Theor Appl Genet (2005) 111: 1022–1031
DOI 10.1007/s00122-005-0014-3

ORIGINAL PAPER

K. Loridon · K. McPhee · J. Morin · P. Dubreuil
M. L. Pilet-Nayel · G. Aubert · C. Rameau
A. Baranger · C. Coyne · I. Lejeune-Hènaut · J. Burstin

Microsatellite marker polymorphism and mapping in pea (*Pisum sativum* L.)

Received: 3 June 2005 / Accepted: 9 June 2005 / Published online: 25 August 2005

Abstract This paper aims at providing reliable and cost effective genotyping conditions, level of polymorphism in a range of genotypes and map position of newly developed microsatellite markers in order to promote broad application of these markers as a common set for genetic studies in pea. Optimal PCR conditions were determined for 340 microsatellite markers based on amplification in eight genotypes. Levels of polymorphism were determined for 309 of these markers. Compared to data obtained for other species, levels of polymorphism detected in a panel of eight genotypes were high with a mean number of 3.8 alleles per polymorphic locus and an average PIC value of 0.62, indicating that pea represents a rather polymorphic autogamous species. One of our main objectives was to locate a maximum number of microsatellite markers on the pea genetic map. Data obtained from three different crosses were used to build a composite genetic map of 1,430 cM (Haldane) comprising 239 microsatellite markers. These include 216 anonymous SSRs developed from enriched genomic libraries and 13 SSRs located in genes. The markers are quite evenly distributed throughout the seven linkage groups of the map, with 85% of intervals between the adjacent SSR markers being smaller than 10 cM. There was a good conservation of marker order and linkage group assignment across the three populations. In conclusion, we hope this report will promote wide application of these markers and will allow information obtained by different laboratories worldwide in diverse fields of pea genetics, such as QTL mapping studies and genetic resource surveys, to be easily aligned.

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00122-005-0014-3

Communicated by F. J. Muehlbauer

K. Loridon · P. Dubreuil · G. Aubert · J. Burstin (✉)
INRA URLEG, Domaine d'Epoisses,
21110, Bretenières, France
E-mail: burstin@epoisses.inra.fr
Tel.: +33-3-80693159
Fax: +33-3-80693263

K. McPhee
USDA-ARS, 646434, Pullman, WA 99164-6434, USA

J. Morin · I. Lejeune-Hènaut
INRA SGV, Domaine Brunehaut, 80200 Peronne, France

M. L. Pilet-Nayel · A. Baranger
UMR APBV, 35327, 35653 Le Rheu cedex, France

C. Rameau
INRA SGAP, Route de St-Cyr,
78026 Versailles Cedex, France

C. Coyne
USDA-ARS, WRPIS, 59 Johnson Hall, Pullman,
WA 99164-6402, USA

Introduction

QTL mapping and association studies are becoming widespread approaches to dissect the genetic determinism of many economically important traits in plant breeding. Groups working on different themes develop populations specific to their goals that segregate for different traits of interest. To relate the results obtained in different populations for the same or for different traits, the development of common tools for applying molecular markers becomes important. The availability of highly polymorphic, locus-specific, easily transferable and cost-effective molecular markers distributed throughout the genome is of great value. Because microsatellite markers have all these qualities and are choice markers for many applications, significant effort has been dedicated to their development in various species during the last decade. They are now widely used for investigating genetic diversity among cultivars and genetic resources, for developing genetic maps suitable for QTL detection studies and marker-assisted selection programs. SSR maps now exist for a large range of species (maize: Sharopova et al. 2002; sunflower: Yu et al. 2003; wheat: Somers et al. 2004; bean: Blair et al. 2003; peanut: De Carvalho Moretzsohn et al. 2004;

Fragaria: Sargent et al. 2004; lytchee: Viruel et al. 2004; *Actinidia*: Fraser et al. 2004; cacao: Pugh et al. 2004; soyabean: Song et al. 2004; grape: Adam-Blondon et al. 2004; *Eucalyptus*: Brondani et al. 2002; melon: Ritschel et al. 2004; cassava: Mba et al. 2001; *Prunus*: Dirlewanger et al. 2004; Rose: Yan et al. 2005).

In pea (*Pisum sativum* L.), genetic maps have been developed using different types of markers (isozymes, RFLPs, RAPDs, AFLPs, and trait loci, see Ellis and Poyser 2002 and McPhee 2005 for a review) and a comprehensive consensus linkage map was developed integrating linkage relationships from multiple maps and linkage studies (Weeden et al. 1998; http://hermes.bionet.nsc.ru/pg/30/map.htm, Ellis and Poyser 2002). However, few studies have used microsatellite markers in pea for either mapping studies (Weeden et al. 1998; Pilet-Nayel et al. 2002; Prioul et al. 2004) or diversity assessment (Burstin et al. 2001; Baranger et al. 2004; Tar'an et al. 2005). In this paper we gathered data obtained from three laboratories in order to provide useful information for microsatellite marker use in mapping and diversity studies in pea. Our main objectives were (1) to design efficient multiplexed SSR marker sets for cost-effective genotyping, (2) to evaluate the level of polymorphism of newly developed SSR markers and (3) to map a maximum number of these markers on a composite map.

Materials and methods

Plant material

Two different sets of genotypes were used in order to evaluate the polymorphism of microsatellite markers under study. The first set of genotypes included 'Térèse', K586, a ramified mutant obtained from 'Torsdag', 'Ballet', 'Caméor', 'Champagne', 'Kazar', 'Melrose', and VavD265 (Set 1). The second set of genotypes included Térèse, Torsdag, 'Puget', 'Baccara', 'Capella', DP, CE101, 'Dark Skin Perfection' (DSP), JI296, 90-2079, 90-2131, 552, and PI180693, (Set 2) that were genotyped in a second experiment in an incomplete manner. These genotypes are parents of various segregating populations being developed. Three segregating populations were used to build a composite genetic map. They consist of recombinant inbred lines (RILs) derived by single seed descent from the crosses Térèse × K586 (Population 1 (Pop1), 139 RILs, Laucou et al. 1998), Champagne × Térèse (Population 2 (Pop2), 164 F_8 RILs produced at INRA-Mons), 'Shawnee' × 'Bohatyr' (Population 3 (Pop3), 187 F_7 RILs produced at USDA-ARS, Pullman, WA, USA). Shawnee and Bohatyr were not included in either Set 1 or Set 2 for evaluation of polymorphism. Thus, a marker declared monomorphic based on the polymorphism survey of Set 1 and Set 2 could be polymorphic between Shawnee and Bohatyr and, therefore, mapped in Pop3. We also report mapping data obtained for the RIL populations derived from DP × JI296 (Prioul et al. 2004) and Puget × 90-2079 (Pilet-Nayel et al. 2002). The genotypes Caméor, Ballet and VavD265 were used to develop marker triplex reactions. Most of these genotypes (Ballet, Caméor, Champagne, Kazar, Melrose, VavD265 (=K4926), Térèse, Torsdag, Puget, Baccara, DP, CE101, JI296, Bohatyr) were described in Baranger et al. 2004. Capella is a feed pea cultivar, 'Dark Skin Perfection' is a wrinkled garden pea cultivar, 90-2079 and 90-2131 are two round-seeded pea germplasm accessions from USDA-ARS, 552 is a garden pea accession from U. Wisconsin, and PI180693 a round-seeded accession from Frankfürt. Shawnee is a smooth-seeded, yellow cotyledon dry pea cultivar developed at the USDA-ARS, Pullman, WA, USA (Muehlbauer 2002).

