Molecular markers assist in the development of diverse inbred backcross lines in European Long cucumber (*Cucumis sativus* L.)

Isabelle Y. Delannay · Jack E. Staub

Abstract The popular fresh-market European Long cucumber (*Cucumis sativus* L.) is grown commercially worldwide under controlled, greenhouse environments. However, it has a narrow genetic base, where private and public improvement programs can trace their origins to comparatively few accessions. Therefore, a project was designed to identify diverse genotypes for use in the formation and analysis of inbred backcross (BC_2S_3) lines (IBL) to broaden the genetic base of this market class. Initially, 42 cucumber accessions were evaluated with a previously defined standard marker array to identify parents for use in backcrossing. The IBL were developed by crossing the elite commercial line NZ1 (Nunhems Vegetable Seeds, Haelen, The Netherlands) and PI 432858 (China), and then backcrossing the most genetically diverse BC_1 and BC_2 progeny to the elite parent as defined by marker analyses (19 polymorphic, mapped SSR, and SCAR marker loci), followed by three generations of single seed descent resulting in 116 IBL (BC_2S_3). The IBL were evaluated under greenhouse conditions for days to anthesis, sex expression, lateral branch number, yield, and exterior fruit quality in Madison, Wisconsin, USA (soil media), and in Haelen and Bergschenhoek, The Netherlands (soilless, hydroponic media). The IBL were genotyped using an expanded marker array (37 polymorphic SSR, SCAR, SNP, EST, BAC end, and gene-associated loci), and genetic relationships were examined by multivariate analyses using phenotypic and genotypic data. The 116 developed IBL possessed considerable morphological and genotypic diversity, where genetic distance (GD) among lines ranged between 0.00 and 0.77. These IBL possessed many commercially acceptable attributes, and, thus, genetic diversity in this market type could be substantially increased by the use of these genetically broad-based IBL during plant improvement.

Keywords Genetic diversity · Genetic distance · Morphological traits · Multivariate analysis

Present Address:
I. Y. Delannay (✉) · J. E. Staub
Department of Horticulture, Vegetable Crops Unit,
U.S. Department of Agriculture, Agricultural Research Service, University of Wisconsin, 1575 Linden Dr, Madison, WI 53706, USA
e-mail: idelannay@yahoo.com

Present Address:
J. E. Staub
USDA, ARS, Forage and Range Research Laboratory,
Utah State University, Logan, UT 84322-6300, USA

Introduction

The European Long market type cucumber (*Cucumis sativus* L.; $2n = 2x = 14$) produces a long (32–40 cm in length), smooth, dark green, fine-spined, seedless (parthenocarpic, non-pollinated fruit set) fruit (Staub et al. 2008). Although the cultivation of this market type is historic, the last intensive, long-term public

breeding effort ended in the 1950s (Andeweg 1956). Since then, plant improvement of this market type has been conducted almost exclusively by private seed companies, primarily in The Netherlands. Current research on European Long cucumbers in public institutions focuses mainly on management practices to maximize productivity within protected growing environments (primarily glasshouse and plastic "hoop houses"; Ehret et al. 2001; Lin and Jolliffe 1994; Papadopoulos and Hao 2000; Van der Gaag and Wever 2005).

A majority of the commercial European Long cultivars are of Dutch origin, and are grown hydroponically in rock wool under highly controlled environmental conditions (Staub et al. 2008). Plants are gynoecious and develop multiple lateral branches that are pruned continuously to form an aerial leaf mat attached to lattice structures. Hydroponic plant culture in protected environments allows for year-round production of high quality fruit. Traits of particular commercial interest include gynoecy, disease resistance, sequential set of parthenocarpic fruit (i.e., fruit set not inhibited by other developing fruit), and the ability to grow in relatively low temperatures and light intensities.

The genetic diversity in commercial cucumber and accessions of exotic origin [i.e., plant introductions (PI)] has been well documented, and the crop's genetic base is considered extremely narrow {3–8% polymorphisms among elite and exotic germplasm and 12% between botanical varieties [*C. sativus* var. *sativus* L. and var. *hardwickii* (R.) Alef.]} (Dijkhuizen et al. 1996; Horejsi and Staub 1999; Meglic and Staub 1996; Meglic et al. 1996; Miliki et al. 2003; Staub et al. 1997 and Staub et al. 1999). In fact, the European Long market class has been found to have the narrowest genetic diversity among the major commercial cucumber market classes, which is in large part due to the initial use of relatively few PIs in germplasm development and the extensive use of the germplasm "Corona" in modern (>1950) breeding (Dijkhuizen et al. 1996; Horejsi and Staub 1999; Staub et al. 2008). Traditionally, commercial cucumber cultivars have been developed through the identification of a unique trait of interest (e.g., disease resistance) in a donor parent (usually a PI), followed by backcross introgression of this trait into an elite genetic background (i.e., commercially acceptable germplasm) (Staub et al. 2008). Inbred

backcross breeding has been used successfully in several crop species including cucumber for the introgression of quantitatively inherited traits during line development in preparation for hybrid production or population development (Owens 1985; Tanksley et al. 1996; Wehrhahn and Allard 1965).

Many yield- and quality-related component traits in cucumber (e.g., fruit number, weight, and shape, and parthenocarpy) are quantitatively inherited and often difficult to manage during plant improvement (Fazio et al. 2003; Sun et al. 2006). Because of the worldwide commercial importance of this market class and its inherent low genetic diversity, a project was designed to enhance and diversify its genetic base by constructing a genetically diverse array of IBL (BC$_2$S$_3$) through marker-assisted selection (MAS). To define the potential utility of these IBL, their phenotypic and genotypic diversity was assessed and the relative breeding value of commercially important traits was estimated. The development of a genetically diverse set of European Long market type IBL with differing morphological characteristics will provide cucumber breeders access to unique germplasm and associated markers for the strategic development of this market type and opportunities for the characterization quantitative trait loci (QTL) for use in MAS (Robbins et al. 2008).

Materials and methods

Identification of parental lines and IBL development

In the fall of 2004, potential parental lines were screened for their genotypic (marker-based assessment using 19 mapped loci) diversity as described by Delannay (2009) and Delannay and Staub (2010). Based on the genetic relationships and phenotypic characteristics (vegetative and reproductive), parents [recurrent (NZ1) and donor (PI 432858); Rogers (1972) genetic distance = GD = 0.81] were chosen and crossed to produce F$_1$ progeny (Fig. 1). The highly inbred gynoecious, lateral branching (2–4/plant in the first ten nodes on the main stem under Wisconsin greenhouse growing conditions) European Long market type line NZ1 was received from Nunhems Vegetable Seeds (Haelen, The Netherlands). It produces relatively long (28–33 cm), moderately dark green, spineless, ribbed, sequentially

setting parthenocarpic fruit (Fig. 1). In contrast, the monoecious landrace PI 432858 originates from China [obtained by the U.S. National Plant Germplasm System (NPGS) in 1979] and produces extremely long (37–39 cm), spiny, dark-green fruits, and was received from the USDA, ARS Regional Plant Introduction Station, Ames, IA, USA. It develops comparatively few lateral branches (0–2/plant) and does not set fruit sequentially under Wisconsin conditions (Fig. 1).

Cuttings of NZ1 were rooted and plants were greenhouse-grown to reproductive maturity in Madison, Wisc to be used as the recurrent paternal parent during backcrossing. A cloned NZ1 plant was mated to a NZ1 × PI 432858-derived F_1 individual to generate the BC_1 generation. The expanding leaves of 288 BC_1 seedlings were harvested at the first-leaf stage, and DNA was extracted according to methods described by Fazio et al. (2003).

Thirty BC_1 individuals were selected (selection intensity = 11%) for the development of the BC_2 generation based on their heterozygosity at 19 polymorphic, mapped SSR and SCAR marker loci (Table 1). These BC_1 individuals were crossed to the cloned recurrent parent (NZ1) to produce BC_2 progeny. Approximately 13 seeds from each of the 30 BC_2 families (384 seeds) were planted, DNA-sampled at the first true-leaf stage, marker-genotyped, and greenhouse-grown for self-pollination. One-hundred twenty BC_2 individuals with the greatest heterozygosity at 44 polymorphic SSR, SCAR, SNP, EST, BAC end, and gene-associated marker loci (Table 1) were self-pollinated to produce 116 BC_2S_3 IBL by single-seed descent.

Wisconsin phenotypic evaluations

During the summer (May–July) of 2007 and spring (April–June) of 2008, greenhouse evaluations were performed at the University of Wisconsin in Madison (UWM) in a randomized complete block design (RCBD) consisting of three blocks (one block per greenhouse) per year. Each of the three experimental blocks consisted of parents (NZ1 and PI 432858), F_2 progeny, and 38 randomly selected IBL, each with three plants per block. Seeds were germinated in vermiculite and seedlings were transplanted to 10 l plastic pots containing METRO-MIX 366P series growing medium (Sun Gro Horticulture Canada CM Ltd., Vancouver, British Columbia, Canada) on 1 m centers 2 weeks after germination. Fertilization was by "Osmocote" granular (Scotts Miracle Grow Co., Marysville, Ohio, 15-9-12, N:P:K) application (∼15 gm) at transplant and then monthly thereafter (∼15 gm), where standard greenhouse cultural practices were applied to plants grown under a 14 h photoperiod supplied by high-pressure sodium lighting [1000 watts; ∼500 µmols m^{-2} s^{-1} at mid-plant height (1.82 m)], at 35°C day and 25°C night temperatures, and ∼65% RH.

