Hybrid origins of cultivated potatoes

Flor Rodríguez · Marc Ghislain · Andrea M. Clausen · Shelley H. Jansky · David M. Spooner

Received: 10 March 2010 / Accepted: 30 July 2010 / Published online: 24 August 2010

© US Government 2010

Abstract Solanum section Petota is taxonomically difficult, partly because of interspecific hybridization at both the diploid and polyploid levels. The taxonomy of cultivated potatoes is particularly controversial. Using DNA sequence data of the waxy gene, we here infer relationships among the four species of cultivated potatoes accepted in the latest taxonomic treatment (S. ajanhuiri, S. curtilobum, S. juzepczukii and S. tuberosum, the latter divided into the Andigenum and Chilotanum Cultivar Groups). The data support prior ideas of hybrid origins of S. ajanhuiri from the S. tuberosum Andigenum Group (2x = S. stenotomum) x S. megistacrolobum; S. juzepczukii from the S. tuberosum Andigenum Group (2x = S. stenotomum) x S. acaule; and S. curtilobum from the S. tuberosum Andigenum Group (4x = S. tuberosum subsp. andigenum) x S. juzepczukii. For the tetraploid cultivar-groups of S. tuberosum, hybrid origins are suggested entirely within much more closely related species, except for two of three examined accessions of the S. tuberosum Chilotanum Group that appear to have hybridized with the wild species S. maglia. Hybrid origins of the crop/weed species S. sucrense are more difficult to support and S. vernei is not supported as a wild species progenitor of the S. tuberosum Andigenum Group.

Introduction

Solanum tuberosum, the major cultivated potato of world commerce, has tremendous diversity of use to breeders. The taxonomy of Solanum L. sect. Petota Dumort. (including both the cultivated and wild potato species) is complicated by sexual compatibility among many species, introgression, interspecific hybridization, auto- and allo-polyploidy, a mixture of sexual and asexual reproduction, possible recent species divergence, phenotypic plasticity, and consequent great morphological similarity and difficulty in defining and distinguishing species (Spooner and van den Berg 1992; Huamán and Spooner 2002; Spooner 2009). These many complicating biological factors have led to great differences among taxonomic treatments of different authors. The latest taxonomic estimate in sect. Petota is about 100 wild species and 4 cultivated species (Spooner et al. 2009) divided into 3 clades (Spooner et al. 2008; Rodríguez and Spooner 2009; Rodríguez et al. 2009). This differs from previous estimates of 217 wild species and 7 cultivated species divided into 21 taxonomic series (Hawkes 1990).

Hawkes’s (1990) treatment of the 7 cultivated species is the most commonly cited, but the Russian potato taxonomists Bukasov (1971) and Lechnovich (1971) recognized 21 cultivated species, Ochoa (1990, 1999) 9 species and 141 infraspecific taxa (subspecies, varieties, and forms; including his unlisted autonyms) for the Bolivian cultivated
species alone. These authors classified cultivated potatoes as distinct species under the International Code of Botanical Nomenclature (ICBN; McNeill et al. 2006). Dodds (1962) anticipated the cultivar-group concept and recognized only three species, S. ×curtilobum, S. ×juzepczukii, and S. tuberosum, with five “groups” recognized in the latter. “Cultivar-groups” (the current terminology under the International Code of Nomenclature of Cultivated Plants; Brickell et al. 2009) are taxonomic categories used by the ICNCP to associate cultivated plants with traits that are of use to agriculturists. In combination with recent morphological analyses of Huamán and Spooner (2002), and microsatellite data, Spooner et al. (2007b) provided a reclassification of the cultivated potatoes into four species (1) S. tuberosum, with two cultivar-groups (a) Andigenum Group of upland Andean genotypes containing diploids, triploids and tetraploids, and (b) the Chilotanum Group of lowland tetraploid Chilean landraces, from which our modern cultivars arise (Ames and Spooner 2008), (2) S. ajanhuiri (diploid), (3) S. juzepczukii (triploid), and (4) S. curtilobum (pentaploid).

DNA sequence data of orthologous genes of polyploids and their related diploid wild species using GBSSI or waxy (Spooner et al. 2008), nitrate reductase (Rodríguez and Spooner 2009), and genomic in situ hybridization data (Pendinen et al. 2008) are providing new insights into polyploid origins of wild members of sect. Petota. These studies have concentrated on the wild species and have supported both auto- and allopolyploid origins for different species in sect. Petota. Elucidating the phylogenetic relationships within the polyploids is crucial for an effective taxonomic treatment of the section and for the utilization of wild potato germplasm in breeding programs. The main purpose of the present study is to examine hybrid origins of S. ajanhuiri, S. juzepczukii, and S. curtilobum using DNA sequence data of waxy, but we also include progenitors of S. tuberosum for completeness. Because these hybrid scenarios involve cultivar-groups of S. tuberosum (Tables 1, 2) we also include the cultivar-groups of S. tuberosum, their putative wild species progenitors, and placeholder species representative of the main clades and outgroups of sect. Petota.