Total DNA was extracted from leaf tissue following different methods: a modified method following Dellaporta et al. (1983) for polymorphism Set 1, and mapping population 1, CTAB method from Doyle and Doyle (1990) for polymorphism Set 2 and mapping populations 2 and 3, the DNeasy kit (Qiagen) method recommended by the manufacturer for multiplex set up.

Microsatellite marker genotyping

Microsatellite markers were obtained from two sources: markers developed by Burstin et al. (2001) by searching publicly available sequence databases for microsatellite motifs, and markers obtained from a Pea Microsatellite Consortium set up by Agrogène Inc., Moissy-Cramayel, France. Microsatellite clones were isolated by Agrogène from four pea libraries (genotype Kelvedon Wonder) enriched for $(CA)_n$, $(GA)_n$, $(AAT)_n$ and $(TAGA)_n$ repeated motifs. The clones were sequenced and primer sets were designed. A total of 434 primer pairs were collected in a database available to members of the Pea Microsatellite Consortium (PMC). In this paper, the name of the markers, as given by Agrogène was shortened (PSMPS at the beginning of the name of each marker was dropped).

In a first step, optimal PCR conditions were determined for microsatellite markers in three INRA laboratories (Dijon, Mons and Rennes). PCR conditions for markers that were subsequently genotyped in polymorphism sets and/or mapping populations 1 and 2 are presented in Table S1 (ESM), including annealing temperatures (T_m), optimal $MgCl_2$ concentration as well as their primer sequences, thanks to the kind authorization for publication from the PMC members. PCR reactions were performed with 20–40 ng of template DNA, 10 mM Tris–HCl pH 9.0, 50 mM KCl, 0.1% Triton x 100, 1.5–2.5 mM $MgCl_2$, 0.2 mM dNTP, 0.2 μM of each primer and 0.5–1 U *Taq* DNA polymerase (Sigma, St Louis, MI, USA) in a total volume of 25 μl, overlaid by paraffin oil. Amplifications were performed in a PTC-200 thermocycler (MJ Research, Waltham, MA, USA), Perkin Elmer GeneAmp 9600

PCR system (Perkin Elmer, Norwalk, CT, USA) or a Mastercycler (Eppendorf, Hamburg, Germany) using the following profile: 3 min at 94°C, 35–45 cycles of 30 s at 94°C, 45–60 s at the required T_m (Table S1), 45–60 s at 72°C and a final extension step of 5 min at 72°C. Marker analysis of Set 1 and Pop1 carried out at INRA Dijon, was accomplished by labeling the reverse primer with [γ^{33}P]ATP using T4 polynucleotide kinase (Promega, Madison, WI, USA) according to the manufacturer's instructions. PCR products were loaded on denaturing polyacrylamide gels (5.7% acrylamide/0.3% bis-acrylamide, 7 M urea, 1x TBE buffer), run 2–5 h at 70 W in a S2 system (Whatmann International Ltd., Maidstone, England) transferred onto 3MM paper (Whatmann International Ltd., Maidstone, England) and dried 1 h at 70°C under vacuum. Radio-labeled fragments were visualized by auto-radiography after 1–5 days exposure with Super RX films (Fuji Photo Film Co. Ltd., Tokyo, Japan); PCR products for polymorphism assessment of set 2 carried out at INRA Rennes were loaded on denaturing polyacrylamide gels and visualized using silver staining as described in Burstin et al. (2001). Genotyping of Pop2, carried out at INRA Mons, involved loading PCR products on 4% Metaphor agarose gels (BMA, Rockland, MA, USA) in 1x TBE buffer, run for 3–5 h at 140 V and stained with ethidium bromide.

PCR amplifications for Pop3, performed at USDA-ARS, Pullman, WA, USA, were carried out in a 20 μl reaction volume comprised of 2 U Promega *Taq* polymerase (Promega, Madison, WI#M1861), 1x Promega buffer (Promega, Madison, WI#M1881), 1.5 mM $MgCl_2$, 10 pmol each primer, 0.2 mM each dNTP and 40 ng template. Template DNA was denatured at 96°C for 2 min prior to amplification for 35 cycles (96°C 20 s, 51 or 61°C depending on specific primer 50 s, 72°C 50 s) on a Gene Amp 9700 thermocycler (Applied Biosystems, Foster City, CA, USA). Final extension was 7 min and products were held at 4°C. The forward primer of each pair was internally labeled with either IRDye 700 or IRDye 800 for the Li-Cor by MWG (High Point, NC, USA). PCR products were viewed on a 6.5% Licor KB+ gel using a Li-Cor 4300 DNA Analyzer (Li-Cor Inc., Lincoln, NE, USA). Gels were prerun for 10 min at 1,000 V, 40 mA, 40 W and 45°C. Samples were run for 2.5 h at 1,500 V, 40 mA, 40 W and 45°C. Gel images were then analyzed and band sizes determined using the Gene ImagIR v. 4.03 software program (Li-Cor, Inc., Lincoln, NE, USA).

Other marker systems

RAPD genotyping in Pop1, Pop2 and Pop3 was performed as described in Laucou et al. (1998). For mapping population 3, isozyme markers were genotyped on zymograms following starch gel electrophoresis for the following enzymes: leucine aminopeptidase (LAP, EC 3.4.11.1), 6-phosphogluconate dehydrogenase (6-PGD, EC 1.1.1.44) and shikimate dehydrogenase (SKDH, EC 1.1.1.25). Plant extracts were collected from 4- to 6-week-old plants and electrophoresis was conducted according to methods described in Muelhbauer et al. (1989).