Fig. 1 Greenhouse-grown fruit of parental cucumber (*Cucumis sativus* L.) lines NZ1 (European Long market type) and PI 432858, and their F_1 progeny as harvested in Madison, Wisconsin

Table 1 Polymorphic markers and their expected fits to genotypic ratios in BC_2S_3 progeny that were used during selection of BC_1 and BC_2 progeny and subsequent genotyping of derived BC_2S_3 lines in cucumber (*Cucumis sativus* L.)

Marker	Marker type	Polymorphism type	Linkage group	Selection in BC_1	Selection in BC_2	Used in final evaluation	Expected allele frequency	Observed allele frequency	G-test
A08	SCAR	Codominant		Yes	Yes	Yes	0.86	0.75	9.68**
AD14SNPG1H2	SNP	Codominant		No	Yes	Yes	0.86	0.86	8.66*
AJ1-BE-L	BAC end	Codominant		No	No	No	0.86	0.13	545.82**
AJ-3-BE-L	BAC end	Dominant		Yes	No	No			
AK16 SCAR	SCAR	Codominant	F2	No	No	No	0.86	0.51	79.48**
AK16SNPG1H2	SNP	Codominant	F2	No	Yes	Yes	0.86	0.71	17.48**
AO12	SCAR	Codominant	F5	No	No	Yes	0.86	0.87	2.50 n.s.
AT1SNPG3H3	SNP	Codominant		No	Yes	Yes	0.86	0.82	7.82*
AW14	SCAR	Codominant	F3	Yes	Yes	Yes	0.86	0.68	25.39**
B-3-BE-R	BAC end	Codominant	Y2	Yes	Yes	Yes	0.86	0.99	27.38**
B-5-BE-L	BAC end	Dominant	Y2	No	Yes	No			
BC231	SCAR	Dominant	F7,Y7	Yes	No	No			
BC523	SCAR	Codominant	F1,Y2	Yes	No	Yes	0.86	0.75	20.27**
C4LG1H2	SNP	Codominant		No	No	Yes	0.86	0.83	5.92 n.s.
CM07	EST-SSR	Codominant		No	No	Yes	0.86	0.85	3.07 n.s.
CM23	EST-SSR	Dominant		No	No	No			
CM26	EST-SSR	Codominant		No	No	No	0.86	0.14	264.78**
CM49	EST-SSR	Codominant		No	No	Yes	0.86	0.85	0.12 n.s.
CS08	EST-SSR	Codominant	Y3	No	No	Yes	0.86	0.82	2.02 n.s.
CS23	EST-SSR	Codominant		No	No	Yes	0.86	0.84	0.46 n.s.
CS41	EST-SSR	Codominant		No	No	Yes	0.86	0.81	1.89 n.s.
CS-PS	GENE	Codominant		No	No	Yes	0.86	0.78	5.17 n.s.
CSWACC02	SSR	Codominant		Yes	No	No	0.86	0.57	57.15**
CSWATT02	SSR	Codominant		No	Yes	Yes	0.86	0.68	30.07**
CSWCT05B	SSR	Codominant		Yes	No	Yes	0.86	0.76	8.23*
CSWCT11	SSR	Codominant	Y7	Yes	No	Yes	0.86	0.97	21.12**
CSWCT24B	SSR	Codominant		Yes	No	No	0.86	0.64	38.53**
CSWCT25	SSR	Codominant	F1, Y2	Yes	No	Yes	0.86	0.92	4.34 n.s.
CSWGAAA02	SSR	Codominant		Yes	Yes	Yes	0.86	0.76	7.94*
CSWGCA01	SSR	Codominant	F1	Yes	No	Yes	0.86	0.86	5.93 n.s.
CSWTA11B	SSR	Codominant	Y7	Yes	No	Yes	0.86	0.82	16.58**
CSWTAAA01	SSR	Codominant	F4,S4	Yes	NO	Yes	0.86	0.83	9.57**
CT269	SSR	Codominant		Yes	No	Yes	0.86	0.76	8.29*

Table 1 continued

Marker	Marker type	Polymorphism type	Linkage group	Selection in BC$_1$	Selection in BC$_2$	Used in final evaluation	Expected allele frequency	Observed allele frequency	G-test
D11SNPG3H1	SNP	Codominant		No	No	Yes	0.86	0.98	28.75**
EST-2057	EST	Codominant		No	No	Yes	0.86	0.79	17.74**
EST-206	EST	Codominant		No	No	No	0.86	0.54	116.24**
EST-285	EST	Codominant		No	No	No	0.86	0.42	255.91**
EST-302	EST	Codominant		No	No	Yes	0.86	0.80	4.27 n.s.
EST-315	EST	Codominant		No	No	Yes	0.86	0.79	12.14**
F4	SCAR	Dominant	F2	Yes	No	No			
F62-89	SSR	Codominant		No	No	Yes	0.86	0.90	7.30*
F8-90A	SSR	Codominant		No	No	Yes	0.86	0.72	11.19**
L1LG3H3	SNP	Codominant		No	No	No	0.86	0.70	89.66**
M-5-BE-L	BAC end	Codominant		Yes	No	Yes	0.86	0.78	6.21*
M-6-BE-R	BAC end	Codominant		Yes	No	Yes	0.86	0.82	1.58 n.s.
S_BC82	SNP	Codominant	Y6	No	No	Yes	0.86	0.82	9.60**
S_C9	SNP	Codominant		No	No	Yes	0.86	0.89	0.92 n.s.
S_F2	SNP	Codominant		No	No	Yes	0.86	0.75	21.08**
S_G9	SCAR	Dominant		No	Yes	No			
S_H5	SNP	Codominant		No	No	Yes	0.86	0.85	0.30 n.s.
S-AB14-2	SCAR	Dominant		No	Yes	No			
S-AV16-3	SCAR	Dominant	Y2	No	Yes	No			
S-M2-1	SCAR	Codominant		No	No	No	0.86	0.79	42.23**
S-ME9EM18	SCAR	Codominant	Y5	No	No	No	0.86	0.16	567.57**
W7SNPG1H3	SNP	Codominant		No	Yes	Yes	0.86	0.86	3.07 n.s.

Linkage groups F1, F2, F3, F4, F5, F6, and F6 are from Fazio et al. 2003; linkage groups S1, S2, S3, S4, S5, S6, and S7 are from Sun et al. 2007; linkages groups Y1, Y2, Y3, Y4, Y5, Y6, and Y6 are from Yuan et al. 2008

BC$_1$ indicates markers used during the marker assisted selection (*MAS*) of the BC$_1$ generation

BC$_2$ indicates markers used during MAS of the BC$_2$ generation

Yes indicates the use of that marker during selection or evaluation, No indicates the exclusion of the marker

Expected allele frequency of the homozygous genotype of the recurrent parent in the BC$_2$S$_3$ generation

Observed allele frequency towards the homozygous genotype of the recurrent parent at BC$_2$S$_3$ generation

G-test [G = 2 $\sum$ O * ln (O/E)] on observed allelic frequencies of the BC$_2$S$_3$ lines compared to their expected BC$_2$S$_3$ allelic frequencies

n.s., *, and ** are not significant, $P \leq 0.05$, and ≤ 0.01, respectively

Individual plants were evaluated for days to anthesis, sex expression, number of lateral branches, yield over two harvests, fruit length, fruit weight, and exterior fruit quality. Days to anthesis was recorded as the number of days after sowing until the appearance of the first fully expanded corolla. For sex expression classification, plants were given a numerical value (designated by a sex "score") based on their relative gynoecy as gynoecious (2), predominantly female (1), or monoecious (0) sex expression. A plant was considered gynoecious if all flowers within the first ten nodes of the plant were pistillate. Plants were classified as predominately female if greater than 51% of flowers on the first ten nodes were pistillate. If plants possessed 50% or fewer pistillate flowers within the first ten nodes, they were designated as monoecious. Number of lateral branches was evaluated on branches greater than 5 cm in length in the first ten nodes on the main stem for each plant at anthesis.

Fruit yield, length, weight, and exterior quality were collected twice for each plant, 1 week apart. Only plants having marketable fruit > 20 cm long were harvested, such that plants were divided into three harvest groups in 2007 and five harvest groups in 2008. Harvest group evaluation (i.e., plants harvested at the same time) allowed for an assessment of earliness (days from sowing to harvest). Fruit weight (gm) and length (cm) was measured for each mature fruit ($\sim$5 cm in diameter), and fruit were judged by subjective ratings for spines (absence vs. presence), ridges (absence vs. presence) and sunburst coloration at the blossom-end (absence vs. presence). The fruit yield was calculated as the cumulative number of fruits collected over the two harvests.