Materials and methods

Plant materials

Twenty-six wild species and four cultivated species (S. ajanhuiri, S. curtilobum, S. juzepczukii and S. tuberosum) of Solanum sect. Petota, and three additional out-group species in sects. Etiuberosum and Lycopersicon were examined (Table 2). These include both diploid and tetraploid populations of S. tuberosum from the Andes (Andigenum Group) and from lowland Chile (exclusively tetraploid, Chilotanum Group). Two to fourteen accessions were examined for the cultivated species and their proposed wild species progenitors. We examined S. bukasovii, S. candolleaum, S. sparsipilum, and S. vernei as putative diploid wild species progenitors of S. tuberosum; S. sucrense as a possible hybrid with tetraploid S. tuberosum Andigenum Group and S. oplocense, S. berthaultii as a

Table 1 Hypotheses of origins of the polyploid cultivated potato species as summarized by Hawkes (1990), classified in taxonomic system of Spooner et al. (2007b), and explanations in the footnotes why some wild species were examined

<table>
<thead>
<tr>
<th>Cultivated species</th>
<th>Putative origins (Hawkes 1990)</th>
<th>Ploidy</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Solanum ajanhuiri</em> Juz. &amp; Bukasov</td>
<td>Diploid populations of <em>S. tuberosum</em> Andigenum Group</td>
<td>2x</td>
</tr>
<tr>
<td>(S. stenotomum) × <em>S. megistacrolobum</em> Bitter</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. curtilobum</em> Juz. &amp; Bukasov</td>
<td><em>S. juzepczukii</em> × tetraploid populations of <em>S. tuberosum</em> Andigenum Group</td>
<td>5x</td>
</tr>
<tr>
<td>(S. tuberosum subsp. <em>andigenum</em>)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. juzepczukii</em> Bukasov</td>
<td>Diploid populations of <em>S. tuberosum</em> Andigenum Group × <em>S. acaule</em> Bitter</td>
<td>3x</td>
</tr>
<tr>
<td><em>S. tuberosum</em> L. Andigenum Group</td>
<td>Diploid populations of <em>S. tuberosum</em> Andigenum Group</td>
<td>4x</td>
</tr>
<tr>
<td>4x [= <em>S. tuberosum</em> subsp. <em>andigenum</em> (Juz. &amp; Bukasov) Hawkes]</td>
<td>(=S. stenotomum) × <em>S. sparsipilum</em> (Bitter) Juz. and Bukasov(^a)</td>
<td></td>
</tr>
<tr>
<td><em>S. tuberosum</em> Chilotanum Group</td>
<td>Selection for long-day adaptation from tetraploid populations</td>
<td>4x</td>
</tr>
<tr>
<td>of <em>S. tuberosum</em> Andigenum Group(^b)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Brücher (1964) proposed the diploid wild species *S. vernei* as the progenitor of Andean populations of *S. tuberosum*. Astley and Hawkes (1979) proposed *S. sucrense* as a hybrid between *S. tuberosum* (tetraploid populations) and the wild species *S. oplocense*. Spooner et al. (2005) showed *S. tuberosum* Andigenum Group to have progenitors in the northern members of the *S. brevicaule* Group (to include *S. bukasovii* and *S. candolleaum*).

\(^b\) Hosaka (2003) proposed *S. tarijense* as a maternal parent in a hybrid origin of *S. tuberosum* Chilotanum Group. Spooner et al. (2007a) showed that *S. tarijense* was a synonym of *S. berthaultii*, but we examine here two accessions previously identified as *S. tarijense* (PI 545922, PI 566799). Raker and Spooner (2002) showed the wild species *S. maglia* Schltdl. to group with *S. tuberosum*.
Table 2 Species and accessions examined with waxy DNA sequence data