Multiplex marker set up

A set of polymorphic markers mapped in at least one of the five populations described in Plant material, was chosen to provide good coverage of the seven linkage groups of the pea genome. Preliminary screening was performed to set up standardized PCR conditions and estimate the allele size range. In a second step, loci that could be satisfactorily amplified in the same standard PCR conditions, and whose banding patterns did not overlap were used for multiplexing. As allele size ranges were very similar among loci, only sets of three loci were considered. Triplexed marker sets were checked for the formation of primer dimers using the FastPCR software (version 2.9.36, http://www.biocenter.helsinki.fi) and those with expected primer dimers having high T_m were not assayed. The reaction buffer included 10–20 ng of genomic DNA, 0.2 mM of each dNTP, 10 mM Tris–HCl, 50 mM KCl, 0.1% Triton X 100, 1.5–2.5 mM $MgCl_2$, 0.27 μM locus-specific reverse primer, 0.07 μM locus-specific forward primer with M13 tail (5′-cacgacgttgtaaaacgac-3′) at the 5′ end, 0.20 μM fluorescently labeled (IRD700 or 800) M13 sequence-specific primer, and 0.5–1 U of *Taq* polymerase. Loci were amplified in reaction buffer at a final $MgCl_2$ concentration of either 1.5 or 2.5 mM using the following cycling program: ten cycles of 30 s at 94°C, 45 s at 56°C or 61°C depending on the T_m of the primer combination, 60 s at 72°C, and 25 cycles of 30 s at 94°C, 45 s at 51°C ($\sim T_m$ of the M13 tail), and 60 s at 72°C followed by a final extension of 10 min at 72°C (Schuelke 2000). Fluorescently labeled PCR products from multiplex amplification with IRD700 and IRD800 were bulked in equal proportion before loading and electrophoresed onto 6.5% or 5.0% denaturing polyacrylamide gels in a Licor IR4200 DNA sequencer (Li-Cor Inc., Lincoln, NE, USA).

Characterization of the level of microsatellite marker polymorphism

The level of polymorphism revealed by the different microsatellite markers was estimated following two approaches. Polymorphism information content (PIC) was calculated from the complete data set obtained for the eight genotypes of Set 1, using the following equation: $PIC_j = 1-\sum_i pij^2$, where *pij* is the frequency of the *i*th allele at locus *j*. The PIC estimates the number of polymorphic pairs among all possible pairs in a population. Total number of pairs of genotypes compared and the total number of polymorphic pairs

among them for each marker were calculated for the incomplete data set from Set 2. A polymorphism index, which is an estimate of PIC in Set 2, was derived by dividing the number of polymorphic pairs by the number of pairs compared. In order to see if the structure of the microsatellite marker had an effect on the level of polymorphism we performed an ANOVA (proc GLM, SAS Institute, 1988) with the PIC and the polymorphism index as independent variables and the repeated motif and the type of repetition ('simple' when only one motif was repeated and 'complex' when compound motifs were repeated) as dependent variables.

Construction of the composite microsatellite marker map

A first map was constructed based on Population 1. Microsatellite markers described in Table S1 were added to other molecular markers (RAPD, RFLP) from a previously published map (Laucou et al. 1998). All markers for which the number of missing data exceeded 25 were excluded from the linkage analysis. The remaining markers were tested for a 1:1 segregation ratio using chi-square tests ($\alpha = 0.01$). Markers that did not fit the expected Mendelian ratios were excluded from the linkage analysis. Linkage groups were built using the 'group', 'assign', 'build', and 'ripple' commands of MAPMAKER/EXP version 3.0b (Lander et al. 1987; Lincoln et al. 1992). Marker order was refined using the 'annealing' algorithms from the CarthaGene software (Schiex et al. 2001). The composite map was built using the map information obtained from Population 1. The 'group' command was run separately for the 3 datasets (LOD = 5). A set of microsatellite and morphological markers distributed all over the genome and genotyped in at least two of the three populations was chosen as anchor markers. Data obtained for the three populations were merged and all markers assigned to linkage groups using the 'assign' command of Mapmaker at a LOD score 4, and ordered using the 'order' command. Markers that were not clearly assigned at first and/or that displayed segregation distortion were added on this framework map using the 'try' command only when they did not change the order of neighboring markers. When a marker showed significant segregation distortion in one mapping population and not in another, only data from the population that did not show a distorted segregation were used to construct the composite map. Finally, marker order was refined using the 'annealing' algorithms from the CarthaGene software. Both Haldane and Kosambi functions were used to calculate centiMorgan (cM) distances in order to compare the length of the map to previously published maps. MapChart 2.1 was used to draw the maps on Figs. S1 and 3 (Voorrips 2002, Plant Research International, Wageningen, The Netherlands).

Results

Characterization of the microsatellite markers

Four hundred thirty-four microsatellite primer pairs were developed by Agrogène for the Pea Microsatellite Consortium and optimal PCR conditions were determined for 340 of these markers. A quality score, A, B, C or D, was given to each marker (Table S1, Electronic supplementary material). Score A was given to 151 markers (45%) of good quality displaying clear single band patterns, score B to 49 medium quality markers (14%) displaying smears and fainter bands difficult to read, score C to 76 markers (22%) displaying multiband profiles and score D to 64 markers (19%) producing smears, non reproducible bands or no product and discarded after several attempts to improve PCR conditions. Among the 67 microsatellite markers obtained from the search of public databases, 43 had score A (64%), 4 score B (6%), 6 score C (9%) and 14 score D (21%). All microsatellite markers with score A, B, and C, are listed in Table S1 with their primer sequences, optimal PCR conditions and quality score. Twenty-four triplexed PCR marker sets were set up to assess multiplexing markers for more cost-effective genotyping (Table 1, Fig. 1). Using two fluorescent dyes on an automated sequencer, six markers could be genotyped at one time.

In order to know if any Pea Microsatellite Consortium markers were located in known sequences public databases were searched with pea sequences. Numerous

Table 1 Summary and amplification conditions for the triplexed primer sets including annealing temperature (°C) and $MgCl_2$ (mM) concentration

Set	T_m (°C)	$MgCl_2$ (mM)	Marker triplex
1	56	2.5	AB72, AB91, AB92
2	61	1.5	AB122, AB64, AD73
3	56	2.5	AA155, AB111, AD68
4	56	1.5	AB113, AD148, AD270
5	56	1.5	AA372.1, AB20, AD160
6	56	1.5	AA163.2, AA99, AB140
7	56	1.5	AB90, AD175, D21
8	56	1.5	AB53, AD174, AA206
9	56	1.5	AA90, AB136, C20
10	56	1.5	AA355, AB28, AB47
11	56	1.5	AA107, AA81, AD158
12	61	1.5	AB128, AD147, AD79
13	61	1.5	AA200, AB101, AB68
14	56	1.5	AA19, AC74, AC76b
15	61	1.5	AA335, AA5, AA5
16	56	1.5	AA285, AA317, AC22
17	61	1.5	AB112, AD186, B14
18	61	1.5	AB109, AB133, AD83
19	56	1.5	AA399, AA456, AD57
20	61	1.5	AA122, AA92, AB100
21	61	1.5	AA18, AD141, AD171
22	61	1.5	AA303, AA497, AD56
23	61	1.5	AA103, AB23, AB77
24	61	2.5	A9, AA1, AB88

markers had similarities to *Psat* repetitive sequences identified by Neumann et al. (2001). Three markers had similarities to a sequence Ty3-type retrotransposon (AC40, AC50.1, AC96), 11 with the gibberellin C20-oxidase gene (AA61, AA372.1, AA415, AD59, AB44, AB98, AB104, AB127, AB152, AB200, C19), 5 with the chalcone synthase gene (AA415, AB98, AB152, AB200, C19), 1 with an Hsp70 gene (AA135), 2 with an IAA4/5 gene (AA47, AA250), 1 with a root expressed trypsin inhibitor gene (AA260), and 1 with a ribosomal protein gene (AA170). Only a few of these markers were mapped in our study: AA61 mapped at two loci on LGIV and VII, AB44 on two loci on LGIV and III, AB104 on LGIII, AA135 on LGVII. AD59, AA47, AA250 were monomorphic in the mapping populations.