The Netherlands phenotypic evaluations

During the fall of 2008 (July–October) and winter of 2009 (January–April), all IBL (116), the donor parent (PI 432858), and F$_2$ progeny (self of NZ1 × PI 432858) were evaluated in Haelen (by Nunhems Vegetable Seeds) and Bergschenhoek (by De Ruiter Zonen Seeds), The Netherlands in commercial, European-style greenhouses. The experiment was arranged in a randomized complete block design (RCBD), consisting of two blocks [one block (GH) in each of the two locations]. Plots within a block consisted of three plants of each entry. Seeds of each entry were sown and seedlings ($\sim$1- to 2-leaf stage) were transplanted to

hydroponic culture (rock wool media) approximately 3 weeks post-seeding, where they were grown under a 14 h photoperiod at maximum and minimum temperatures of 25 and 20°C, respectively, using standard European cultural practices (Papadopoulos and Hao 2000). The same trait evaluations and collection procedures conducted in the USA (i.e., UWM) were performed in The Netherlands.

Location comparative analyses

For comparative analyses, the same subset of 38 IBL and parental lines were evaluated using data from all three locations (Table 2). Analyses of variances (ANOVA) were performed using the *proc glm* procedure in SAS software (version 9.1 for Windows; SAS Institute, Cary NC), treating location, blocks within locations, lines, and location-by-lines as random effects. Only the USA location consisted of multiple blocks repeated over years, and thus each year was considered as a separate location for this analysis. Block effects were solely derived from the USA location since the USA was the only location to examine multiple blocks within a location.

Least square means were calculated for each trait by block (for the USA 2007 and 2008 experiments) and location (for the two locations in The Netherlands) using the *lsmeans* statement within *proc glm* in SAS. Rank correlations (r$_s$) between locations were performed using *proc corr* (spearman option) in SAS. The least square means for each trait were plotted by location (based on block number) to define genotype-by-environment (G × E) interactions.

Repeatability measures were performed for all traits to predict potential performance within the IBL population (Falconer and Mackay 1989). The variance components for the location, block, and line for each trait were estimated using the *covtest* option in the *proc mixed* procedure in SAS. All variables were treated as random effects, and repeatability (r) measures were calculated according to Falconer and Mackay (1989). The standard error (SE) of the repeatability measures was adapted from Hallauer and Miranda (1988).

Phenotypic analyses of the IBL population

The morphological characteristics of the complete set of IBL (116) were evaluated through the *proc glm* procedure in SAS using data collected in Haelen and

Table 2 Trait effects partitioned between location, replication, line, and location-by-line interaction for 38 inbred backcross cucumber (*Cucumis sativus* L.) lines (BC$_2$S$_3$) in three locations (Wisconsin, USA; Haelen, The Netherlands; and Bergschenhoek, The Netherlands) in a European Long market type background

Trait	Location		Block (location)		Line		Location-by-line	
	df	F-score	df	F-score	df	F-score	df	F-score
Days to anthesis	3	41.30**	4	11.90**	40	1.73*	115	1.65**
Sex expression	3	9.16**	4	0.75 n.s.	40	24.70**	117	1.36*
Lateral branch number	3	54.96**	4	15.88**	40	3.26**	117	2.61**
Harvest group	3	24.20**	4	6.40**	40	2.39**	113	3.09**
Yield over two harvests	3	0.67 n.s.	4	42.20**	40	3.96**	109	1.15 n.s.
Fruit traits								
Fruit length	3	0.06 n.s.	4	9.47**	40	5.24**	113	2.42**
Fruit weight	3	27.90**	4	7.58**	40	2.55**	113	1.47**
Sunburst coloration	3	5.39*	4	7.84**	40	3.01**	110	1.85**
Ribbed fruit	3	2.12 n.s.	4	14.90**	40	4.10**	110	1.57**
Spines on fruit	3	2.17 n.s.	4	4.88**	40	7.55**	110	2.25**

Days to anthesis is recorded as the number of days between transplanting and the appearance of the first fully expanded corolla; Sex expression recorded is as sex score were gynoecious, predominately female, and monoecious plants were given values of 2, 1, and 0, respectively; Lateral branch number is recorded as number of lateral branches within the first ten nodes of individual plants; Harvest group is the assessment of earliness (days from sowing to harvest); Yield over two harvests is the total number of fruit per plant over two harvests; Fruit length is average length of fruit per plot; Fruit weight is the average weight of fruit per plot; Sunburst coloration is the presence of yellow at the blossom-end of the fruit measured as presence or absence; Ribbed fruit is the presence or absence of ribs on fruit; and Spine is the presence or absence of spines on exterior of fruit

df degrees of freedom for the numerator

*,**, n.s. indicate significance at $P \leq 0.05$, ≤ 0.01, and not significant, respectively

Bergschenhoek, The Netherlands. Locations (i.e., single blocks), lines, and location-by-lines interaction were treated as random effects. Least square means were calculated for all traits for both locations using the *lsmeans* option within *proc glm* in SAS.

Spearman rank correlations (r_s) were performed using the *proc corr* (Spearman option) procedure in SAS for the locations in The Netherlands across all traits. Repeatability measures (r) were calculated using *proc mixed* with the *covtest* option to obtain trait variance components (Falconer and Mackay 1989).

Principal component analysis (PCA) was performed using phenotypic data to further elucidate relationships between parental lines (NZ1 and PI 432858) and IBL (116). Germplasm relationships were visualized using the *proc princomp* procedure within SAS followed by multidimensional graphic presentation via the 3D plot option within NTsys (Rohlf 1998). Since only sex expression, fruit number, fruit length, and spine occurrence were uniformly evaluated in 38 IBL and correlated ($r_s = 0.25$–0.80; $P < 0.05$; Table 3) across locations, data for these traits as recorded across IBL were included in PCA.

Molecular data collection

Seeds of 116 IBL were sown in a greenhouse at the UWM in 2008, and tissue from seedlings at the first true-leaf stage was harvested and bulked by IBL for DNA analysis. DNA was extracted according to Fazio et al. (2003), and PCR was performed using DNA from IBL and parental lines (NZ1 and PI 432858) as primed by 55 markers (Table 1; Fazio et al. 2003; Kong et al. 2006; Ritschel et al. 2004). Marker polymorphisms were visualized by digital imaging according to Fazio et al. (2003).

Genotypic analysis

The markers employed were evaluated for predicted segregation of codominant (1:2:1) and dominant marker (3:1) ratios by goodness of fit testing. The G-test was employed for assessment of marker distributions (Sokal and Rohlf 1994), and the degrees of freedom (df) equaled two in all tests due to the possibility of identifying either homozygosity or heterozygosity at a locus. Homogeneity tests were

Table 3 Rank correlations between three locations (Wisconsin, USA; and Haelen and Bergschenhoek, The Netherlands) for 10 traits evaluated in 38 cucumber (*Cucumis sativus* L.) inbred backcross lines (BC$_2$S$_3$) partitioned by block