<table>
<thead>
<tr>
<th>Species</th>
<th>Plastid clade</th>
<th>Ploidy(^a)</th>
<th>Accession number (no. of clones sequenced, no. of alleles discovered)(^b)</th>
<th>NCBI sequence number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Solanum</em> L. sect. <em>Petota</em> Dumort. wild species</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. acaule</em> Bitter</td>
<td>4</td>
<td>4x</td>
<td>472735 (7,4)</td>
<td>HM561721–HM561724</td>
</tr>
<tr>
<td><em>S. andreanum</em> Baker</td>
<td>3</td>
<td>2x</td>
<td>320345</td>
<td>HM561729</td>
</tr>
<tr>
<td><em>S. avilesii</em> Hawkes &amp; Hjert.</td>
<td>4</td>
<td>2x</td>
<td>498091</td>
<td>HM561730</td>
</tr>
<tr>
<td><em>S. berthaultii</em> Hawkes(^c)</td>
<td>4</td>
<td>2x</td>
<td>498096, 498105, 545851, 545922, 566799</td>
<td>HM561731, HM561732, HM561733, HM561734, HM561735</td>
</tr>
<tr>
<td><em>S. bukasovii</em> Juz.</td>
<td>4</td>
<td>2x</td>
<td>210042</td>
<td>HM561736</td>
</tr>
<tr>
<td><em>S. bulbocastanum</em> Dunal</td>
<td>2</td>
<td>2x</td>
<td>347757</td>
<td>HM561737</td>
</tr>
<tr>
<td><em>S. candolleaum</em> P. Berthault</td>
<td>4</td>
<td>2x</td>
<td>545972</td>
<td>HM561738</td>
</tr>
<tr>
<td><em>S. cardiophylum</em> Lindl.</td>
<td>2</td>
<td>2x</td>
<td>595465</td>
<td>HM561739</td>
</tr>
<tr>
<td><em>S. circarifolium</em> Bitter</td>
<td>4</td>
<td>2x</td>
<td>498117</td>
<td>HM561741</td>
</tr>
<tr>
<td><em>S. clarum</em> Correll</td>
<td>2</td>
<td>2x</td>
<td>283099</td>
<td>HM561742</td>
</tr>
<tr>
<td><em>S. euhrenbergii</em> (Bitter) Rydb.</td>
<td>1</td>
<td>2x</td>
<td>611097</td>
<td>HM561751</td>
</tr>
<tr>
<td><em>S. infundibuliforme</em> Phil.</td>
<td>4</td>
<td>2x</td>
<td>472857</td>
<td>HM561753</td>
</tr>
<tr>
<td><em>S. jamesii</em> Torr.</td>
<td>1</td>
<td>2x</td>
<td>458424</td>
<td>HM561754</td>
</tr>
<tr>
<td><em>S. maglia</em> Schltdl., Chile</td>
<td>4</td>
<td>3x</td>
<td>558316 (8,3)</td>
<td>HM561760–HM561762</td>
</tr>
<tr>
<td><em>S. maglia</em>, Argentina</td>
<td>4</td>
<td>2x</td>
<td>C1M862 (8,3), C1M865 (8,3), C1M868 (8,3), C1M870 (8,3), C1M871 (9,3)</td>
<td>HM561763–HM561765, HM561766–HM561768, HM561769–HM561771, HM561772–HM561774, HM561775–HM561777</td>
</tr>
<tr>
<td><em>S. megistacrolobum</em> Bitter</td>
<td>4</td>
<td>2x</td>
<td>233125, 310932</td>
<td>HM561778, HM561779</td>
</tr>
<tr>
<td><em>S. oplocense</em> Hawkes</td>
<td>4</td>
<td>4x</td>
<td>265885 (12,1), 458390 (12,1)</td>
<td>HM561780, HM561781</td>
</tr>
<tr>
<td><em>S. pascoense</em> Ochoa</td>
<td>3</td>
<td>2x</td>
<td>365339</td>
<td>HM561782</td>
</tr>
<tr>
<td><em>S. piurae</em> Bitter</td>
<td>3</td>
<td>2x</td>
<td>310997</td>
<td>HM561783</td>
</tr>
<tr>
<td><em>S. polyadenium</em> Greenm.</td>
<td>1</td>
<td>2x</td>
<td>161728</td>
<td>HM561784</td>
</tr>
<tr>
<td><em>S. raphanifolium</em> Card. &amp; Hawkes</td>
<td>4</td>
<td>2x</td>
<td>265862</td>
<td>HM561785</td>
</tr>
<tr>
<td><em>S. santolalae</em> Vargas</td>
<td>4</td>
<td>2x</td>
<td>195168</td>
<td>HM561786</td>
</tr>
<tr>
<td><em>S. sparsiplum</em> (Bitter) Juz. &amp; Bukasov</td>
<td>4</td>
<td>2x</td>
<td>230503 (8,2), 234014, 265871, 275153, 310933, 414151, 473531, 498073, 498074, 498305 (5,2), 546007, 597715, 597766, 597696</td>
<td>HM561787–HM561788, HM561789, HM561790, HM561791, HM561792, HM561793, HM561794, HM561795, HM561796, HM561797–HM561798, HM561799, HM561801, HM561802, HM561800</td>
</tr>
<tr>
<td><em>S. sucrense</em> Hawkes</td>
<td>4</td>
<td>4x</td>
<td>473364 (12,4), 545916 (10,2)</td>
<td>HM561803–HM561806, HM561807–HM561808</td>
</tr>
<tr>
<td><em>S. vernei</em> Bitter &amp; Wittm.</td>
<td>4</td>
<td>2x</td>
<td>458371</td>
<td>HM561828</td>
</tr>
<tr>
<td><em>S. verrucosum</em> Schltdl.</td>
<td>4</td>
<td>2x</td>
<td>545745</td>
<td>HM561829</td>
</tr>
<tr>
<td><em>S. violaceimarmoratum</em> Bitter</td>
<td>4</td>
<td>2x</td>
<td>473396</td>
<td>HM561830</td>
</tr>
</tbody>
</table>
Table 2 continued

