On the origin of *Solanum nigrum*: can networks help?

Péter Poczai · Jaakko Hyvönen

Abstract  Black nightshades are a group of species best known for their ‘poisonous’ or noxious weedy reputation. It is not so well known that species of this group serve as emerging food source in many countries worldwide especially in the African continent. Despite the fact that the section has recently been studied extensively, taxonomy is still unsettled and debated because of inter- and intraspecific hybridization, phenotypic plasticity and polyplodization. In this study we analyze the genetic relationships among diploid, tetraploid and hexaploid species of sect. *Solanum*, which have possibly taken part in the formation of *Solanum nigrum*, utilizing multi-locus (SCoT, ISSR) markers combined with chloroplast *trnL*-*F* sequence data and morphological characters. We scored 51 morphological characters united with SCoT (171), ISSR (224) and *trnL*-*F* (1042), for simultaneous analysis of 49 terminals and 1488 characters. The topology of the tree is concordant with the results of the network analysis. In the phylogenetic networks, all the accessions of the diploid species shared a split with all of the polyploid species. This reflected a high portion of shared ISSR and SCoT bands between diploids and polyploids. In addition, a strong split divided the diploid species. The history of *S. nigrum* might be reticulate with hybrid speciation playing an important role. Genetically differentiated diploids in few combinations have created a series of genetically distinct polyploid populations. The insufficient isolation that permitted further recombination between ancient polyploids and diploids have resulted in high level of genotypic and phenotypic polymorphism. This high level of novel genomic variability obviously enabled species to succeed in their new environment.

Keywords  Black nightshades · *Solanum nigrum* · Splits graphs · Phylogenetic networks · Start codon targeted polymorphism (SCoT) · Inter-simple sequence repeats (ISSR)

Introduction

The genus *Solanum* L. comprises a wide range of economically important crop plants like potato (*Solanum tuberosum* L.), eggplant (*Solanum melongena* L.) and tomato (*Solanum lycopersicum* L.). Beside these major crops, the genus also contains many minor food plants like species of *Solanum* sect. *Solanum*, or black nightshades. They are known as poisonous weeds in many parts of the world, while in others they are used as valuable leafy vegetables. Because of this ‘poisonous reputation’, edible species of the group carry a negative disgrace, and thus are neglected or underutilized in Europe. Black nightshades serve as emerging food sources in Africa, but some varieties are cultivated elsewhere. However, no cultivars have yet been developed using traditional plant breeding techniques [1]. Large number of traditional landraces and varieties exist in regions where they are utilized as food and/or medical plants [2]. In addition to many local uses in Africa the species of sect. *Solanum* are potentially globally important for agriculture, human health, plant breeding and...
biotechnology. The leaves and berries are a potential source of coloring plant extracts, inks and dyes [3, 4], and they are rich in proteins, fibers, vitamins and amino acids [5–7]. The value of the group as a genetic resource for plant breeding is also very significant because *S. nigrum* L. is resistant to all known races of late blight—economically one of the most important diseases of solanaceous plants—caused by an oomycete *Phytophthora infestans* (Mont.) de Bary [8–11]. While resistance has been successfully transferred to the potato in some experiments [12–15]; the species has not proved to be a useful source for traditional breeding. However, the genes responsible for the resistance (R) might be cloned and introduced to the potato [16] offering a promising new gene pool of possible R genes [17, 18].

Section *Solanum*, also known as the *S. nigrum* complex, is composed of 50 species [19] and it is one of the most variable groups of the genus [20]. Polyploidy and hybridization is often reported as an important source of diversity [21], while members of the group are phenotypically diverse. Natural inter- and intraspecific hybrids have been described [22–26], as well as individuals with higher level of ploidy among diploid and hexaploid taxa [24, 27–29]. Since the description of *S. nigrum* by Linnaeus [30] more than 30 names have been published referring to the same species. The great number of junior synonyms has resulted in large number of conflicting and confusing species definitions [31–33]. Despite the fact that the group has recently been studied extensively [34–41], its taxonomy is still unsettled and debated [31].