	UW 1-1	UW 1-2	UW 1-3	UW 2-1	UW 2-2	UW 2-3	H	B
Days to anthesis								
UW 1-1		0.61**	0.43**	0.16 n.s.	0.33*	0.21 n.s.	0.02 n.s.	0.26 n.s.
UW 1-2	0.61**		0.51**	0.09 n.s.	0.27 n.s.	0.34*	0.15 n.s.	0.31*
UW 1-3	0.43**	0.51**		0.11 n.s.	0.20 n.s.	0.52**	0.26 n.s.	0.18 n.s.
UW 2-1	0.16 n.s.	0.09 n.s.	0.11 n.s.		0.15 n.s.	0.14 n.s.	0.14 n.s.	−0.02 n.s.
UW 2-2	0.33*	0.27 n.s.	0.20 n.s.	0.15 n.s.		0.51**	0.15 n.s.	0.07 n.s.
UW 2-3	0.21 n.s.	0.34*	0.52**	0.14 n.s.	0.51**		0.06 n.s.	0.16 n.s.
H	0.02 n.s.	0.15 n.s.	0.26 n.s.	0.14 n.s.	0.15 n.s.	0.06 n.s.		0.31*
B	0.26 n.s.	0.31*	0.18 n.s.	−0.02 n.s.	0.07 n.s.	0.16 n.s.	0.31*	
Sex expression								
UW 1-1		0.77**	0.75**	0.76**	0.69**	0.73**	0.71**	0.80**
UW 1-2	0.77**		0.76**	0.62**	0.72**	0.83**	0.55**	0.79**
UW 1-3	0.75**	0.76**		0.75**	0.63**	0.66**	0.76**	0.77**
UW 2-1	0.76**	0.62**	0.75**		0.52**	0.65**	0.70**	0.79**
UW 2-2	0.69**	0.72**	0.63**	0.52 **		0.71**	0.65**	0.69**
UW 2-3	0.73**	0.83**	0.66**	0.65**	0.71**		0.55**	0.72**
H	0.71**	0.55**	0.76**	0.70**	0.65**	0.55**		0.72**
B	0.80**	0.79**	0.77**	0.79**	0.69**	0.72**	0.72**	
Lateral branch number								
UW 1-1		0.50**	0.54**	0.05 n.s.	0.35*	0.40**	0.42**	0.39*
UW 1-2	0.49**		0.52**	−0.13 n.s.	0.17 n.s.	0.27 n.s.	0.49**	0.50**
UW 1-3	0.54**	0.52**		0.00 n.s.	0.39*	0.48**	0.43**	0.43**
UW 2-1	0.05 n.s.	−0.13 n.s.	0.00 n.s.		0.31*	0.23 n.s.	0.05 n.s.	0.11 n.s.
UW 2-2	0.35*	0.17 n.s.	0.39*	0.31*		0.48**	0.18 n.s.	0.30 n.s.
UW 2-3	0.40**	0.27 n.s.	0.48**	0.23 n.s.	0.48**		0.20 n.s.	0.26 n.s.
H	0.42**	0.49**	0.43**	0.05 n.s.	0.18 n.s.	0.20 n.s.		0.42**
B	0.39*	0.50**	0.44**	0.11 n.s.	0.30 n.s.	0.26 n.s.	0.42**	
Harvest group								
UW 1-1		0.36*	0.03 n.s.	0.28 n.s.	0.27 n.s.	0.25 n.s.	0.19 n.s.	0.34*
UW 1-2	0.36*		0.18 n.s.	0.25 n.s.	0.32*	0.49**	0.18 n.s.	0.09 n.s.
UW 1-3	0.03 n.s.	0.18 n.s.		−0.20 n.s.	0.06 n.s.	−0.02 n.s.	−0.16 n.s.	−0.13 n.s.
UW 2-1	0.28 n.s.	0.25 n.s.	−0.20 n.s.		0.43**	0.42**	0.53**	0.38*
UW 2-2	0.27 n.s.	0.32*	0.06 n.s.	0.43**		0.38*	0.18 n.s.	−0.04 n.s.
UW 2-3	0.25 n.s.	0.48**	−0.02 n.s.	0.42**	0.38*		0.25 n.s.	0.17 n.s.
H	0.19 n.s.	0.18 n.s.	−0.16 n.s.	0.53**	0.18 n.s.	0.25 n.s.		0.40**
B	0.34*	0.09 n.s.	−0.13 n.s.	0.38*	−0.04 n.s.	0.17 n.s.	0.40**	
Yield over two harvests								
UW 1-1		0.61**	0.64**	0.50**	0.48**	0.45**	−0.08 n.s.	0.49**
UW 1-2	0.61**		0.40**	0.52**	0.34*	0.35*	0.24 n.s.	0.43**
UW 1-3	0.64**	0.40**		0.31*	0.38*	0.33*	0.15 n.s.	0.48**
UW 2-1	0.50**	0.52**	0.31*		0.35*	0.52**	0.18 n.s.	0.40**
UW 2-2	0.48**	0.34*	0.38*	0.35*		0.45**	−0.04 n.s.	0.32*
UW 2-3	0.45**	0.35*	0.33*	0.52**	0.45**		0.13 n.s.	0.34*

Table 3 continued

	UW 1-1	UW 1-2	UW 1-3	UW 2-1	UW 2-2	UW 2-3	H	B
H	−0.08 n.s.	0.24 n.s.	0.15 n.s.	0.18 n.s.	−0.04 n.s.	0.13 n.s.		0.32*
B	0.49**	0.43**	0.48**	0.40**	0.32*	0.34*	0.32*	
Fruit traits								
Fruit length								
UW 1-1		0.36*	0.27 n.s.	0.44**	0.43**	0.12 n.s.	0.40**	0.11 n.s.
UW 1-2	0.36*		0.00 n.s.	0.49**	0.46**	0.53**	0.48**	0.34*
UW 1-3	0.27 n.s.	0.00 n.s.		0.15 n.s.	0.34*	0.15 n.s.	0.37*	0.29 n.s.
UW 2-1	0.44**	0.49**	0.15 n.s.		0.60**	0.29 n.s.	0.44**	0.17 n.s.
UW 2-2	0.43**	0.46**	0.34*	0.60**		0.55**	0.51**	0.25 n.s.
UW 2-3	0.12 n.s.	0.53**	0.15 n.s.	0.29 n.s.	0.55**		0.30 n.s.	0.37*
H	0.40**	0.48**	0.37*	0.44**	0.51**	0.30 n.s.		0.35*
B	0.11 n.s.	0.34*	0.29 n.s.	0.17 n.s.	0.25 n.s.	0.37*	0.35*	
Fruit weight								
UW 1-1		0.16 n.s.	0.10 n.s.	0.46**	0.51**	−0.19 n.s.	0.09 n.s.	0.01 n.s.
UW 1-2	0.16 n.s.		−0.28 n.s.	0.24 n.s.	−0.01 n.s.	0.07 n.s.	0.21 n.s.	0.11 n.s.
UW 1-3	0.10 n.s.	−0.28 n.s.		−0.01 n.s.	0.00 n.s.	0.09 n.s.	0.00 n.s.	0.09 n.s.
UW 2-1	0.46**	0.24 n.s.	−0.01 n.s.		0.34*	−0.02 n.s.	0.18 n.s.	0.07 n.s.
UW 2-2	0.51**	−0.01 n.s.	0.00 n.s.	0.34*		0.04 n.s.	0.49**	0.16 n.s.
UW 2-3	−0.19 n.s.	0.07 n.s.	0.09 n.s.	−0.02 n.s.	0.04 n.s.		0.24 n.s.	−0.01 n.s
H	0.09 n.s.	0.21 n.s.	0.00 n.s.	0.18 n.s.	0.49**	0.24 n.s.		0.33*
B	0.01 n.s.	0.11 n.s.	0.09 n.s.	0.07 n.s.	0.16 n.s.	−0.01 n.s.	0.33*	
Sun burst								
UW 1-1		0.39**	0.58**	0.45**	0.33*	0.13 n.s.	0.03 n.s.	0.40**
UW 1-2	0.39*		0.29 n.s.	0.47**	0.11 n.s.	−0.06 n.s.	0.01 n.s.	0.22 n.s.
UW 1-3	0.58**	0.29 n.s.		0.24 n.s.	0.17 n.s.	−0.04 n.s.	0.27 n.s.	0.02 n.s.
UW 2-1	0.45**	0.47**	0.24 n.s.		0.66**	0.22 n.s.	0.18 n.s.	0.59**
UW 2-2	0.33*	0.11 n.s.	0.17 n.s.	0.66**		0.17 n.s.	0.20 n.s.	0.48**
UW 2-3	0.13 n.s.	−0.06 n.s.	−0.04 n.s.	0.22 n.s.	0.17 n.s.		0.18 n.s.	0.58**
H	0.03 n.s.	0.01 n.s	0.27 n.s.	0.18 n.s.	0.20 n.s.	0.18 n.s.		0.21 n.s.
B	0.40**	0.22 n.s.	0.02 n.s.	0.59**	0.48**	0.58**	0.21 n.s.	
Ribbed								
UW 1-1		0.28 n.s.	0.37*	0.24 n.s.	0.60**	0.45**	0.28 n.s.	0.40**
UW 1-2	0.28 n.s.		0.20 n.s.	0.13 n.s.	0.30 n.s.	0.27 n.s.	0.08 n.s.	0.12 n.s.
UW 1-3	0.37*	0.20 n.s.		0.28 n.s.	0.30 n.s.	0.35*	−0.01 n.s.	0.16 n.s.
UW 2-1	0.24 n.s.	0.13 n.s.	0.28 n.s.		0.30 n.s.	0.22 n.s.	0.22 n.s.	0.47**
UW 2-2	0.60**	0.30 n.s.	0.30 n.s.	0.30 n.s.		0.32*	0.26 n.s.	0.38*
UW 2-3	0.45**	0.27 n.s.	0.35*	0.22 n.s.	0.32*		0.19 n.s.	0.23 n.s.
H	0.28 n.s.	0.08 n.s.	−0.01 n.s.	0.22 n.s.	0.26 n.s.	0.19 n.s.		0.54**
B	0.40**	0.12 n.s.	0.16 n.s.	0.47**	0.38*	0.23 n.s.	0.54**	
Spine								
UW 1-1		0.60**	0.39*	0.54**	0.71**	0.46**	0.67**	0.67**
UW 1-2	0.60**		0.33*	0.46**	0.46**	0.65**	0.55**	0.56**
UW 1-3	0.39*	0.33*		0.38*	0.28 n.s.	0.37*	0.35*	0.25 n.s.
UW 2-1	0.54**	0.46**	0.38*		0.68**	0.67**	0.63**	0.60**

Table 3 continued

	UW 1-1	UW 1-2	UW 1-3	UW 2-1	UW 2-2	UW 2-3	H	B
UW 2-2	0.71**	0.46**	0.28 n.s.	0.68**		0.54**	0.54**	0.65**
UW 2-3	0.46**	0.65**	0.37*	0.67**	0.54**		0.70**	0.63**
H	0.67**	0.55**	0.35*	0.63**	0.54**	0.70**		0.74**
B	0.67**	0.56**	0.25 n.s.	0.60**	0.65**	0.63**	0.74**	

UW1-1, UW1-2, and UW1-3 represents evaluations performed at the University of Wisconsin, Madison, Wisconsin (UW) in 2006 for blocks 1, 2, and 3, respectively. UW2-1, UW2-2, and UW2-3 represents evaluations performed at UW in 2007 for blocks 1, 2, and 3, respectively

H represents evaluations performed in Haelen, The Netherlands

B represents evaluations performed in Bergschenhoek, The Netherlands

*,**, n.s. indicate significance at $P \leq 0.05$, ≤ 0.01, and nonsignificant, respectively

also performed using the Pearson χ^2 test (Pearson 1900) with df = (number of lines − 1) * (number of classes − 1) to test whether allelic frequencies were distributed similarly across the markers examined.