<table>
<thead>
<tr>
<th>Species</th>
<th>Plastid clade</th>
<th>Ploidy&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Accession number (no. of clones sequenced, no. of alleles discovered)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>NCBI sequence number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Solanum</em> sect. <em>Petota</em> cultivated species</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. ajanhuiri</em> Juz. &amp; Bukasov</td>
<td>4</td>
<td>2x</td>
<td>703810 (8,2), 704227 (8,2)</td>
<td>HM561725–HM561726, HM561727–HM561728</td>
</tr>
<tr>
<td><em>S. curtilobum</em> Juz. &amp; Bukasov</td>
<td>4</td>
<td>5x</td>
<td>702282 (15,4), 702937 (16,4)</td>
<td>HM561743–HM561746, HM561747–HM561750</td>
</tr>
<tr>
<td><em>S. jupeczukii</em> Bukasov</td>
<td>4&lt;sup&gt;d&lt;/sup&gt;</td>
<td>3x</td>
<td>700895 (12,2), 703262 (9,2)</td>
<td>HM561755–HM561756, HM561757–HM561758</td>
</tr>
<tr>
<td><em>S. tuberosum</em> L. Andigenum Group 2x (includes <em>S. phureja</em> Juz. and Bukasov (705154, 705825), and <em>S. stenotomum</em> Juz. &amp; Bukasov (234011, 703783, 706025))</td>
<td>4</td>
<td>2x</td>
<td>234011, 703783, 705154, 705825, 706025</td>
<td>HM561809, HM561810, HM561811, HM561812, HM561813</td>
</tr>
<tr>
<td><em>S. tuberosum</em> Andigenum Group 4x (=<em>S. tuberosum</em> subsp. <em>andigenum</em> (Juz. &amp; Bukasov) Hawkes)</td>
<td>4</td>
<td>4x</td>
<td>265882 (9,4)</td>
<td>HM561814–HM561817</td>
</tr>
<tr>
<td><em>S. tuberosum</em> Chilotanum Group 4x (=<em>S. tuberosum</em> subsp. <em>tuberosum</em>)</td>
<td>4</td>
<td>4x</td>
<td>245835 (6,4), 703606 (15,3), 703610 (15,3)</td>
<td>HM561818–HM561821, HM561822–HM561824, HM561825–HM561827</td>
</tr>
<tr>
<td>Outgroups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. lycopersicum</em> L.</td>
<td>2x</td>
<td></td>
<td>LA1673</td>
<td>HM561759</td>
</tr>
<tr>
<td><em>S. peruvianum</em> L.</td>
<td>2x</td>
<td></td>
<td>LA2744</td>
<td>HM561740</td>
</tr>
<tr>
<td><em>Solanum</em> sect. <em>Etuberosum</em> (Bukasov &amp; Kameraz) A. Child</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. etuberosum</em> Lindl.</td>
<td>2x</td>
<td></td>
<td>498311</td>
<td>HM561752</td>
</tr>
</tbody>
</table>

<sup>a</sup> Ploidy (2n = 2x = 24; 2n = 4x = 48; 2n = 5x = 60)

<sup>b</sup> The six-digit 600,000 number and lower series are US Plant Introduction Numbers; the six-digit 700,000 numbers series are International Potato Center Numbers; LA numbers are from the C.M. Rick Tomato Genetics Stock Center; CIM refer to collections at INTA made by A.M. Clausen and R.W. Masuelli

<sup>c</sup> Hosaka (2003) supported *S. tarijense* as a maternal genome contributor to the evolution of the *S. tuberosum* Chilotanum Group but Spooner et al. (2007a) reclassified *S. tarijense* under *S. berthaultii*. We examine some accessions formerly identified as *S. berthaultii* and some as *S. tarijense*.  

<sup>d</sup> Not examined in plastid DNA studies but expected to be in clade 4 because both putative parents are in clade 4
possible maternal species contributor to *S. tuberosum* Chilotanum Group, *S. megistacrolobum* as a possible contributor to *S. ajanhuiri*, and *S. acaule* as a possible parent in the hybrid origin of *S. curtifolium* and *S. juzepczukii* (Table 1). We included additional species of sect. *Petota* (e.g., *S. jamesii*, *S. piurae*) to provide good cladistic structure within sect. *Petota*.

**Waxy** primer design, cloning and sequencing

DNA sequences were obtained from the first to the eighth exon of the *waxy* gene. We cloned and sequenced amplicons of all the polyploid cultivated species and those diploid species showing polymorphic sequences from initial direct sequencing. With cloned products, 299 total DNA sequences from 69 accessions were examined in our study. All techniques regarding DNA isolation, purification, primer design, cloning, and sequencing follow Peralta and Olmstead and Palmer (1994): (1) 1 million replicates initiated from random trees with a sequential heat of 0.05 (determined empirically) for 5 million generations with trees sampled every 100 generations. At the end of the run, convergence was evaluated by two methods: (1) a visual inspection of a graph of likelihood as a function of the restriction site (binary) model for DNA data and (4) the resulting trees were searched with TBR and MULPARS. The last two analyses were terminated at 10,000 trees. The resulting trees were used to compute a strict consensus tree. A bootstrap analysis was conducted on 500 replicates with TBR and MULPARS. The above analyses were done twice, once with the entire dataset and once with the diploids.