The section is placed to the Morelloid clade in the 13 clade scheme of Bohs [42]—which is being used for example to arrange the taxa for the PBI *Solanum* project based on ndhF sequence data. According to Weese and Bohs [43] and Poczai et al. [44] the Morelloid clade is closely related to the representatives of the Dulcamaroid clade (e.g. *Solanum dulcamara* L.). While the infrageneric relationships have been studied in detail [45–51], much less is known about the relationships of the species. To address important question about biogeography, genetic diversity and relationships, markers covering the whole genome (e.g. AFLP, RAPD, ISSR) are often used [52, 53]. In the case of the *S. nigrum* complex molecular markers like RAPDs have been used by Stracke et al. [54] to study genetic variation among GenBank accessions of *Solanum americanum* Mill., *S. nigrum* and *Solanum melanocerasum* All. The same technique was successfully used by Poczai et al. [44] to assess genetic diversity among common species of the section found in Europe and later to clarify the origin of rarely cultivated edible varieties [55]. AFLP and SSR markers have also been used [56–58] to characterize genetic diversity among the germplasm accessions of the Gatersleben GenBank as well as to classify taxonomically unknown material. These studies simultaneously uncovered significant differences among the accessions of *S. americanum*. AFLPs were used to study different genotypes of black nightshades alongside with interbreeding studies to reveal interspecific relationships between species occurring in South Africa [59], Uganda [60, 61] and other parts of Africa [62, 63]. DNA sequence analyses have not been so widely used in studies of this group, but Poczai et al. [64] used rDNA ITS data, along with ITS2 RNA transcript secondary structure modeling, to reveal the origin of an ambiguous historical *Solanum scabrum* Mill. herbarium specimen from the eighteenth century collected by Paulus Kitaibel.

The cited studies clearly reveal that even with the use of molecular methods the species of the group cannot be identified easily. In addition, it is quite obvious that new phylogenetic methods (e.g. phylogenetic network modeling) with the aim to detect non-tree like evolutionary histories will be needed for successful phylogenetic studies of this group. Conventional phylogenetic methods assume that evolutionary history can be represented by means of bifurcating trees [65]. However, it is likely that large number of angiosperm species have arisen through hybridization [66], and thus their history would be better represented as a network [67]. Many new algorithms have been developed to model and visualize reticulating events [68–70].

In this study we aimed to analyze genetic relationships among diploid, tetraploid and hexaploid species of sect. *Solanum*, which have possibly taken part in the formation of *S. nigrum*, utilizing multi-locus (SCoT, ISSR) markers along with chloroplast *trnL-F* sequence data and morphological characters. We aimed to investigate: (1) whether relationships of these species will be better represented as a reticulate rather than tree-like, and (2) by using a network method we aimed to study relationships between these species at genome level.

**Materials and methods**

Plant material and DNA extraction

A total of 48 accessions of *S. nigrum* related, diploid (*S. americanum*, *Solanum physalifolium*, *Solanum che nopodioides*), tetraploid (*Solanum retroflexum*, *Solanum villosum*) and hexaploid (*Solanum opacum*, *S. scabrum*) taxa were used in this study. Although, almost 50 taxa belong to sect. *Solanum*, we have chosen to analyze the most relevant species which have possibly taken part in the formation of *S. nigrum* as implicated in previous studies [44, 55, 58, 62]. Identification of the species followed Edmonds and Chweya [2]. We recognize *S. americanum*
and *S. nodiflorum* here as separate taxa after Manoko et al. [62]. *S. americanum* is an aggregate species but detailed analysis of this complex is outside the scope of the present study. One *S. dulcamara* entry was used as an outgroup following the results by Manoko et al. [62, 63]. Details about the plant material used in this study are listed in Table 1.