Microsatellite marker polymorphism

Length polymorphism was first assessed among a set of eight genotypes (Set 1). Of 257 markers analysed, 188 (73%) were polymorphic and the number of alleles ranged from 2 to 7 with the mean number of alleles per marker was 3.8. The markers displaying 7 alleles were AA355, AA5, AB36, AB53, AD146, AD147, AD159. The polymorphism information content (PIC) varied from 0.22 to 0.85 and averaged 0.63. Fifty-two markers (20%) displayed null alleles for one to six genotypes and were verified at least twice. Length polymorphism was also assessed among a second set of 13 genotypes (Set 2) for 279 markers and 224 (80%) were polymorphic. The Polymorphism Index calculated as the ratio of polymorphic pairs among the second set of 13 genotypes varied from 0.04 to 1 and averaged 0.62. The PIC calculated for the first set of genotypes and the polymorphism index which estimated PIC in the second set of genotypes were significantly correlated ($R^2 = 0.57$, Fig. 2). In total, 309 markers were assessed for polymorphism and 236 (76%) were polymorphic either in Set 1 or in Set 2.

Trying to relate the level of polymorphism observed for the different markers with the structure of the repeated sequence, an analysis of variance was performed for PIC and the polymorphism index using the repeated motif and the type of repetition ('simple' or 'complex') as factors. Markers containing complex repetitions were slightly more polymorphic than markers containing simple repetitions ($\alpha = 1\%$, $R^2 = 0.03$) as measured by the PIC and the polymorphism index.

Construction of the 'composite' microsatellite pea map

The initial map was constructed for Population 1 (139 RILs derived from 'Térèse' × K586) based on genotyping data for 255 markers (83 SSR markers from this study, and 160 RAPD, 10 RFLP, 2 morphological markers from Laucou et al. 1998). Markers lacking a significant number of data points or deviating from the expected Mendelian ratios were excluded from the linkage analysis. The final genetic map (Fig. S1, Elec-

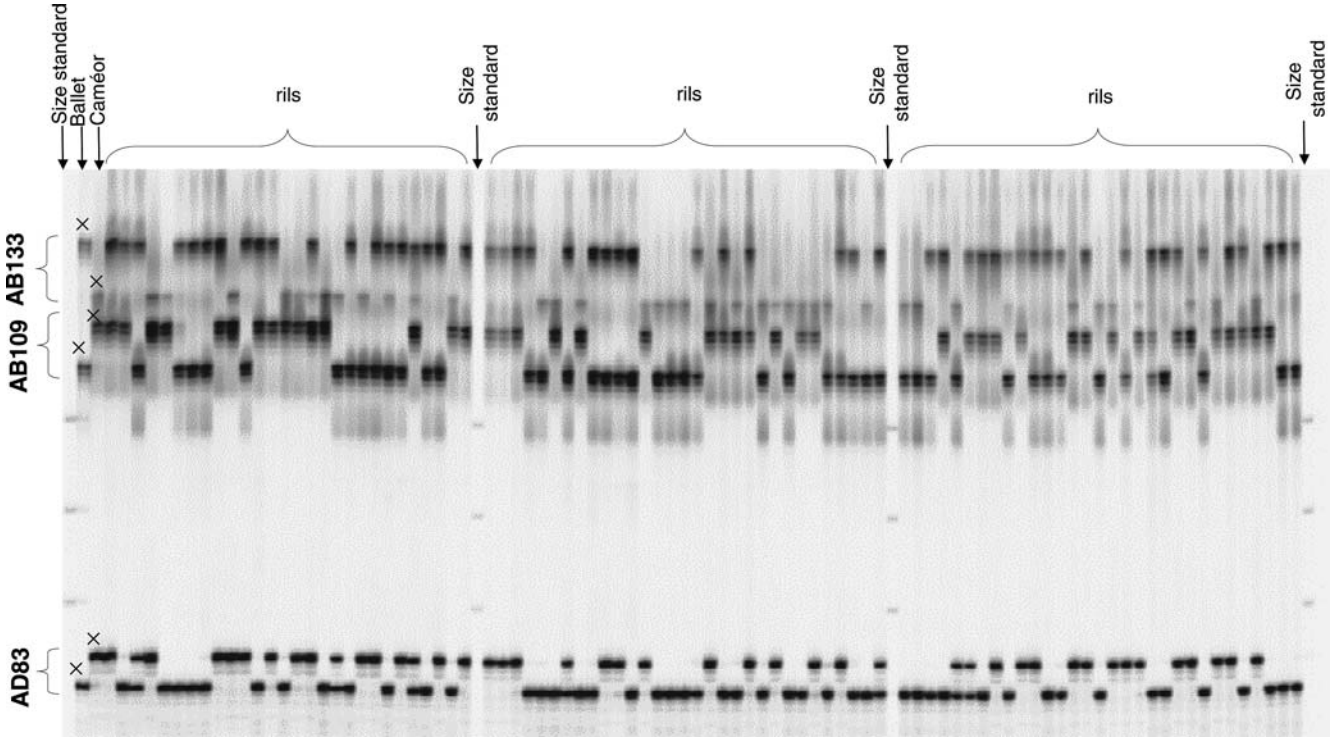

Fig. 1 Segregation pattern of three multiplexed SSR markers (Triplex 18; see Table 1, size standard is 50–350 size standard from LICOR)

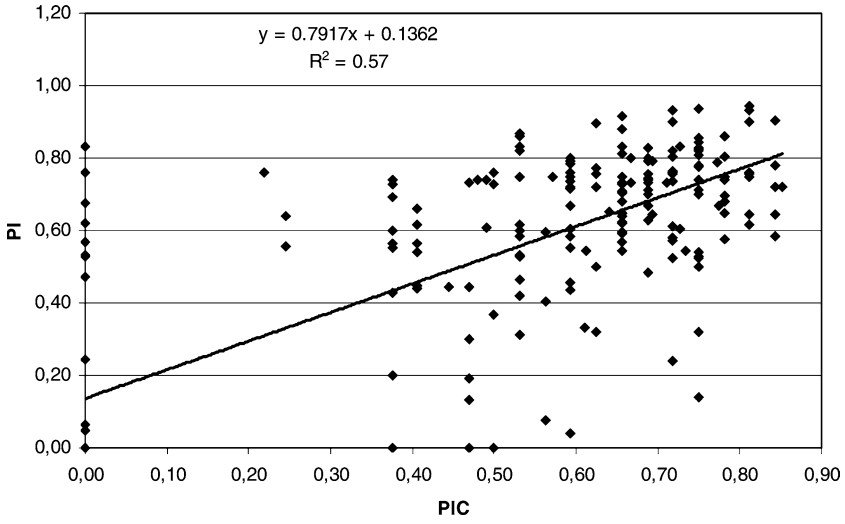

Fig. 2 Relationship between the polymorphism information content (PIC) calculated for 8 genotypes (Set 1) and the polymorphism index calculated for 13 genotypes (Set 2)

tronic supplementary material) comprised 242 loci, including 150 RAPD, 82 SSR, 10 RFLP and 2 morphological markers distributed over nine linkage groups (LOD = 5). Two groups corresponded to LGI and two groups to LGIII as reported in Laucou et al. (1998). This map was used as a framework for construction of the 'composite' map.