Based on the χ^2 segregation analyses and marker consistency (repeatability and banding intensity) during genotyping, the 37 codominant markers (two alleles/locus) were employed in multidimensional scaling (MDS) to evaluate relationships among the 116 IBL and parental lines (NZ1 and PI 432858) using NTsys 2.01 computer software (Rohlf 1998). While homozygous individuals were assigned an allelic frequency of 1 at a given locus, heterozygous lines were assigned a locus allelic frequency of 0.5. The GD of Rogers (Rogers 1972), as modified by Wright (Wright 1978), was used in the Simgend procedure in NTsys to calculate a GD matrix for the 116 IBL and the two parental lines. An MDS analysis was performed using data from the GD matrix in NTsys, where results were visualized using its 3D-plot option (Rohlf 1998).

Results

Comparative morphological analyses

The main effects of location, block, and line (IBL) and the first order interaction (location-by-line) for the 38 IBL evaluated in The Netherlands (Haelen and Bergschenhoek) in 2008 and over 2 years in the USA (2007 and 2008) are presented in Table 2. Locations performed similarly for the cumulative two-harvest yield, fruit length, and the occurrence of ribbed and spiny fruit ($P \leq 0.05$). Likewise, the degree of gynoecy was similar

between blocks within a test location (i.e., USA). Also, no significant ($P \leq 0.05$) location-by-line interaction was detected for cumulative two-harvest yield. Nevertheless, lines differed ($P \leq 0.05$) in all traits.

Pairwise rank correlations for traits among the 38 IBL examined in all locations are presented in Table 3. Sex expression, cumulative two-harvest yield, fruit length, and the occurrence of spines on fruits were traits that demonstrated the greatest correlations between locations ($r_s = 0.11$–0.80). While sex expression and the occurrence of spines on fruit correlated across locations ($r_s = 0.25$–0.80), correlations in fruit length were highest between the USA and Haelen ($r_s = 0.51$), and correlations for cumulative two-harvest yield were highest between the USA and Bergschenhoek ($r_s = 0.48$).

Sex expression was highly repeatable in the 38 lines evaluated (Table 4). The traits, fruit length and spines on fruits, had moderately high repeatability measures. In contrast, days to anthesis, lateral branch number, harvest group, yield over two harvests, fruit weight, the presence of ribs on fruit, and sunburst coloration at the blossom-end of fruit demonstrated comparatively low measures of repeatability. The standard error of the repeatability measures for these traits was less than half that of the repeatability estimate themselves.

Phenotypic analysis of the IBL population

The location and line (IBL) main effects, and location-by-line interaction for all traits in 116 IBL are presented in Table 5. With the exception of sex expression and lateral branch number, all traits were significantly affected ($P \leq 0.05$) by growing location.

Table 4 Components of variances and repeatability measures with associated standard errors of morphological traits evaluated in 38 cucumber (*Cucumis sativus* L.) inbred backcross lines (BC$_2$S$_3$) taken collectively over three locations {USA (Madison, Wisconsin; 2007, 2008) and The Netherlands [Haelen (2008) and Bergschenhoek (2009)]} during three evaluation years

Trait	σ^2_{loc}	σ^2_{rep}	σ^2_{line}	σ^2_{int}	r
Traits examined across all locations					
Days to anthesis	42.94 ± 35.96	1.37 ± 1.06	1.51 ± 0.73	2.45 ± 0.70	0.03 ± 0.02
Sex expression	0.01 ± 0.01	0.00 ± 0.00	0.42 ± 0.10	0.01 ± 0.01	0.95 ± 0.22
Lateral branch	12.14 ± 10.08	0.29 ± 0.22	0.64 ± 0.21	0.67 ± 0.16	0.05 ± 0.02
Harvest group	0.86 ± 0.74	0.06 ± 0.05	0.20 ± 0.08	0.32 ± 0.08	0.14 ± 0.06
Yield over two harvests	0.00 ± 0.00	0.37 ± 0.21	0.35 ± 0.10	0.07 ± 0.04	0.45 ± 0.12
Fruit traits					
Length	0.00 ± 0.00	0.24 ± 0.15	4.29 ± 1.13	1.23 ± 0.30	0.75 ± 0.20
Weight	3542.39 ± 3008.75	186.17 ± 152.16	537.18 ± 191.42	285.29 ± 111.38	0.12 ± 0.04
Sunburst	0.03 ± 0.03	0.01 ± 0.01	0.03 ± 0.01	0.02 ± 0.01	0.35 ± 0.12
Ribbed	0.02 ± 0.03	0.02 ± 0.02	0.04 ± 0.01	0.01 ± 0.01	0.44 ± 0.13
Spine	0.00 ± 0.00	0.00 ± 0.00	0.09 ± 0.02	0.03 ± 0.01	0.73 ± 0.19
Traits examined across locations used in the principal component analysis					
Sex expression	0.01 ± 0.01	0.00 ± 0.00	0.42 ± 0.10	0.01 ± 0.01	0.95 ± 0.22
Yield over two harvests	0.00 ± 0.00	0.37 ± 0.21	0.35 ± 0.10	0.07 ± 0.04	0.45 ± 0.12
Fruit length	0.00 ± 0.00	0.24 ± 0.15	4.29 ± 1.13	1.23 ± 0.30	0.75 ± 0.20
Spine	0.00 ± 0.00	0.00 ± 0.00	0.09 ± 0.02	0.03 ± 0.01	0.73 ± 0.19

Days to anthesis is recorded as the number of days between transplanting and the appearance of the first fully expanded corolla; Sex expression recorded is as sex score were gynoecious, predominately female, and monoecious plants were given values of 2, 1, and 0, respectively; Lateral branch number is recorded as number of lateral branches within the first ten nodes of individual plants; Harvest group is the assessment of earliness (days from sowing to harvest); Yield over two harvests is the total number of fruit per plant over two harvests; Fruit length is average length of fruit per plot; Fruit weight is the average weight of fruit per plot; Sunburst coloration is the presence of yellow at the blossom-end of the fruit measured as presence or absence; Ribbed fruit is the presence or absence of ribs on fruit; and Spine is the presence or absence of spines on exterior of fruit

σ^2_{loc} = variance within location

σ^2_{line} = variance within replications

σ^2_{line} = variance within lines

σ^2_{int} = variance of the genotype by environment interaction

r repeatability of trait, where r = the σ^2_{lines} divided by σ^2_{P} and $\sigma^2_{\text{P}} = \sigma^2_{\text{lines}} + \sigma^2_{\text{rep}} + \sigma^2_{\text{loc}} + \sigma^2_{\text{int}}$

Likewise, differences among lines and lines across locations were detected for all traits ($P \leq 0.05$).

All traits, except for the presence of blossom-end sunburst in fruit, demonstrated significant correlations ($P \leq 0.05$) between growing locations in The Netherlands (Table 6), which varied between r$_s$ = 0.29 and 0.81.

Repeatability measures for all traits within the growing locations in The Netherlands are presented in Table 7. Sex expression and presence of spines on fruits demonstrated high repeatability. Lateral branch number, harvest group, fruit length, fruit weight, and the presence of ribs demonstrated moderate repeatability across growing locations, where the standard error of the repeatability measures for these traits was less than half that of the repeatability estimates themselves. In contrast, the repeatability of days to anthesis, cumulative two-harvest yield, and sunburst coloration at the blossom end of the fruit were comparatively low.

Principal components (PC) 1–3 accounted for 88% of the observed phenotypic variation among IBL (PC1 = 49%, PC2 = 22%, and PC3 = 17%) (Fig. 2). While the four traits evaluated contributed equally to the ordination of IBL by PC1, spines on fruit was mainly responsible for IBL ordination in PC2, and fruit length largely determined IBL ordination in PC3.