DNA was extracted according to the following protocol: 50 mg of fresh young leaves were suspended in 1 ml of Solution I (50 mM Tris–HCl, 50 mM EDTA, 500 mM NaCl, 15% sucrose, pH 8.0). The tissues were crushed and homogenized, for complete cell lyses and disruption, using a mixer mill (Retsch MM 301, Germany) and sterile steal beads for 10 min, 30 s. The lysated samples were centrifuged at room temperature for 5 min at 8,000 rpm. The supernatant was discarded and the remaining pellet was dissolved in 300 μl of 20T-10E buffer (20 mM Tris–HCl, 10 mM EDTA, pH 8.0) and 20 μl 10% SDS solution was added. The samples were incubated for 15 min at 70°C and gently shaken a few times. After the incubation 150 μl of 7.5 M sodium acetate was added, and the samples were placed on ice for 1 h and centrifuged for 10 min, 15,000 rpm at 4°C. The flow through (approx. 400 μl) was transferred to new Eppendorf tubes and 500 μl of ice-cold isopropanol was added to precipitate DNA. The samples were kept at 4°C for 30 min and then centrifuged for 5 min, 8,000 rpm at 4°C. The flow through was removed; DNA was recovered as a pellet and dissolved in 500 μl of TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) equal volume of CIA (chloroform:isoamyl-alcohol, 24:1) was added and the samples were shaken for 15 min at 300 rpm and then centrifuged for 15 min, 15,000 rpm at room temperature. The upper phase was transferred to a new Eppendorf. 1 ml of 99.9% ethanol and 40 μl of ammonium acetate was added to the samples and then centrifuged for 10 min, 15,000 rpm at 4°C. The supernatant was removed and the pellet was washed again with 500 μl of 70% ethanol and centrifuged for 5 min, 15,000 rpm at 4°C. The tubes were placed upside down and air dried. Finally the DNA was dissolved in 200 μl TBE (Tris–Borate-EDTA pH 8.0) and stored at −20°C. The DNA concentration was measured with fluorodensitometric analysis (GeneTools, Syngene UK) after electrophoresis and ethidium-bromide post-staining using the GeneRuler 1 kb Ladder (Fermentas, Lithuania) separated on 1% agarose gel.

Morphological analysis

We compiled a matrix of 51 morphological characters measured and coded as multistate characters. Forty of the characters were qualitative and 11 were quantitative traits. Characters were taken and measured from approximately the same position of each plant. Leaf characters were taken from the fourth true leaf of each individual while floral characters from the uppermost inflorescence. Four individuals from each accession were measured. The mean values of these plants and accessions were used as representatives of each taxon. Data collected here was compared with previous descriptions and studies on section *Solanum* [2, 20, 40]. The list and the matrix of the morphological characters are presented in Online Resources 1 and 2.

ISSR and SCoT amplification

We chose arbitrary amplified DNA (AAD) markers for our study to generate fragments that are randomly generated by inter-simple sequence repeats (ISSR) or start codon targeted (SCoT) over the whole genome [71]. Recent studies [72–75] have shown that AAD markers can solve phylogenetic relationships of closely related, recently radiated taxa at low taxonomic levels [76–78]. However, one of the arguments against the use of AADs is that they are homoplastic—co-migration of non-identical bands—causing noise instead of phylogenetic signal in the datasets [79, 80]. The species of the sect. *Solanum* are assumed to be very closely related and homoplasy becomes a greater problem where distantly related species are involved; it is less likely to cause problems for studies of very closely related species [72, 81]. This assumption certainly applies to *Solanum* where the utility of multi-locus methods in phylogenetic reconstruction have repeatedly been used at species level [82–86].