Maximum likelihood (ML; Felsenstein 1973) analysis was used to assess 56 models of sequence evolution evaluated in Modeltest ver. 3.07 (Posada and Crandall 1998) using the hierarchical likelihood ratio test (hLRT) at \( \alpha = 0.01 \). A maximum likelihood heuristic search was conducted using a genetic algorithm as implemented in GARLI 0.951 (Zwickl 2006). The run was repeated three times from random starting trees using the default settings for the components of the genetic algorithm but the default termination conditions were changed allowing the software to terminate the run after 300,000 generations if no better scoring topology (branching pattern) was encountered. A HKY + \( \Gamma \) model was used. To evaluate the stability of clades on the optimal tree, a bootstrap analysis was performed with 100 bootstrap replicates performing the same search, under the best model of evolution with parameters fixed.

Bayesian Markov chain Monte Carlo (MCMC) phylogenetic analysis (Yang and Rannala 1997) was also performed using MrBayes version 3.1 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003), as an independent measurement of phylogenetic relationship and clade composition. The HKY + \( \Gamma \) model for DNA data was chosen according to the criteria described above and the restriction site (binary) model for gaps. Model parameters were estimated separately for DNA and gaps; for gaps the restriction site (binary) model was used as “let coding = variable” which accounts for the ascertainment bias produced by characters that are constant (either state 0 or 1) in all taxa and are not observed. Tree searching using MrBayes was performed by running six linked chains, initiated from random trees with a sequential heat of 0.05 (determined empirically) for 5 million generations with trees sampled every 100 generations. At the end of the run, the proportion of searches in which any given
node is found during the post-burn-in portion of the chain constitutes the Bayesian posterior probability for that node.

**Results**

The aligned DNA data matrix was 994 characters long and gap scores added an additional 83 characters. MP analyses produced more than 5,000 trees (the upper limit saved), using 151 characters as parsimony informative to produce a tree of 477 steps long, with a consistency index of 0.82 and retention index of 0.92. The three independent Bayesian MCMC runs (each with two internal runs) mixed well as indicated by the two convergence diagnostics mentioned above. The Bayesian tree was nearly identical to the MP tree using all characters except for minor differences in topology at the terminal branches. Figure 1 shows the Bayesian tree with Bayesian posterior probabilities.

The MP tree excluding the 83 gap scores (not shown), produced more than 5,000 trees (the upper limit saved), using 99 characters as parsimony informative to produce a tree of 366 steps long, with a consistency index of 0.85 and a retention index of 0.92. However, it was highly unresolved and placed the tomato clade within the potato clade.

For ML, the best model that explained the data was HKY $+ \Gamma$. ML does not use gap characters, and like the MP tree it was similarly unresolved. Within sect. *Petota*, clade 3 is at the base of the tree, followed by a grade of species from clades 1 to 2 and *S. circiacifolium*, and then species on clade 4a + 4b. Hence, we discuss only the phylogenetic results using gaps.

All accessions of cultivated potato and their putative wild species progenitors fall in clade 4a or 4b. All accessions of the three cultivated hybridogenetic species *S. ajchanuiri*, *S. curtloibum*, and *S. juzepczukii* have their alleles in every case partitioned into the two well-supported clades 4a and 4b. Two of the 13 accessions of the diploid wild species *S. sparsipilum* (PI 230503, PI 498305) also have their alleles partitioned into clades 4a and 4b. Two of the three accessions of *S. tuberosum* Chilotonanum Group have alleles in both clades 4a and 4b. Within clade 4a, two of the three accessions of the *S. tuberosum* Chilotonanum Group (703606, 703610) resolve with all accessions of the wild species *S. maglia* on a well-supported clade. Within clade 4b, alleles of the *S. tuberosum* Chilotonanum Group fall in two different places. Two accessions of the *S. tuberosum* Chilotonanum Group (245385, 703606) cluster with *S. berthaultii* (=*S. tarjense*), also on a well-supported clade.

All accessions of the *S. tuberosum* Andigenanum Group at both the 2x and 4x ploidy levels fall in clade 4b, with wild species in the northern *S. brevicaule* group (*S. bukasovii, S. candolleaum*). However, clade 4b also contains wild species in the southern members of this group, as well as species outside the group (e.g., *S. avilesii, S. berthaultii, S. infundibuliforme, S. oplocense, S. sparsipilum, S. sucreense, S. verrucosum*).

**Discussion**

Cladistic structure of sect. *Petota*

The MP tree excluding the 83 gap scores was highly unresolved. Similar to the *waxy* (Spooner et al. 2008) and nitrate reductase (Rodrıíguez and Spooner 2009) analysis of the wild potato species, this analysis shows the importance of insertion and deletion characters as phylogenetic markers in these genes.