The 'composite' genetic map reported here was based on three segregating populations (in total, 490 RILs and 579 markers): Population 1 described above, Population 2 (164 RILs derived from Térèse × Champagne and characterised for 189 markers—including 97 SSR markers, 87 RAPDs and 5 morphological markers), and Population 3 (187 RILs derived from Shawnee × Bohatyr and characterized for 302 markers including 186 SSR markers, 107 RAPDs, 5 isozymes and 4 morphological markers). A total of 118 markers were common to at least two of the three mapping populations: 31 markers including 28 SSRs were genotyped in all three populations; 52 markers including 21 SSRs were genotyped both in Pop1 and 2, 17 markers including 7 SSRs in Pop1 and 3, and 18 markers including 17 SSRs in Pop2 and 3. Eleven of these 'bridge' markers were assigned to LGI, 19 to LGII, 25 to LGIII, 15 to LGIV, 14 to LGV, 14 to LGVI and 20 to LGVII. Individual maps for the three populations are presented in Fig. S1 (Electronic supplementary material). Marker order was conserved between the populations with few exceptions. P11_600 and AA121 on LG1 were reversed in Pop1 compared to Pop2. The segment between E12_490 and AA504 on LGII appeared inverted in Pop2 compared to Pop1 and 3. Map position of AD73 on LGIII was different in Pop3 compared to Pop1 and Pop2, and the order of closely linked *Le*, AB64 and A6 was different on all 3 maps. Position of AD60 on LGVI was not the same in Pop2 and Pop3; however, this marker is significantly distorted in Pop2. All other orders were conserved between Pop1 and Pop2, Pop1 and Pop3, and Pop2 and Pop3.

Among the well-distributed 'bridge' SSR and morphological markers we defined a set of 'anchor' markers for each of the seven larger linkage groups obtained for Pop1: *Af*, D21, AB28, AD147, AA474, for LGI; AA205, AB149, AA372.1, AD148, AD186.1, AA474 for LGII; AA107, AA278, AD73, AB140, AA355, AD270, AB53, *Le*, AB64 for LGIII; AD61, AD171, AB45, AA122, AA92, AD186.2 for LGIV; AA460, AC58, AD55.2, AD79, AA99, AB47, AA81 for LGV; AA103, AB20, AA335, AD59, AD51 for LGVI; and AA456, AA90, AA19, AD146, AA317, AD237 for LGVII. Using the 'assign' command of Mapmaker, two zones presented some conflicting results with different mapping populations. The first zone between G10_980 and AA467.1 on LGIII comprised some markers which were assigned to LGV when genotyped in Pop1 and 2 and some to LGVII in Pop3. Following Laucou et al. (1998), we assigned this segment to LGIII. The second zone between UBC256 and AD230-135 on LGVI comprised some markers which were assigned to LGVI in Pop1 and 2 and some to LGII in Pop3. We assigned this region to LGVI. These results suggest that these two chromosomal regions may have undergone rearrangements in the different parental genotypes. After this assignment, the map was built. We did not try to integrate all markers but included all SSR markers and 'bridge' markers. The composite genetic map (Fig. 3) covers 1,430 cM (Haldane) or 1,350 cM (Kosambi) and comprises 462 loci distributed over 7 linkage groups, including 216 anonymous SSRs and 13 SSRs belonging to genes. The mean spacing between adjacent markers is 3.1 and 6.2 cM between SSR markers. Ninety-five percent of intervals between adjacent markers, and 85% percent of intervals between adjacent SSR markers were smaller than 10 cM. Five intervals between adjacent markers were between 20 and 30 cM (1%) and nine intervals between adjacent SSR markers were between 20 and 38 cM (4%).

Among the 581 markers, 7 markers showed significant segregation distortion ($\alpha = 5\%$) including 4 SSRs and 3 RFLPs in Pop1, 44 markers including 23 SSRs and 21 RAPDs in Pop2, and 55 markers including 22 SSRs, 32 RAPDs in Pop 3. Fifty-six of these markers were mapped (Fig. 3). The majority of distorted loci