Sixteen ISSR and 12 SCoT primers were selected for further study, after a screening and optimization process with 50 primers for each method. Degenerate nucleotide anchoring was used for some ISSR primers. The selected primers yielded stable and reproducible banding patterns. Amplification reactions were performed in 10 μl volumes in 384-well plates containing: 5 μl NFW (Nuclease Free Water, Promega, USA), approx. 20 ng template DNA, 0.5 μM of each primer, 0.2 mM dNTP (Fermentas, Lithuania), 1 μl 10× PCR buffer (1 mM Tris–HCl, pH 8.8 at 25°C, 1.5 mM MgCl₂, 50 mM KCl and 0.1% Triton X-100) and 0.5 U of DyNazyme II (Finnzymes, Finland) polymerase. All reactions were done in a MasterCycler ep384 (Eppendorf, Germany) with the following conditions: 2 min at 94°C for initial denaturation, 35 cycles of 30 s denaturation at 94°C, 1 min annealing at 50°C, and 2 min extension at 72°C, followed by a final extension for 5 min at 72°C. Amplification products were separated on 1.5% agarose gels (Promega, USA) in 0.5× TBE buffer (300 V, 1.5 h) and post-stained with ethidium–bromide. The gels were documented using the GeneGenius Bio Imaging System (Syngene, UK). Although, ISSRs and SCoTs were found to be reproducible in our experiments—due to higher annealing temperature and longer anchored primers—replicate
<table>
<thead>
<tr>
<th>No.</th>
<th>Accession number</th>
<th>Accession code</th>
<th>Taxon</th>
<th>Received</th>
<th>Origin</th>
<th>Voucher a</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HUGEO09022</td>
<td>Ame01</td>
<td>S. americum Mill.</td>
<td>G</td>
<td>Australia</td>
<td>UPG0001</td>
</tr>
<tr>
<td>2</td>
<td>HUGEO09020</td>
<td>Ame02</td>
<td>S. americum Mill.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0002</td>
</tr>
<tr>
<td>3</td>
<td>HUGEO09023</td>
<td>Ame03</td>
<td>S. americum Mill.</td>
<td>G</td>
<td>Australia</td>
<td>UPG0003</td>
</tr>
<tr>
<td>4</td>
<td>904750023</td>
<td>Ame04</td>
<td>S. americum Mill.</td>
<td>N</td>
<td>USA</td>
<td>UPG0004</td>
</tr>
<tr>
<td>5</td>
<td>HUGEO09024</td>
<td>Ame05</td>
<td>S. americum Mill.</td>
<td>G</td>
<td>Australia</td>
<td>UPG0005</td>
</tr>
<tr>
<td>6</td>
<td>HUGEO09021</td>
<td>Ame06</td>
<td>S. americum Mill.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0006</td>
</tr>
<tr>
<td>7</td>
<td>884750042</td>
<td>Che01</td>
<td>S. chenopodioides Lam.</td>
<td>N</td>
<td>Switzerland</td>
<td>UPG0007</td>
</tr>
<tr>
<td>8</td>
<td>HUGEO09025</td>
<td>Che02</td>
<td>S. chenopodioides Lam.</td>
<td>G</td>
<td>France</td>
<td>UPG0008</td>
</tr>
<tr>
<td>9</td>
<td>UHBG211-1470</td>
<td>Che03</td>
<td>S. chenopodioides Lam.</td>
<td>H</td>
<td>Germany</td>