Plastid DNA phylogenies document section *Petota* to be divided into four clades that often show little relationship to Hawkes’s taxonomic series (Spooner and Castillo 1997). These clades contain: (1) North and Central American diploid species, with the exception of *S. bulbocastanum, S. cardiophyllum, and S. verrucosum*; (2) *S. bulbocastanum, and S. cardiophyllum*; (3) all examined members of series *Piurana* and some South American species in series *Conicibaccata, Megistacroloba, Tuberosa, Yungasensis*; (4) all remaining South American species, *S. verrucosum* from Mexico, and North and Central American polyploid species. Phylogenies from *waxy* (Spooner et al. 2008) and nitrate reductase (Rodrıíguez and Spooner 2009) recover the same four clades except unite clades 1 and 2 as a single clade, support two well-supported subclades in clade 4 (4a, 4b), and support allopolyploid origins for some of the polyploids. Our results (Fig. 1) show clade 3 to be at the base of the tree, followed by a grade of species from clades 1 + 2 and *S. circiacifolium*, and then species from clade 4a + 4b. This differs from the prior *waxy* analysis (Spooner et al. 2008) where all of these three clades formed a polytomy; and from the dominant COSII results (Rodrıíguez et al. 2009) where clade 1 + 2 are basal and clades 3 and 4 are sister. This is the first nuclear orthologous DNA sequence phylogeny including *S. circiacifolium*, which is placed in clade 4 in the plastid results (Spooner and Castillo 1997) but in the grade of species (grade 1 + 2) in the present nuclear results. Of interest is that *S. circiacifo- lium* has white stellate corollas similar to the Mexican diploid species in grade 1 + 2. Removing *S. circiacifolium* from the present analysis maintains the same main clades.

Cladistic structure of the cultivated potato clades 4a, 4b

The results show alleles of all accessions of the three cultivated species *S. ajchanuiri, S. curtloibum, and S. juzepczukii* in every case be partitioned into the two well-supported clades 4a and 4b, grouping with their
Fig. 1 Bayesian phylogram of the waxy data of *Solanum* sect. *Petota* and outgroups with branch lengths drawn in proportion to the estimated number of substitutions per site and representing an average of the branch length of all trees sampled in the Markov chain that have that branch. Posterior probability values are indicated for nodes with a probability value greater than 0.50. The number after the species name is the accession number (Table 2), followed by the allele number, and in parenthesis the number of identical allele variants falling on branch. Accessions in **black color** are wild species, in **red** are the *S. tuberosum* Andigenum Group including both diploids and tetraploids, in **green** are the *S. tuberosum* Chilotanum Group (all tetraploid), and in **blue** the cultivated hybrid potatoes (Table 1). The numbering of the clades follows the convention of the plastid and waxy DNA results (Spooner et al. 2008).
parental species as outlined in prior hypotheses (Table 1). For example, alleles of S. ajanhuii group with both parents S. megistacrolorum and diploid accessions of S. tuberosum Andigenum Group, and similar support for published hybrid origins (Table 1) are found with S. curtillobum and S. juzepczuki.

An unexpected result was for two of the 14 accessions of the diploid wild species S. sparsipilum (PI 230503, PI 498305) to have their alleles partitioned into clades 4a and 4b, also suggesting hybrid origins. We examined multiple (14) accessions of S. sparsipilum only after our initial sequencing of S. sparsipilum PI 230503 showed its two allelic variants in clades 4a and 4b. A similar partitioning of alleles from diploid wild species into clades 4a and 4b was found with nitrate reductase in the diploid wild species S. brevicaule and another accession of S. sparsipilum not examined here (Rodríguez et al. 2009). S. brevicaule and S. sparsipilum are two diploid wild species members of the taxonomically complex S. brevicaule group (Spooner et al. 2005). Morphological data (Van den Berg et al. 1998; Alvarez et al. 2008) and DNA marker data (Miller and Spooner 1999; Spooner et al. 2005) suggest that these two species may be conspecific. Irrespective of their taxonomic status, allelic data from this study support their hybrid origins (Baum 2007) across the well-supported clades 4a and 4b, and may help explain the taxonomic difficulty in the S. brevicaule group.

The finding that two of the three accessions of the S. tuberosum Chilotanum Group (703606, 703610) have alleles in both well-supported clades 4a (grouping with all accessions of the wild species S. maglia) and 4b, reflects the microsatellite results of Raker and Spooner (2002) that grouped S. maglia with the S. tuberosum Chilotanum Group. The fact that within clade 4b, alleles of two accessions (245835, 703606) of the S. tuberosum Chilotanum Group cluster with S. berthaultii (=S. tarijense), also on a well-supported clade, match hypotheses of a wild species maternal contributor from S. berthaultii to the S. tuberosum Chilotanum Group by Hosaka (2002). Brücher’s (1964) hypothesis of S. vernei as a progenitor of S. tuberosum Andigenum Group (4×) is not supported by the present waxy results, concordant with AFLP data (Spooner et al. 2005). Although S. vernei is a member of clade 4b, it is removed from the bulk of the species in this clade by high support values (1 posterior probability).