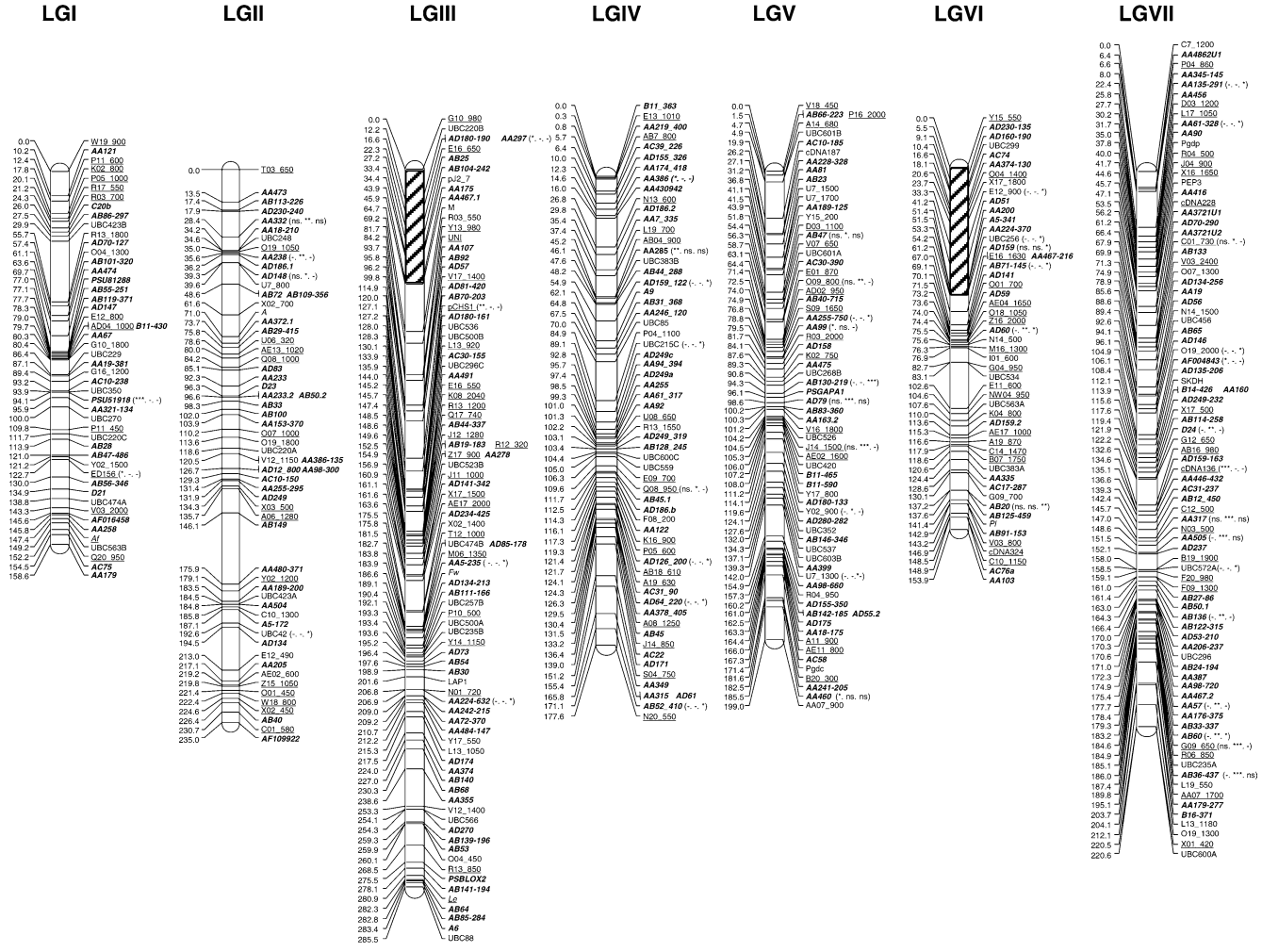

Fig. 3 Composite microsatellite marker map of pea. Microsatellite markers are indicated in *bold*, markers that are common to the map published by Laucou et al. (1998) are *underlined*. For markers that show a distorted segregation, the levels of the distortion as revealed by a chi-square test in the three mapping populations are indicated within brackets (Pop1, Pop2, Pop3), with * = significant at 5%, ** = significant at 1%, *** = significant at 0.1%, *ns* = not significant and - not mapped in this population. The chromosomal regions that were ambiguously assigned are in grey. Distances are in centiMorgan (Haldane)

mapped on LGIV (20%), LGV (18%), and LGVII (30%). Some regions contained markers that were significantly distorted in two or three populations. Three markers in Pop2, one in Pop1, and two in Pop3 were distorted within the region from AA163.2 to AA255–750 on LGV; two markers in Pop2 and three in Pop3 were distorted in the region from AD60 to E12_900 on LGVI; and two markers in Pop1, three in Pop2, and two in Pop3 were distorted within the region O19_2000 to UBC572A on LGVII (Fig. 3).

Discussion

This paper aims at providing reliable and cost-effective genotyping conditions, level of polymorphism in a range of genotypes and map position for newly developed microsatellite markers in order to promote broad application of these markers as a common set for genetic studies in pea. The high transferability of microsatellite markers allowed us to compile information obtained from different laboratories using different genotyping procedures.

Characterization and levels of polymorphism of SSR markers

All together, 309 markers were tested for level of polymorphism, 257 in a set of 8 genotypes (Set 1) and 279 in a larger, but incomplete, set of 13 genotypes (Set 2). In total, 235 markers were polymorphic either within Set 1 or Set 2. In Set 1, 73% of the markers were polymorphic, with up to seven alleles identified among the eight genotypes for seven markers, and a mean of 3.8 alleles per polymorphic marker. In Set 2, 80% of markers were polymorphic. The level of polymorphism as estimated by the PIC within the first set of genotypes or by a similar polymorphism index within the second set of genotypes were significantly correlated (R^2 = 0.57) and averaged around 0.62. These indices estimate markers'

level of polymorphism and allow the most useful markers to be chosen for genetic studies in pea.

The level of polymorphism reported here is consistent with data obtained in Burstin et al. (2001) where 3.6 alleles per polymorphic marker were observed for 31 markers derived from gene sequences, even though the panel of 12 genotypes used included some more exotic accessions than the panel of eight genotypes of Set 1. These levels of polymorphism, when compared to data obtained in other studies for other species, indicate that pea represents a rather polymorphic autogamous species. In barley, 15 SSRs revealed a mean number of 3.5 alleles per polymorphic marker and an average PIC value of 0.45 among 26 accessions (Hamza et al. 2004); in sunflower, 170 SSRs revealed 3.5 alleles per locus with a mean PIC of 0.55 in 16 accessions (Paniego et al. 2002) and 300 SSRs revealed a mean number of 3.6 alleles per polymorphic marker and an average heterozygosity of 0.53 among 24 accessions (Tang et al. 2003); in tomato, 65 polymorphic SSRs revealed 2.7 alleles per locus with a mean PIC around 0.35 among 19 accessions (He et al. 2003); in *Oryza sativa* L., 22 SSRs revealed an average PIC value of 0.73 among 13 diverse accessions (Coburn et al. 2002).

Correspondence of the composite microsatellite map with previous maps

In many species, 'composite' maps integrate data obtained from different crosses in order to take advantage of different sources of polymorphism and of the mapping efforts of different research groups to accumulate a maximum number of markers on a single map (Causse et al. 1996; Weeden et al. 1998; Song et al. 2004; Somers et al. 2004; Adam-Blondon et al. 2004). As one of our main objectives was to locate a maximum number of microsatellite markers on the pea genetic map, we gathered data obtained from three different crosses. The composite genetic map (Fig. 3) covers 1,430 cM (Haldane) or 1,350 cM (Kosambi) and comprises 229 microsatellite markers, including 216 anonymous SSRs developed from enriched genomic libraries and 13 SSRs located in genes (Burstin et al. 2001). The markers are evenly distributed throughout the seven linkage groups of the map. Eighty-five percent of intervals between adjacent SSR markers were smaller than 10 cM and nine intervals were between 20 and 38 cM (4%). The size of the map is similar to that of previously published maps (Ellis et al. 1992; Gilpin et al. 1997; Pilet-Nayel et al. 2002; Prioul et al. 2004; Tar'an et al. 2003). It also has roughly the same size as the map published by Laucou et al. (1998, 1,139 cM), considering that this latter map lacks the segments between markers AD04_1000 and R03_700 (55.4 cM) on linkage groups I and between markers E16_650 and Y13_980 (59.4 cM) on group III. However, it is longer than the *Pisum* consensus linkage map (Weeden et al. 1998) and exceeds the size predicted from chiasma counts (Hall et al. 1997).