<td>UPG0009</td>
</tr>
<tr>
<td>10</td>
<td>HUGEO09005</td>
<td>Che04</td>
<td>S. chenopodioides Lam.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0010</td>
</tr>
<tr>
<td>11</td>
<td>HUGEO09026</td>
<td>Che05</td>
<td>S. chenopodioides Lam.</td>
<td>G</td>
<td>Romania</td>
<td>UPG0011</td>
</tr>
<tr>
<td>12</td>
<td>HUGEO09027</td>
<td>Che06</td>
<td>S. chenopodioides Lam.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0012</td>
</tr>
<tr>
<td>13</td>
<td>HUGEO09004</td>
<td>Nig01</td>
<td>S. nigrum¹ L.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0013</td>
</tr>
<tr>
<td>14</td>
<td>HUGEO06005</td>
<td>Nig02</td>
<td>S. nigrum² L.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0014</td>
</tr>
<tr>
<td>15</td>
<td>HUGEO06006</td>
<td>Nig03</td>
<td>S. nigrum² L.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0015</td>
</tr>
<tr>
<td>16</td>
<td>HUGEO09028</td>
<td>Nig04</td>
<td>S. nigrum² L.</td>
<td>G</td>
<td>Croatia</td>
<td>UPG0016</td>
</tr>
<tr>
<td>17</td>
<td>HUGEO09029</td>
<td>Nig05</td>
<td>S. nigrum² L.</td>
<td>G</td>
<td>Romania</td>
<td>UPG0017</td>
</tr>
<tr>
<td>18</td>
<td>HUGEO09030</td>
<td>Nig06</td>
<td>S. nigrum¹ L.</td>
<td>G</td>
<td>Italy</td>
<td>UPG0018</td>
</tr>
<tr>
<td>19</td>
<td>884750223</td>
<td>Opa01</td>
<td>S. opacum A.Braun&amp;Bouché</td>
<td>N</td>
<td>Unknown</td>
<td>UPG0019</td>
</tr>
<tr>
<td>20</td>
<td>HUGEO09015</td>
<td>Opa02</td>
<td>S. opacum A.Braun&amp;Bouché</td>
<td>G</td>
<td>Australia</td>
<td>UPG0020</td>
</tr>
<tr>
<td>21</td>
<td>HUGEO09016</td>
<td>Opa03</td>
<td>S. opacum A.Braun&amp;Bouché</td>
<td>G</td>
<td>Australia</td>
<td>UPG0021</td>
</tr>
<tr>
<td>22</td>
<td>HUGEO09017</td>
<td>Opa04</td>
<td>S. opacum A.Braun&amp;Bouché</td>
<td>G</td>
<td>Australia</td>
<td>UPG0022</td>
</tr>
<tr>
<td>23</td>
<td>HUGEO09018</td>
<td>Opa05</td>
<td>S. opacum A.Braun&amp;Bouché</td>
<td>G</td>
<td>Australia</td>
<td>UPG0023</td>
</tr>
<tr>
<td>24</td>
<td>HUGEO09019</td>
<td>Opa06</td>
<td>S. opacum A.Braun&amp;Bouché</td>
<td>G</td>
<td>Australia</td>
<td>UPG0024</td>
</tr>
<tr>
<td>25</td>
<td>894750076</td>
<td>Phy01</td>
<td>S. physalifolium³ Rusby</td>
<td>N</td>
<td>Germany</td>
<td>UPG0025</td>
</tr>
<tr>
<td>26</td>
<td>HUGEO09010</td>
<td>Phy02</td>
<td>S. physalifolium³ Rusby</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0026</td>
</tr>
<tr>
<td>27</td>
<td>HUGEO09011</td>
<td>Phy03</td>
<td>S. physalifolium³ Rusby</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0027</td>
</tr>
<tr>
<td>28</td>
<td>HUGEO09012</td>
<td>Phy04</td>
<td>S. physalifolium³ Rusby</td>
<td>G</td>
<td>Romania</td>
<td>UPG0028</td>
