lactuca spp.). Theor Appl Genet 115:735–746

Van Ooijen JW, Voorrips RE (2001) JoinMap 3.0: software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Voorrips RE (2002) MapChart, software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Wang G, Pan JS, Li XZ, He HL, Wu AZ, Cai R (2005) Construction of a cucumber genetic linkage map with SRAP markers and location of the genes for lateral branch traits. Sci China C 48:213–220

Watcharawongpaiboon N, Chunwongse J (2007) Development and Characterization of Microsatellite Markers from an Enriched Genomic Library of Cucumber (*Cucumis sativus*). Plant breed 127:74–81

Xie J, Wehner CT. (2001) Gene list 2001 for cucumber. Cucurbit Genetics Cooperative Report

Yuan XJ, Li XZ, Pan JS et al (2008a) Genetic linkage map construction and location of QTLs for fruit-related traits in cucumber. Plant Breeding 127(2):180–188

Yuan XJ, Pan JS, Cai R et al (2008b) Genetic mapping and QTL analysis of fruit and flower related traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Euphytica 164:473–491

Zhang WW, He HL, Guan Y, Du H, Yuan LH, Li Z, Yao DQ, Pan JS, Cai R (2010) Identification and mapping of molecular markers linked to the tuberculate fruit gene in the cucumber (*Cucumis sativus* L.). Theor Appl Genet 120(3):645–654