Ellis et al. (1992) emphasized the danger, in pea, of compiling a single map from different crosses, due to frequent rearrangements occurring in the pea genome. There was a good conservation of marker order and marker linkage group assignment in the different populations considered, showing that data generated independently among three different laboratories agree well (Fig. S1). The composite map presented on Fig. 3 was the best model that we obtained by integrating data from Pop1, 2 and 3. However some discrepancies in marker assignments or orders, as for LGII in Pop2 as compared to Pop1 and 3 (Fig. S1), suggested possible rearrangements in these regions for one of the parents of the crosses. We also checked the congruence of marker locations with previously published maps. One hundred and twenty-nine markers, mainly RAPDs, from the map published by Laucou et al. (1998) are located on the composite map (Fig. 3): most of the marker orders have been conserved except a few rearrangements between close markers. Out of the 71 microsatellite markers mapped by Prioul et al. (2004), 47 were also mapped on the composite map. Again, most of the marker orders have been conserved except a few rearrangements between close markers and except marker AB64 which is located on LGIII, near *Le*, on the composite map and which is located on LGIV in Prioul et al. (2004). More generally, we checked that the assignment to linkage groups was consistent between the present composite map and two other maps obtained for populations of recombinant inbred lines derived from DP × JI296 (Prioul et al. 2004), and Puget × 90-2079 (Pilet-Nayel et al. 2002) (Table S1, Pilet-Nayel and Baranger, data not shown). Of 72 markers located on the composite map and either of these two latter populations, 64 were located on identical linkage groups and 8 to different groups in the different linkage maps. Different causes may explain different assignments in different populations: different bands may be mapped in the different populations for multi-band markers and/or when contrasting methods with different resolutions were used to separate products. These issues may be solved using automated sequencer technology to determine the precise band size. Another source of conflicting assignment is distorted segregation of a marker within a population. And finally, true chromosomal rearrangements can lead to differences in genetic maps.

More generally, our map can be related to historical maps. It is connected to the Ellis et al. (1992) map by 30 markers (22 RAPDs, 2 morphological, and 6 RFLP) as reported by Laucou et al. (1998). It is also connected to the pea consensus map (Weeden et al. 1998; Ellis and Poyser 2002) but only by few common morphological and isozyme markers : *Af* located on LG1, *A* on LGII, *M, UNI, Fw*, LAP1, and *Le* on LGIII, none on LGIV, Pgdc on LGV, *Pl* on LGVI and Pgdp, Pep3, SKDH on LGVII. On LGIII and LGVII, except for LAP1, all the other markers map in the same order. In the present study *Fw* and LAP1 were mapped simultaneously in Pop3 and since the number of markers used was quite

large, we expect their relative positions to be accurate. As for previous estimates, the *Fw* and LAP1 genes were approximately placed on the consensus map of Weeden et al. (1998) based on deduction and congruence from different maps. In Grajal-Martin and Muehlbauer (2002), seven populations were characterised simultaneously for *Fw* and morphological and isozyme markers including LAP1, but only four linked markers were used to establish linkage relationships in the region of *Fw*. Additional marker data would be necessary to confirm the relative positions of *Fw* and LAP1 in these populations. An important challenge now will be to relate our microsatellite composite map with the consensus map. Indeed, trait genetic loci present on the consensus map can provide interesting information to understand the genetic polymorphisms underlying QTLs. The next step could then be to map some of the proposed anchor microsatellites on the JI1794×Slow primary map (Weeden 1998).

Conclusion

Microsatellite markers are choice markers for genetic studies compared to other markers such as RAPD, AFLP and SSAP markers derived from retroelements (Ellis et al. 2002) because they are codominant and easily transferable to other populations because of their high level of polymorphism. In this study, more than 300 SSR markers were tested. Among the 194 that displayed a good quality and clear banding pattern, 168 were polymorphic, either in the first or second set of genotypes, 106 were mapped, and 45 were mapped in more than one population. We hope that this report will promote wide use of these markers and will allow information obtained from different laboratories around the world in different fields of pea genetics such as future QTL mapping and genetic resources surveys to be easily related to each other. This would open avenues for (1) the development of a more dense genetic map in pea with the microsatellite markers serving as anchor markers for mapping genes, (2) QTL validation in different genetic backgrounds, (3) a coordinated assessment and management of pea genetic resources worldwide, and finally (4) their use in marker assisted selection programs.

Acknowledgements We are very thankful to F. Jacquin, G. Deniot, A.L. Lainé, M. Nicolas, D. Yost, I. Legoff, and AM Jacob for their tireless dedication and precious technical assistance for the microsatellite genotyping. We acknowledge the significant role of K. Edwards for the generation of enriched libraries and of Agrogène (Eurofins Scientific Analysis, BP42301, 44323 Nantes cedex 3, France, http://www.eurofins.fr) for the development of markers from these libraries. We warmly thank all the partners of the Pea Microsatellite Consortium who have allowed the publication of the primer sequences that they own (Axel Toft, Danisco, DLF Trifolium, Florimond-Desprez, Génoplante-Valor, John Innes Centre, Lochow-Petkus, Momont, Nickerson, Procosem, Seminis, Serasem, Svalof-Weibull, Unilever, Unisigma, U. Saskatchewan, U. Melbourne). The mapping of microsatellite markers on Population 1 was funded by the Genoplante Program GOP-Pea-A. We specially thank the Genoplante Pea coordinators, B. Desprez and G. Duc. Thanks to R. Thompson for his suggestions on the manuscript.

References

Adam-Blondon AF, Roux C, Claux D, Butterlin G, Merdinoglu D and This P (2004) Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics. Theor Appl Genet 109:1017–1027

Baranger A, Aubert G, Arnau G, Lainé AL, Deniot G, Potier J,Weinachter C, Lejeune-Hénaut I, Lallemand J, Burstin J (2004) Genetic diversity within *Pisum sativum* using protein- and PCR-based markers. Theor Appl Genet 108:1309–1321

Blair MW, Pedraza F, Buendia HF, Gaitán-Solís E, Beebe SE, Gepts P, Tohme J (2003) Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 107:1362–1374

Brondani RPV, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. Mol Genet Genomics 267:338–347

Burstin J, Deniot G, Potier J, Weinachter C, Aubert G, Baranger A. (2001) Microsatellite polymorphism in *Pisum sativum*. Plant Breed 120:311–317

Causse M, Santoni S, Damerval C, Maurice A, Charcosset A, Deatrick J, de Vienne D (1996) A composite map of expressed sequences in maize. Genome 39:418–432

Coburn JR, Temnykh SV, Paul EM, McCouch SR (2002) Design and Application of microsatellite marker panels for semi-automated genotyping of rice (*Oryza sativa* L.). Crop Sci 42:2092–2099

De Carvalho Moretzsohn M, Hopkins MS, Mitchell SE, Kresovich S, Valls JFM, Ferreira ME (2004) Genetic diversity of peanut (*Arachis hypogea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. BMC Plant Biol 4:11

Dellaporta SL, Woods J, Hicks JB (1983) A plant DNA mini preparation II. Plant Mol Biol 1:9–21

Dirlewanger E, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Voisin R, Poizat C, Lafargue B, Baron O, Laigret F, Kleinhentz M, Arús P, Esmenjaud D (2004), Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid—location of root-knot nematode resistance genes. Theor Appl Genet 109:827–832

Doyle JL, Doyle JJ (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Ellis THN, Poyser SJ (2002) An integrated and comparative view of pea genetic and cytogenetic maps. New Phytologist 153:17–25

Ellis THN, Turner L, Hellens RP, Lee D, Harker CL, Enard C, Domoney C, Davies DR (1992) Linkage maps in pea. Genetics 130:649–663

Ellis THN, Poyser SJ, Knox MR, Vershinin AV, Ambrose MJ (1998) Polymorphism of insertion sites of Ty1-copia class retrotransposons and its use for linkage and diversity analysis in pea. Mol Gen Genet 260:9–19