Table 5 Effects partitioned between location, line and location-by-line interaction for 116 cucumber (*Cucumis sativus* L.) inbred backcross lines (BC$_2$S$_3$) as evaluated in Haelen and Bergschenhoek, The Netherlands

Trait	Location		Line		Location-by-line	
	df	F-score	df	F-score	df	F-score
Days to anthesis	1	1565**	114	2.08**	102	1.66**
Sex expression	1	0.01 n.s.	114	17.84**	113	1.58**
Lateral branch number	1	0.33 n.s.	144	1.81**	113	3.15**
Harvest group	1	82.84**	113	2.31**	107	1.52**
Yield over two harvests	1	77.59**	114	1.76**	114	1.93**
Fruit traits						
Fruit length	1	9.87**	114	3.18**	108	4.09**
Fruit weight	1	38.98**	114	3.27**	108	1.89**
Sunburst	1	13.26**	113	1.45*	104	6.51**
Ribbed	1	23.59**	114	2.22**	107	4.50**
Spines	1	4.32*	114	7.30**	107	1.88**

Days to anthesis is recorded as the number of days between transplanting and the appearance of the first fully expanded corolla; Sex expression recorded is as sex score were gynoecious, predominately female, and monoecious plants were given values of 2, 1, and 0, respectively; Lateral branch number is recorded as number of lateral branches within the first ten nodes of individual plants; Harvest group is the assessment of earliness (days from sowing to harvest); Yield over two harvests is the total number of fruit per plant over two harvests; Fruit length is average length of fruit per plot; Fruit weight is the average weight of fruit per plot; Sunburst coloration is the presence of yellow at the blossom-end of the fruit measured as presence or absence; Ribbed fruit is the presence or absence of ribs on fruit; and Spine is the presence or absence of spines on exterior of fruit

df degrees of freedom for the numerator

*,**, n.s. indicate significance at $P \leq 0.05$, ≤ 0.01, and not significant, respectively

The morphology of a subset of IBL (4, 10, 16, 33, 36, 41, 43, 44, 46, 50, 51, 56, 68, 72, 78, 80, 82, 83, 89, 90, 91, 95, 99, 103, 104, 116, 121, and 130) were found to be greatly divergent from each other and the remaining IBL based on the four traits evaluated. This subset circumscribed the variation of the IBL examined (Fig. 2).

Molecular analysis

Results of Pearson-based χ^2 and G-test analyses are presented in Table 1. The expected gene frequencies in IBL that was homozygous for recurrent parent (AA), heterozygous (Aa), and homozygous for donor parent (aa) alleles are 0.86, 0.03, and 0.11, respectively. Pearson-based marker homogeneity tests indicated that allelic frequencies at the marker loci examined were independent ($P \leq 0.01$).

Eighteen of the 55 markers used herein were previously mapped by Fazio et al. (2003) or Yuan et al. (2008). These eighteen markers were dispersed throughout the cucumber genome, with markers on all linkage groups—except Linkage Group 6 as defined by Fazio et al. (2003).

Genetic relationships between all IBL after MDS are presented in Fig. 3. IBL 28, 30, and 31 were determined to be genetically identical by marker analyses (GD = 0.00), and IBL 36 and 66 were the most distant from other IBL (GD = 0.77). Although a majority of the IBL evaluated were found to be dissimilar to PI 432858 (GD = 0.60–0.90), IBL 4, 28, 42, 55, 69, 79, 96, 117, and 125 were the least similar to PI 432858 (GD = 0.90).

Discussion

The genetic diversity of European Long cucumber is the narrowest (GD = 0.00–0.24) of all major commercial cucumber market types (Dijkhuizen et al. 1996). The information provided herein is the first large-scale, publicly initiated, strategic genetic assessment of European Long cucumber germplasm directed towards the enhancement of this market class. This MAS-driven assessment led to the development of a genetically diverse set of IBL that holds value for genetic studies and plant improvement in cucumber.

Table 6 Rank correlations between greenhouse locations in Haelen and Bergschenhoek, The Netherlands for cucumber (*Cucumis sativus* L.) traits in 116 inbred backcross lines (BC$_2$S$_3$) evaluated in 2008 and 2009

Trait	Correlation
Days to anthesis	0.42*
Sex expression	0.81*
Lateral branch number	0.29*
Harvest group	0.38*
Yield over two harvests	0.22*
Fruit traits	
Length	0.44*
Weight	0.54*
Sunburst	0.12
Ribbed	0.40*
Spine	0.75*

Days to anthesis is recorded as the number of days between transplanting and the appearance of the first fully expanded corolla; Sex expression recorded is as sex score were gynoecious, predominately female, and monoecious plants were given values of 2, 1, and 0, respectively; Lateral branch number is recorded as number of lateral branches within the first ten nodes of individual plants; Harvest group is the assessment of earliness (days from sowing to harvest); Yield over two harvests is the total number of fruit per plant over two harvests; Fruit length is average length of fruit per plot; Fruit weight is the average weight of fruit per plot; Sunburst coloration is the presence of yellow at the blossom-end of the fruit measured as presence or absence; Ribbed fruit is the presence or absence of ribs on fruit; and Spine is the presence or absence of spines on exterior of fruit

* Indicate the correlation is significant at $P \leq 0.05$

Location comparisons

Evaluations of IBL were conducted under controlled conditions in Haelen and Bergschenhoek, The Netherlands, which mirror commercial hydroponic greenhouse conditions. A comparative evaluation of IBL grown in soil media under greenhouse conditions in the USA provides for a distinct horticultural production contrast. However, because of this difference and because of the fact that it only contains the subset of 38 IBL, the USA evaluations were used only as a point of reference and were not employed in PCA or used to develop subsequent conclusions.

Trait repeatability measures and rank correlations of the subset of 38 IBL between traits were used to determine which trait and location data were most appropriate for PCA (Table 3, 4). Cumulative two-harvest yield was the only trait of those evaluated in the IBL that was not substantially affected by environment (i.e., no G × E) (Table 2). Although cumulative two-harvest yield demonstrated marginal repeatability (0.51), trait performance of IBL at the Bergschenhoek location correlated well with other evaluation locations (Table 3 and 4). In contrast, cumulative two-harvest yield performance of IBL at the Haelen location correlated to a limited extent with Bergschenhoek ($r_s = 0.32$), but not with the USA locations ($r_s = -0.08$ to 0.24). Such location differences may partially explain the comparatively low repeatability of yield in the subset IBL (38) evaluated. Given the low correlation between the Haelen evaluation and the USA evaluations, only cumulative two-harvest yield performance from Bergschenhoek was used in PCA.

The repeatability values of sex expression (0.98), spines on fruit (r = 0.91) and fruit length (r = 0.95) between the subset of 38 IBL was relatively high (Table 4). Since rank correlations were high in both sex expression and occurrence of spines on fruit for both locations in The Netherlands ($r_s = 0.52$–0.80 and 0.25–0.71, respectively) (Table 3), combined location averages for these traits were used in PCA. In contrast, the correlations for fruit length between Bergschenhoek and the USA locations were relatively low ($r_s = 0.11$–0.37; Table 3), and, thus, only fruit length data from Haelen were used in PCA.

Genetic diversity in IBL

The trait attributes of cumulative two-harvest yield, sex expression, occurrence of spines on fruit, and fruit length allowed for their inclusion in PCA to provide for the identification of morphologically unique IBL (Fig. 2). Gynoecious IBL clustered along the right-hand side and top of the graphic projection where the gynoecious commercial parent, NZ1, was also located. High yielding IBL resided at the top and right-hand side of the projection, and IBL that developed spiny, long fruits clustered to the upper left-hand side of the figure where the donor parent, PI 432858 (spiny fruit, 37–39 cm in length) was also located. IBL exhibiting varying yield potential and fruit length were among the most divergent lines, and were located at the periphery of the figure. For instance, while IBL 99 developed, on average, the longest fruits (35.1 cm) and IBL 51 the shortest fruits (25.2 cm), IBL 46 bore the lowest number of fruit (0.7 fruits/plant) and IBL 5 and 116 the highest fruit

Table 7 Components of variances and their standard errors of pre- and post-harvest traits in 116 inbred backcross lines (BC$_2$S$_3$) European Long market class cucumber (*Cucumis sativus* L.) lines based on evaluations in Haelen and Bergschenhoek, The Netherlands in 2008 and 2009

Trait	σ^2_{loc}	σ^2_{line}	σ^2_{int}	r
Days to anthesis	27.64 ± 39.11	0.95 ± 0.31	0.80 ± 0.28	0.03 ± 0.01
Sex score	0.00 ± 0.00	0.49 ± 0.07	0.02 ± 0.01	0.96 ± 0.13
Lateral branch	0.00 ± 0.00	0.74 ± 0.24	1.12 ± 0.24	0.40 ± 0.13
Harvest group	0.78 ± 1.12	0.68 ± 0.17	0.29 ± 0.14	0.39 ± 0.10
Two harvest yield	0.74 ± 1.06	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Fruit traits				
Length	0.14 ± 0.22	1.86 ± 0.39	1.25 ± 0.24	0.57 ± 0.12
Weight	271.91 ± 392.90	821.35 ± 164.39	278.51 ± 88.32	0.60 ± 0.12
Sunburst	0.01 ± 0.02	0.02 ± 0.02	0.11 ± 0.02	0.16 ± 0.10
Ribbed	0.02 ± 0.03	0.06 ± 0.02	0.07 ± 0.01	0.38 ± 0.10
Spine	0.00 ± 0.00	0.15 ± 0.02	0.02 ± 0.01	0.87 ± 0.14