</tr>
<tr>
<td>29</td>
<td>HUGEO09013</td>
<td>Phy05</td>
<td>S. physalifolium³ Rusby</td>
<td>G</td>
<td>Romania</td>
<td>UPG0029</td>
</tr>
<tr>
<td>30</td>
<td>HUGEO09014</td>
<td>Phy06</td>
<td>S. physalifolium³ Rusby</td>
<td>G</td>
<td>Slovakia</td>
<td>UPG0030</td>
</tr>
<tr>
<td>31</td>
<td>PI6347502SD</td>
<td>Ret01</td>
<td>S. retroflexum Dunal</td>
<td>U</td>
<td>USA</td>
<td>UPG0031</td>
</tr>
<tr>
<td>32</td>
<td>HUGEO09006</td>
<td>Ret02</td>
<td>S. retroflexum Dunal</td>
<td>G</td>
<td>USA</td>
<td>UPG0032</td>
</tr>
<tr>
<td>33</td>
<td>HUGEO09007</td>
<td>Ret03</td>
<td>S. retroflexum Dunal</td>
<td>G</td>
<td>Australia</td>
<td>UPG0033</td>
</tr>
<tr>
<td>34</td>
<td>HUGEO09009</td>
<td>Ret04</td>
<td>S. retroflexum Dunal</td>
<td>G</td>
<td>South Africa</td>
<td>UPG0034</td>
</tr>
<tr>
<td>35</td>
<td>904750228</td>
<td>Ret05</td>
<td>S. retroflexum Dunal</td>
<td>N</td>
<td>Australia</td>
<td>UPG0035</td>
</tr>
<tr>
<td>36</td>
<td>HUGEO09008</td>
<td>Ret06</td>
<td>S. retroflexum Dunal</td>
<td>G</td>
<td>Australia</td>
<td>UPG0036</td>
</tr>
<tr>
<td>37</td>
<td>824750011</td>
<td>Sca01</td>
<td>S. scabrum Mill.</td>
<td>N</td>
<td>Unknown</td>
<td>UPG0037</td>
</tr>
<tr>
<td>38</td>
<td>Grif1419801SD</td>
<td>Sca02</td>
<td>S. scabrum Mill.</td>
<td>U</td>
<td>USA</td>
<td>UPG0038</td>
</tr>
<tr>
<td>39</td>
<td>HUGEO09004</td>
<td>Sca03</td>
<td>S. scabrum Mill.</td>
<td>G</td>
<td>Uganda</td>
<td>UPG0039</td>
</tr>
<tr>
<td>40</td>
<td>UHBG211-1465</td>
<td>Sca04</td>
<td>S. scabrum Mill.</td>
<td>H</td>
<td>Germany</td>
<td>UPG0040</td>
</tr>
<tr>
<td>41</td>
<td>HUGEO09002</td>
<td>Sca05</td>
<td>S. scabrum Mill.</td>
<td>G</td>
<td>South Africa</td>
<td>UPG0041</td>
</tr>
<tr>
<td>42</td>
<td>HUGEO09003</td>
<td>Sca06</td>
<td>S. scabrum Mill.</td>
<td>G</td>
<td>South Africa</td>
<td>UPG0042</td>
</tr>
<tr>
<td>43</td>
<td>HUGEO09031</td>
<td>Vil01</td>
<td>S. villosum⁴ L.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0043</td>
</tr>
<tr>
<td>44</td>
<td>HUGEO09032</td>
<td>Vil02</td>
<td>S. villosum⁴ L.</td>
<td>G</td>
<td>Romania</td>
<td>UPG0044</td>
</tr>
<tr>
<td>45</td>
<td>HUGEO09033</td>
<td>Vil03</td>
<td>S. villosum⁴ L.</td>
<td>G</td>
<td>Hungary</td>
<td>UPG0045</td>
</tr>
<tr>
<td>46</td>
<td>804750186</td>
<td>Vil04</td>
<td>S. villosum⁴ L.</td>
<td>N</td>
<td>Unknown</td>
<td>UPG0046</td>
</tr>
<tr>
<td>47</td>
<td>HUGEO09034</td>
<td>Vil05</td>
<td>S. villosum⁴ L.</td>
<td>G</td>
<td>Unknown</td>
<td>UPG0047</td>
</tr>
</tbody>
</table>
experiments were performed, containing one negative and positive control, to check the reliability of the primers and the patterns produced.