Fraser LG, Harvey CF, Crowhurst RN, De Silva HN (2004) EST-derived microsatellites from *Actinidia* species and their potential for mapping. Theor Appl Genet 108:1010–1016

Gilpin BF, McCallum JA, Frew TJ, Timmerman-Vaughan GM (1997) A linkage map of the pea (*Pisum sativum* L.) genome containing cloned sequences of known function and expressed sequence tags (ESTs). Theor Appl Genet 95:1289–1299

Grajal-Martin MJ and Muehlbauer FJ (2002) Genomic location of the *Fw* gene for resistance to fusarium wilt race 1 in peas. J Hered 93:291–293

Hall KJ, Parker JS, Ellis THN (1997) The relationship between genetic and cytogenetic maps of pea. I. Standard and translocation karyotypes. Genome 40:744–754

Hamza S, Hamida WB, Rebaï A, Harrabi M (2004) SSR-based genetic diversity assessment among Tunisian winter barley and relationship with morphological traits. Euphytica 135:107–118

He C, Poysa V, Yu K (2003) Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. Theor Appl Genet 106:363–373

Julier B, Flajoulot S, Barre P, Cardinet G, Santoni S, Huguet T, Huyghe C (2003) Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biol 3:9 (http://www.biomecentral.com/1471-2229/3/9)

Lander ES, Green P, Abrahamson J, Barlow A, Daly M, Lincoln S, Newberg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Laucou V, Haurogné K, Ellis N, Rameau C (1998) Genetic mapping in pea. 1. RAPD-based genetic linkage map of *Pisum sativum*. Theor Appl Genet 97:905–915

Lincoln S, Daly M, Lander ES (1992) Constructing genetic maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, 3rd edn. Whitehouse Technical Institute, Cambridge

Mba REC, Stephenson P, Edwards K, Melzer S, Nkumbira J, Gullberg U, Apel K, Gale M, Tohme J, Fregene M (2001) Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. Theor Appl Genet 102:21–31

Mc Phee KE (2005) Pea. In: Kole C (ed) Genome mapping and molecular Breeding, vol III: Pulse and Tuber Crops. Science Publishers Inc., Enfield, (Accepted)

Muehlbauer FJ (2002) Registration of 'Shawnee' yellow dry pea. Crop Sci 42:299

Muehlbauer FJ, Weeden NF, Hoffman DL (1989) Inheritance and linkage relationships of morphological and isozyme loci in lentil (Lens Miller). J Heredity 80:298–303

Neumann P, Nouzova M, Macas J (2001) Molecular and cytogenetic analysis of repetitive DNA in pea (*Pisum sativum* L.). Genome 44:716–728

Paniego N, Echaide M, Munoz M, Fernandez L, Torales S, Faccio P, Fuxan I, Carrera M, Zandomeni R, Suarez EY, Hopp HE (2002) Microsatellite isolation and characterization in sunflower (*Helianthus annuus* L.). Genome 45:34–43

Pilet-Nayel ML, Muehlbauer FJ, McGee RJ, Kraft JM, Baranger A, Coyne CJ (2002) Quantitative trait loci for partial resistance to *Aphanomyces* root rot in pea. Theor Appl Genet 106:28–39

Prioul S, Frankewitz A, Deniot G, Morin G and Baranger A (2004) Mapping of quantitative trait loci for partial resistance to *Mycosphaerella pinodes* in pea (*Pisum sativum* L.), at the seedling and adult plant stages. Theor Appl Genet 108:1322–1334

Pugh T, Fouet O, Risterucci AM, Brottier M, Abouladze M, Deletrez C, Courtois B, Clement C, Larmande P, Goran JAKN, Lanaud C (2004) A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. Theor Appl Genet 108:1151–1161

Ritschel PS, de Lima Lins TC, Trdistan RL, Cortopassi Buso GS, Buso JA, Ferreira ME (2004) Development of microsatellite markers from an enriched genomic library for genetic analyses of melon (*Cucumis melo* L.). BMC Plant Biol 4:9

Sargent DJ, Davis TM, Tobutt KR, Wilkinson MJ, Battey NH, Simpson DW (2004) A genetic linkage map of microsatellite, gene-specific and morphological markers in diploid *Fragaria*. Theor Appl Genet 109:1385–1391

SAS Institute (1988) SAS/STAT Users' guide. SAS Institute Inc. Cary, NY

Schiex T, Chabrier P, Bouchez M, Milan D (2001) Boosting EM for radiation hybrid and genetic mapping. In: Proceedings of WABI'2001 (First workshop on algorithms in bioinformatics), LNCS 2149, Aug 2001, pp 41–51

Sharopova N, McMullen MD, Schultz L, Schroeder S, Villeda HS, Gardiner J, Bergstrom D, Houchins K, Hancock SM, Musket T, Polacco NDM, Edwards K, Ruff T, Register JC, Brouwer C, Thompson R, Velasco R, Chin E, Lee M, Woodman-Clikeman W, Long MJ, Liscum E, Cone K, Davis G, Coe EH Jr (2002) Development and mapping of SSR markers for maize. Plant Mol Biol 48:463–481

Somers DJ, Isaac P, Edwards K (2004) A High density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theor Appl Genet 109:1105–1114

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128

Tang S, Kishore VK, Knapp SJ (2003) PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor Appl Genet 107:6–19

Tar'an B, Warkentin T, Somers DJ, Miranda D, Vandenberg A, Blade S, Woods S, Bing D, Xue A, DeKoyeer D, Penner G (2003) Quantitative trait loci for lodging resistance, plant height and partial resistance to *mycosphaerella blight* in field pea (*Pisum sativum* L.). Theor Appl Genet 107:1482–1491

Tar'an B, Zhang C, Warkentin T, Tullu A, Vandenberg A (2005) Genetic diversity among varieties and wild species accessions of pea (*Pisum sativum* L.) based on molecular markers, morphological and physiological characters. Genome (in press)

Viruel MA, Hormaza JI (2004) Development, characterization and variability analysis of microsatellites in lychee (*Litchi chinensis* Sonn. *Sapindaceae*). Theor Appl Genet 108:896–902

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Heredity 93(1):77–78

Weeden NF, Ellis THN, Timmerman-Vaughan GM, Swiecicki WK, Rozov SM, Berdnikov VA (1998) A consensus linkage map for *Pisum sativum*. Pisum Genet 30:1–3

Yan Z, Denneboom C, Hattendof A, Dolstra O, Debener T, Stam P, Visser PB (2005) Construction of an integrated map of rose with AFLP, SSR, PK, RGA, RFLP, SCAZR and morphological markers. Theor Appl Genet 110:766–777

Yu JK, Tang S, Slabaugh MB, Heesacker A, Cole G, Herring M, Soper J, Han F, Chu WC, Webb DM, Thompson L, Edwards KJ, Berry S, Leon AJ, Grondona M, Olungu C, Maes N, Knapp SJ (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Sci 43:367–387