Days to anthesis is recorded as the number of days between transplanting and the appearance of the first fully expanded corolla; Sex expression recorded is as sex score were gynoecious, predominately female, and monoecious plants were given values of 2, 1, and 0, respectively; Lateral branch number is recorded as number of lateral branches within the first ten nodes of individual plants; Harvest group is the assessment of earliness (days from sowing to harvest); Yield over two harvests is the total number of fruit per plant over two harvests; Fruit length is average length of fruit per plot; Fruit weight is the average weight of fruit per plot; Sunburst coloration is the presence of yellow at the blossom-end of the fruit measured as presence or absence; Ribbed fruit is the presence or absence of ribs on fruit; and Spine is the presence or absence of spines on exterior of fruit

σ^2_{loc} = variance within location

σ^2_{line} = variance within lines

σ^2_{int} = variance of the genotype by environment interaction

r repeatability of trait, where r = the σ^2_{lines} divided by σ^2_{P} and $\sigma^2_{\text{P}} = \sigma^2_{\text{lines}} + \sigma^2_{\text{loc}} + \sigma^2_{\text{int}}$

number (4.3 fruits/plant). Since a line's position in the graphical depiction of the PCA is a function of its morphological attributes, IBL can be strategically selected for use in future genetic enhancement projects based on their graphical position.

Relative breeding value of traits

The evaluations (blocks) conducted in The Netherlands were considered random variables given that data were collected during different times of the year (i.e., Haelen fall vs. Bergschenhoek winter) in different greenhouses. Although G × E interactions between these growing locations were predictable and detected for all traits (Table 5), only the blossom-end starburst trait data was not correlated between locations ($r_s = 0.12$; Table 6). These results suggest that detected G × E interactions are of relatively minor importance for IBL characterization and that traits examined, except for blossom-end starburst color development on fruit, will perform similarly in different locations and seasons under standard European hydroponic culture conditions. This is in contrast to findings reported using Beit Alpha market class inbred lines (Delannay and Staub 2010), in which it was determined that selection among Beit Alpha IBL must be performed in specific growing environments due to low repeatability measures between the evaluated locations.

Concordance between phenotypic and genotypic appraisals

In some cases morphological and molecular data were concordant. For example, IBL 28, 30, and 31 were genetically identical (GD = 0.00) based on the markers employed in this study and possess morphological characteristics that are similar to one another (Fig. 2). However, generally, morphological and molecular appraisals of genetic diversity were not equivalent. For instance, IBL 5 and 116 share common morphological characteristics (Fig. 2) but do not possess substantial genetic affinities (GD = 0.39) (Fig. 3). This suggests that, in fact, genetic and
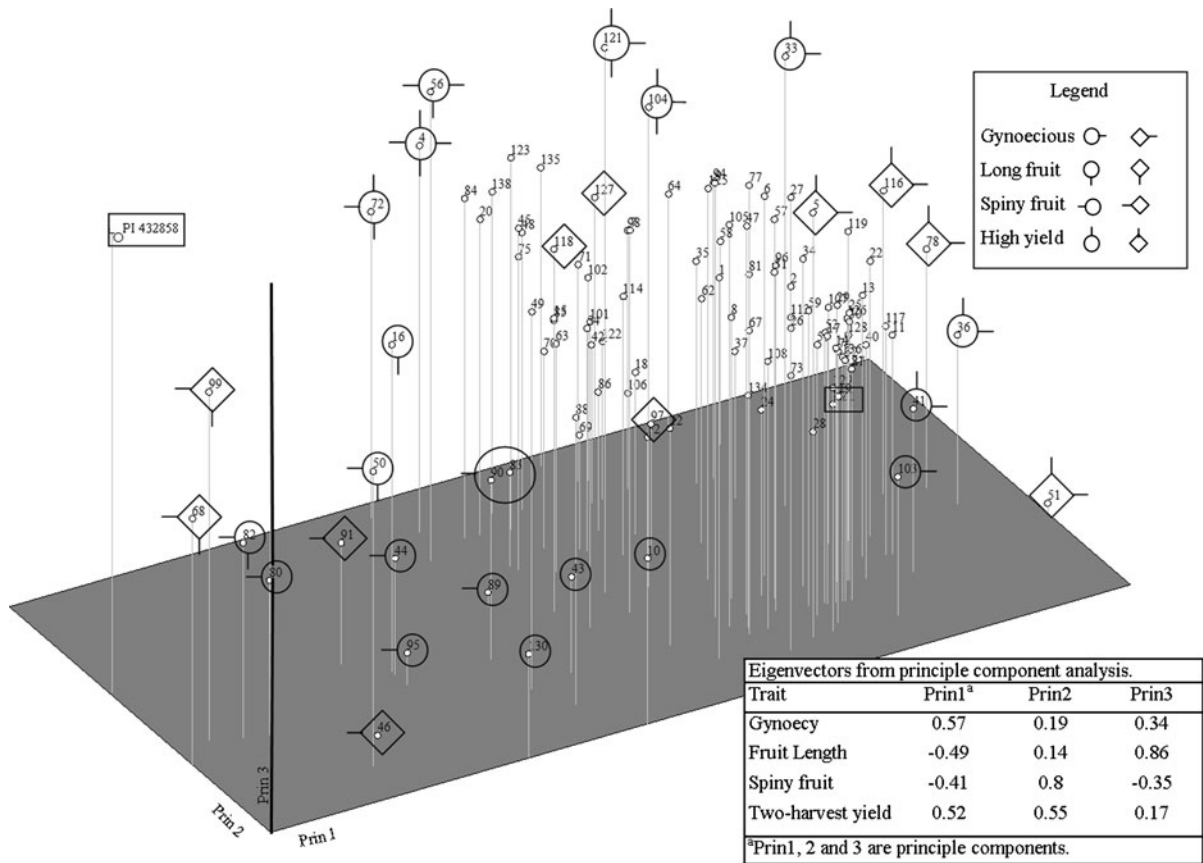
Fig. 2 Principal component-based graphic depiction of cucumber (*Cucumis sativus* L.) parental lines [*Square*; NZ1 (European Long market type) and PI 432858 (China)] and their derived BC$_2$S$_3$ progeny based on traits evaluated at Haelen and Berschenhoek, The Netherlands, in greenhouses in the fall of 2008 and winter 2009, respectively. *Backcross lines circled* are lines that deviated substantially from the progeny population

mean for a particular trait, and those designated by *diamonds* represent lines that had either the maximum or minimum value for a particular trait. *Lines* radiating from the *circles* and *diamonds* indicate which trait the *line* is uniquely characteristic for; gynoecious lines (*right*), long fruit (*bottom*), spiny fruit (*left*), and high yield (*top*)

morphological differences between European Long cucumber IBL can assist in the development of MAS strategies for increasing genetic diversity while fixing traits of economic importance. For instance, a cross between IBL 5 and 116 may enhance genetic diversity in this market class and will likely result in the recovery of gynoecious, early-flowering, high-yielding progeny in advanced generations through continued marker-assisted and phenotypic selection.

Potential use of IBL in plant improvement

The genetic improvement of this market type has reached a critical juncture, in which its narrow genetic base cannot predictably support the improvements needed in commercial yield and quality. The