**Chloroplast trnL-F amplification, cloning and sequencing**

PCR amplifications and treatment of the trnL-F chloroplast region followed the procedures described by Bohs and Olmstead [49]. Primers described by Taberlet et al. [87] were used in all reactions. For the cloning procedure, the amplified products were excised from the agarose gel and cleaned using the SpinPrep Gel DNA Kit (Novagen, Germany). The fragments were cloned using the pGem-T Vector System (Promega, USA) and DH5α competent *Escherichia coli* strains. The procedure was carried out according to the manufacturer’s protocols. The desired white colonies holding the insertion were selected and screened applying a colony-PCR. For the procedure white colonies were picked up and used directly as DNA template in the amplification process. For the procedure white colonies were picked up and used directly as DNA template in the amplification process. The cycling parameters were: 94°C for 2 min, followed by 35 cycles of 94°C, 30 s annealing, 72°C 1 min extension, and one additional extension cycle at 72°C for 1 min. Plasmids were extracted from the selected white colonies using the EZ-10 Spin Column Plasmid DNA Kit (Bio Basic Inc., Canada). Sequencing was done using the T7 and pUC/M13 (5’-CAGGAAACAGCTATGAC-3’) reverse primers, in an ABI3100 for three selected colonies. All sequences are deposited at NCBI GenBank. Accession numbers and further details are given in Table 2.

### Band scoring and data analysis

The evaluation of the binding patterns was carried out with the program GeneTools (Syngene, UK). Kingston and Rosel [88] described a conservative scoring protocol that was used also here to prevent problems associated with multi-locus methods, e.g.: uneven amplification among samples and poor amplification of larger fragments for degraded DNA samples. Only well-resolved, distinct bands were scored. Amplicons found in replicate reactions were considered reliable. The amplified fragments were coded as absent/present (0/1). It was presumed that fragments with equal length had been amplified from homologous loci and represents a single, dominant locus with two possible alleles. To measure the information content detected with each primer the Polymorphic Information Content (PIC) value was calculated, according to Botstein et al. [89]. The heterozigosity (H) value was also calculated according to Liu [90]. For all calculation the test version of the online program PICcalc was used [91]. As PIC and H are both influenced by the number and frequency of alleles, the maximum number for a dominant marker is 0.5, since two alleles per locus are assumed in the analysis [92–94]. For all data about the primers and calculated values see Online Resource 3.

### Table 2 Accession numbers for chloroplast trnL-F sequences used in this study

<table>
<thead>
<tr>
<th>Species name</th>
<th>Sequence accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. americanum</em></td>
<td>GU169332</td>
</tr>
<tr>
<td><em>S. chenopodioides</em></td>
<td>GU169331</td>
</tr>
<tr>
<td><em>S. nigrum</em></td>
<td>AY349002</td>
</tr>
<tr>
<td><em>S. opacum</em></td>
<td>DQ855059</td>
</tr>
<tr>
<td><em>S. physalifolium</em></td>
<td>EF068253</td>
</tr>
<tr>
<td><em>S. retroflexum</em></td>
<td>AY348999</td>
</tr>
<tr>
<td><em>S. scabrum</em></td>
<td>GU169330</td>
</tr>
<tr>
<td><em>S. villosum</em></td>
<td>DQ180459</td>
</tr>
<tr>
<td><em>S. dulcamara</em></td>
<td>AY266231</td>
</tr>
</tbody>
</table>

#### Note

The sequences will be published by NCBI GenBank after acceptance New sequence provided by this study are in bold
Phylogenetic network analysis

The network analysis was carried out to study whether reticulating or tree-like evolutionary graphs would be better representations of the phylogeny for the selected set of terminals. A NeighborNet (NNet) analysis was carried out with SplitsTree v. 4.6 [95]. We conducted a NNet analysis with the separate ISSR and SCoT dataset and also with the joined dataset. All calculations were based on uncorrected p-distances where splits (bipartitions) were represented as parallel edges. The resulting splits graphs were filtered with a 95% threshold level to include only the splits with high support values.

We also constructed a supernetwork from the collection of trees resulting from the parsimony analysis of the combined ISSR and SCoT dataset using the supernetwork construction function of SplitsTree. The congruence or incongruence of the resulting trees and graphs was used to infer recombination, which was tested using the Phi-test implemented in SplitsTree. This test uses the pairwise homoplasy index, \( \Phi \) statistic, to detect refined incompatibility indicating recombination [96]. This was done with the filtered taxon set including only the analyzed diploid species (\( S. \) americanum, \( S. \) chenopodioides, \( S. \) physalifo- lium) to detect splits resulting from false negatives.