The grouping of all accessions of S. tuberosum (both the Andigenum and Chilotanum Groups) at both the 2x and 4x ploidy level in clade 4b with their wild species progenitors in the northern S. brevicaule group (S. bukasovii and S. candolleanum) is consistent with AFLP data (Spooner et al. 2005). However, clade 4b also contains wild species in the southern members of this group, as well as species outside the group (e.g., S. avilesii, S. berthaultii, S. infundibuliforme, S. oplocense, S. sparsipilum, S. succrense and S. verrucosum). Clearly, the region of waxy examined here lacks the resolving power to address progenitors of the S. tuberosum Andigenum Group that was possessed by the large AFLP database of Spooner et al. (2005; 438 character bands from 6 AFLP primer combinations). Waxy and AFLPs are very different markers. AFLPs, RAPDs, and nRFLPs are largely anonymous marker data spread throughout the genome, representing overall genetic relationships but only suitable for investigating closely related species. Waxy, on the other hand, is a well-characterized ortholog in potatoes representing much better knowledge of homologs, but representing only a single area of the genome.

In summary, DNA sequence data of orthologous nuclear genes, including waxy, in sect. Petota and in the sister clade sect. Lycopersicon, have been proven to be very useful to support prior molecular phylogenies and to elucidate allopolyploid origins (Spooner et al. 2008; Rodríguez and Spooner 2009; Rodríguez et al. 2009). Waxy similarly supports hypotheses of allopolyploid origins for S. juzepczuki and S. curtillobum, and a diploid hybrid origin for S. ajanhuii. For the tetraploid cultivar-groups of S. tuberosum Andigenum and Chilotanum Groups, hybrid origins are suggested entirely within much more closely related species within clade 4b, except for two of the three accessions of the S. tuberosum Chilotanum Group that appear to have hybridized with the wild Chilean species S. maglia from clade 4a. These data support an original upland Andean origin of the S. tuberosum Chilotanum Group from within the S. tuberosum Andigenum Group, followed by long distance transport to Chile. Hybrid origins of the crop/weed species S. succrense are more difficult to support because of poor structure within clade 4b. S. vernei is not supported as a wild species progenitor of the S. tuberosum Andigenum Group.

Implications for potato breeding

The synthesis of the four cultivated species in nature can be predicted based on knowledge of rules for interspecific hybridization. In potato, ploidy does not always predict crossing success. Consequently, each Solanum species has been assigned an endosperm balance number (EBN) based on its ability to hybridize with standard species for which EBN values have been assigned empirically (Hanneman 1994). If no other crossing barriers exist, then successful hybridization is expected when male and female gametes have matching EBN values, regardless of ploidy. The diploid 2 EBN species S. ajanhuii resulted from hybridization between the two diploid 2 EBN species S. stenotomum and S. megistacrolorum. In this case, the parents and their interspecific hybrids had matching ploidy and EBN
values. The triploid 2 EBN species *S. juzepczukii* resulted from hybridization between diploid 2 EBN *S. stenotomum* and tetraploid 2 EBN *S. acaule*. While the parents differed in ploidy levels, their EBN values matched, so hybridization was successful. In this case, triploid *S. juzepczukii* is expected to be male and female sterile. However, if it carries a mutation that interrupts meiosis, then it may produce functional 2n (numerically unreduced) gametes. These meiotic mutations occur naturally and frequently in cultivated and wild potatoes (Peloquin et al. 1999; Carputo et al. 2000). The 2n (triploid, 2 EBN) gametes from *S. juzepczukii* could then unite with normal (diploid, 2 EBN) gametes from the *S. tuberosum* Andigenum Group to produce pentaploid 4 EBN offspring. The pentaploid 4 EBN species *S. curtilobum* resulted from hybridization between triploid 2 EBN *S. juzepczukii* and tetraploid 4 EBN *S. tuberosum* Andigenum Group. Pentaploid 4 EBN clones readily cross with tetraploid 4 EBN clones, producing fertile aneuploid offspring (Johnston and Hanneman 1982; Camadro and Esperillio 1990; Adiwilaga and Brown 1991; Carputo et al. 2003b). Because 2n gametes are genetically determined, repeated sexual polyploidization events likely occurred in nature. This would also contribute to genetic diversity in polyploid species (Carputo and Barone 2005). Tetraploid 4 EBN *S. tuberosum* probably arose through bilateral sexual polyploidization, in which the diploid 2 EBN female parent produced 2n eggs and the diploid 2 EBN male parent produced 2n pollen. A high allele frequency of a 2n gamete-producing gene supports the hypothesis that tetraploid cultivated potato originated via sexual polyploidization (Carputo et al. 2003a).