stringent commercial standards for European Long cucumber market types would preclude the release of the IBL described herein as cultivars (i.e., commercial cultivars must have uniform fruit quality traits whereas plants within individual IBL tend to segregate for quality traits). However, some of the genetically diverse IBL possess phenotypic differences that are complementary (e.g., high yield potential and culinary attributes) and have utility in plant improvement. For instance, the gynoecious IBL 33 develops a relatively large number (~ 4 fruit/plant) of spineless, relatively smooth, uniformly green fruit. Moreover, it shares little genetic affinity with parental line NZ1 (GD = 0.46), and thus its inclusion in narrow-based breeding efforts would likely enhance genetic diversity. Likewise, the
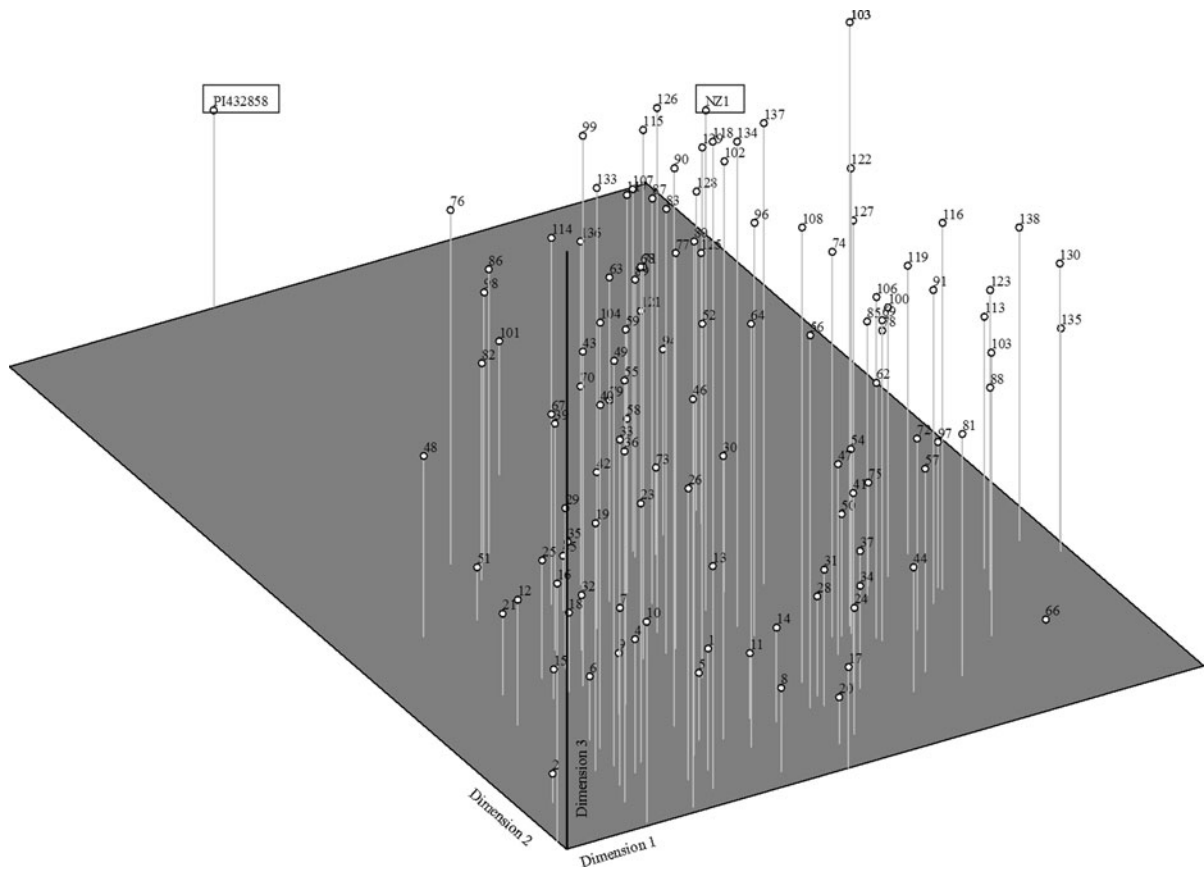
Fig. 3 Genetic relationships among cucumber (*Cucumis sativus* L.) parental lines NZ1 (European Long market type) and PI 432858 (China), and their derived BC$_2$S$_3$ progeny after multidimensional scaling as defined by 37 marker loci (SSR, SCAR, SNP, EST, and BAC-end) used for calculation of Rogers (1972) genetic distance as modified by Wright (1978)

gynoecious IBL 5 exhibits high yield (∼4 fruit/pant), multiple lateral branching (∼13 branches) and early flowering (∼12 days to anthesis) traits. This IBL could be intermated with the monoecious IBL 20 (1 lateral branches) or 127 (∼3 lateral branches) to develop genetically and phenotypically diverse populations for selection and subsequent inbred extraction. In fact, traits demonstrating consistency and repeatability (i.e., sex expression, fruit number, fruit length, and spine frequency) over diverse growing locations should be relatively easy to manipulate during selection (Table 7; Fazio et al. 2003).

Genetic mapping efforts in European Long cucumber could be enhanced by the development of novel mapping populations (Dijkhuizen et al. 1996). Given its morphological characteristics and genotypic diversity, IBL 33 could be considered as a mapping parent for the identification of yield and quality components. In this regard, IBL 90 contrasts with IBL 33 in

morphology and allelic constitution (GD = 0.67; Fig. 2). While the high yielding (∼4 fruits/plant), gynoecious IBL 33 produces fruit of relatively high quality, monoecious IBL 90 produces comparatively few fruit (∼3 fruits/plant) that are non-uniformly green and spiny (100% fruits with spines). The use of IBL such as these in the identification of QTL for yield and quality components would allow the use of markers in MAS for optimal genetic diversity and improved plant performance in this economically important cucumber market class.

References

Andeweg JM (1956) The breeding of scab-resistant frame cucumbers in The Netherlands. Euphytica 5:185–195

Delannay IY (2009) Use of molecular markers to increase genetic diversity of Beit Alpha, European Long, and U.S.

processing market classes of cucumber (*Cucumis sativus* L.) through marker-assisted selection. Dissertation, University of Wisconsin

Delannay IY, Staub JE (2010) Use of molecular markers aids in the development of diverse inbred backcross lines in Beit Alpha cucumber (*Cucumis sativus* L.). Euphytica 175:65–78

Dijkhuizen A, Kennard WC, Havey MJ, Staub JE (1996) RFLP variation and genetic relationships in cultivated cucumber. Euphytica 90:79–87

Ehret DL, Koch C, Menzies J, Sholberg P, Garland T (2001) Foliar sprays of clay reduce the severity of powdery mildew on long English cucumber and wine grapes. HortScience 36:934–936

Falconer DS, Mackay TFC (1989) Introduction to quantitative genetics. Benjamin Cummings, San Francisco

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

Hallauer AR, Miranda JB (1988) Quantitative genetics in maize breeding, 2nd edn. Wiley, Hoboken, NJ

Horejsi T, Staub JE (1999) Genetic variation in cucumber (*Cucumis sativus* L.) as assessed by random amplified polymorphic DNA. Genet Resour Crop Ev 46:337–350

Kong Q, Xiang C, Yu Z (2006) Development of EST-SSRs in *Cucumis sativus* from sequence database. Mol Ecol Notes 6:1234–1236

Lin WC, Jolliffe PA (1994) Canopy light affects shelf life of long English cucumber. Postharvest Phys Fruits 398:249–255

Meglic V, Staub JE (1996) Genetic diversity in cucumber (*Cucumis sativus* L): 2. An evaluation of selected cultivars released between 1846 and 1978. Genet Resour Crop Ev 43:547–558

Meglic V, Serquen F, Staub JE (1996) Genetic diversity in cucumber (*Cucumis sativus* L):1. A re-evaluation of the US germplasm collection. Genet Resour Crop Ev 43:533–546

Miliki A, Staub JE, Sun ZY, Ghorbel A (2003) Genetic diversity in African cucumber (*Cucumis sativus* L.) provides potential for germplasm enhancement. Genet Resour Crop Ev 50:461–468

Owens KW, Bliss FA, Peterson CE (1985) Genetic variation within and between two cucumber populations derived via the inbred backcross line method. J Amer Soc Hort Sci 110:437–441

Papadopoulos AP, Hao X (2000) Effects of day and night air temperature on growth, productivity and energy use of long English cucumber. Can J Plant Sci 80:143

Pearson K (1900) Primary: on a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philos Mag 50:158

Ritschel P, de Lima Lins T, Tristan R, Cortopassi-Buso G, Amauri-Buso J, Ferreira M (2004) Development of microsatellite markers from an enriched genomic library for genetic analysis of melon (*Cucumis melo* L.). BMC Plant Biol 4:9

Robbins MD, Casler M, Staub JE (2008) Pyramiding QTL for multiple lateral branching in cucumber using nearly isogenic lines. Mol Breed 22:131–139

Rogers JS (1972) Measures of genetic similarity and genetic distance. Studies Genet 7213:145–153

Rohlf FJ (1998) NTSYS-Pc v2.0 numerical taxonomy and multivariable analysis system. Appl Biostat, New York

Sokal RR, Rohlf FJ (1994) Biometry: the principles and practice of statistics in biological research, 3rd edn. Freeman, New York

Staub JE, Serquen FC, McCreight JD (1997) Genetic diversity in cucumber (*Cucumis sativus* L.): III. An evaluation of Indian germplasm. Genet Resour Crop Ev 44:315–326

Staub JE, Serquen FC, Horejsi T, Chen JF (1999) Genetic diversity in cucumber (*Cucumis sativus* L.): IV. An evaluation of Chinese germplasm. Genet Resour Crop Ev 46:297–310

Staub JE, Robbins MD, Wehner TC (2008) Cucumber. In: Prohens J, Nuez F (eds) Vegetables I: asteraceae brassicaceae chenopodiaceae, and cucurbitaceae. Springer, New York, pp 241–282

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

Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, BeckBunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor Appl Genet 92:213–224

Van der Gaag DJ, Wever G (2005) Conduciveness of different soilless growing media to pythium root and crown rot of cucumber under near-commercial conditions. Eur J Plant Pathol 112:31–41

Wehrhahn C, Allard RW (1965) The detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. Genetics 51:109–119

Wright S (1978) Evolution and the genetics of populations. In: Variability within and among natural populations, vol 4. U Chicago Press, Chicago, IL

Yuan X, Pan J, Cai R, Guan Y, Liu L, Zhang W et al (2008) Genetic mapping and QTL analysis of fruit and flower related traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Euphytica 164:473–491