Tremendous allelic diversity is apparent in cultivated potato species. Genomes from clades 4a and 4b in all species (except the *S. tuberosum* Andigenum Group) contribute to this diversity. Polyploid hybrid cultivated potato species likely arose through the functioning of 2n gametes (Carputo et al. 2003a). Compared to somatic doubling, sexual polyploidization results in high levels of genotypic diversity (Watanabe et al. 1991; Werner and Peloquin 1991). While the mutational decay of functionally duplicated genes in polyploids can reduce useful genetic variability, recent studies indicate that alternative fates are possible. Instead of losing their functionality, duplicated genes may find new roles or partition of their ancestral roles (Adams et al. 2004; Wang et al. 2004). Partitioning of the levels of expression of duplicate genes allows for subfunctionalization. These gene pairs can diverge to produce parallel networks that are expressed in different tissues or environments. The epigenetic alteration of gene expression may be stable throughout even long evolutionary periods in polyploids (Adams and Wendel 2005). Since both members of a duplicated gene pair would be subjected to selection pressure, mutational decay would not occur and the genetic variability contributed by the progenitor species would be maintained. In addition, dormant transposable elements may become transcriptionally activated as a result of polyploidization (Madlung et al. 2005).

The allelic diversity in cultivated *Solanum* species may be particularly valuable to potato breeders. Potato exhibits strong hybrid vigor due to genetic interactions within and among loci (Tai 1976; Mendiburu and Peloquin 1977). Because high potato yield and vigor depend on these allelic interactions, breeding strategies typically aim to maximize heterozygosity. The introgression of germplasm with high levels of allelic diversity into potato breeding clones is likely to contribute to high yield. In addition, the ability to express different combinations of duplicate gene alleles in variable environments is likely to result in phenotypically stable clones. Yield stability across environments has been reported in genetically diverse potato populations and is attributed to the buffering provided by allelic diversity (Darmo and Peloquin 1990; Ortiz et al. 1991).

Allopolyploid cultivated species, such as *S. juzepczukii* and *S. curtilobum*, may be a source of superior germplasm for domestication. The union of two diverged genomes provides the opportunity to utilize loci from both genomes during domestication. Human selection may act on the regulation of new transcriptional networks that are produced by combining genomes when polyploids are formed. These novel pathways may contribute to important domestication traits (Hovav et al. 2008a). In addition, genes on homoeologous chromosomes have been demonstrated to be differentially transcribed in allopolyploids (Hovav et al. 2008b). The level of dominance of expression of one genome over another varies among tissues and developmental time points, and may be influenced by domestication events (Flagel et al. 2008). For example, D genome expression bias was observed during fiber development of domesticated cotton, implying that human selection was responsible for the preferential recruitment of that genome. The extent to which expression bias occurs appears to increase during the evolution of allopolyploids (Flagel et al. 2008). Studies are being planned to compare genome expression bias in *S. juzepczukii* and *S. curtilobum*, and in newly formed polyploids in order to determine whether expression bias occurs in potato allopolyploids.

The predominant potato cultivars in high altitude regions (above 3,700 m) are members of *S. ajanhuiri*, *S. juzepczukii*, and *S. curtilobum* (Huamán and Spooner 2002). The frost tolerance contributed by the progenitor species *S. acaule* and *S. megistacrolobum* (Hijmans et al. 2003) likely contributed to their success in these regions. In addition, however, the novel gene interactions generated by combining divergent genomes may have been important for establishment and maintenance of these cultivar-groups in the severe growing conditions of the Andean highlands.
The four cultivated species of potato represent a rich germplasm resource for potato breeders. Modern potato cultivars are closely related to each other and, consequently, are low in genetic diversity (Plaisted and Hoopes 1989; Ortiz 1998). They have been separated from their wild relatives since they were brought to Europe almost 500 years ago (Rios et al. 2007; Ames and Spooner 2008). However, cultivated Solanum species continue to hybridize with wild potato species in South America (Ugent 1970; Grun 1990). Unlike wild Solanum species, which also offer breeders a source of genetic diversity, the cultivated Solanum species have been selected for traits of interest to humans, including large tuber yield and size, low glycoalkaloid levels, desirable flavor, short cooking times, and high nutritional value (Johns and Alonso 1990; De Maine et al. 1993; Winfield et al. 2005; Bradshaw et al. 2006; Morris et al. 2007, 2008; Ducreux et al. 2008). All cultivated Solanum species can be readily introgressed into modern potato cultivars by breeders. Members of both tetraploid Solanum tuberosum cultivar-groups can be crossed directly to cultivars, while diploid Solanum ajanhuiri clones selected for 2n gamete production can also be crossed to cultivars, producing tetraploid offspring. Although Solanum juzepczukii is a sterile triploid, 2n gamete producing clones are directly crossable to cultivars. As discussed earlier, 2n gametes are common in Solanum polyploids. Finally, pentaploid Solanum curtloibum will cross directly to cultivars, producing fertile aneuploid offspring.

Acknowledgments This study was supported by NSF DEB 0316614 entitled PBI Solanum: a worldwide treatment (http://www.nhm.ac.uk/researchcuration/projects/solanaceae/​) to DS and by the USDA National Research Initiative Grant 2008-35300-18669 to DS.

References

Ochoa CM (1999) Las papas de Sudamerica: Peru. Centro Internacional de La Papa (CIP), Lima, Peru