Simultaneous analyses

ClustalX version 1.83.1 [97] with default settings was used to align \( trnL-F \) sequences. Last 24 nucleotides by 3’ prime for \( S. \) nigrum, \( S. \) retroflexum and \( S. \) villosum, and 12 nucleotides for \( S. \) chenopodioides and \( S. \) scabrum were removed as unalignable with other species. Conventional phylogenetic analyses were performed using parsimony as an optimality criterion. These were performed with NONA [98] in conjunction with a Winclada shell [99]. Fifty-one morphological characters were united with SCoT (171), ISSR (224) and \( trnL-F \) (1042 characters, the latter converted into numeric form using winclada) data to a concatenated matrix of 49 terminals and 1,488 characters. Three hundred and eighty-three of these were parsimony informative (34 morphology, 142 SCoT, 200 ISSR and 6 \( trnL-F \)). While there were several accessions for each species for SCoT and ISSR, both morphology and \( trnL-F \) were scored as identical for all terminals of the same species (an assumption that can be challenged). We analyzed also a reduced matrix so that each species was represented by the least divergent (based on the results of larger matrix) terminal in order to see if the same topology for species level relationships was obtained also in this case. Analyses of larger matrix were performed with the following settings: hold 90,000 (holding defined number of trees), 30,000 replications (search performed with multiple tree-bisection-reconnection algorithm mult*max*), hold/3 (keeping three starting trees for each replication). Analysis was further enlarged by holding up to 180,000 trees (hold 180,000) and keeping 60 starting trees for each replication (hold/60). Also a small, restricted analysis on the large matrix was performed with the settings hold 100, 50 replications (mult*max*), hold/1. A smaller matrix with nine terminals (based on the analyses of the larger matrix the terminal in the basal position within each species clade) and 383 informative characters was analyzed with TNT version 1.1 [100]. Because of its small size we were able to use a command (jenum) that ensures finding the shortest tree for this matrix. The assembled data matrix is given in Online Resource 4.

Results

Network analysis

The snapshot of the combined ISSR and SCoT dataset (Fig. 1) obtained by NNet analysis provided a general representation of the genetic relationships of the terminals. The planar property of the NNet algorithm—which is an explicit extension of Neighbor-Joining [101]—represents conflicting signals in the dataset. In addition, NNet split graphs highlight the predominant signals in the data and the extent to which these signals may or may not be tree-like [70]. Reticulate evolutionary history is displayed as the occurrence of incompatible splits or parallel edges in the resulting split graphs. It is often reported that taxa formed by reticulate evolutionary patterns or hybridizations are linked by splits to their potential parents [102]. However, the constructed graphs by NNet algorithm display only the contradictory signals which are visualized in a planar graph. These graphs can not be considered explicitly as a model of reticulate evolution although it was recently shown that there is a strong correlation between split graphs and reticulate history [103–106].

In this study, all the accessions of the diploid species shared a split with all of the polyploid species. This split reflects a high portion of shared ISSR and SCoT bands between diploids and polyploids. Another major split divides \( S. \) nigrum (6\( \times \)) and \( S. \) villosum (4\( \times \)) together with \( S. \) scabrum (6\( \times \)). In addition, a strong split also occurs—dividing the diploid species—and linking them together with the polyploids in a different context. The accessions of \( S. \) americanum (2\( \times \)) share a weak split with \( S. \) nigrum, \( S. \) villosum and \( S. \) scabrum (Fig. 1) and they form a circular ‘splits rose’ (Fig. 2). A clear parallel edge links the terminals of \( S. \) chenopodioides (2\( \times \)) and \( S. \) physalifolium (2\( \times \)) with \( S. \) retroflexum (4\( \times \)) and \( S. \) opacum (6\( \times \)). While the constructed supernetwork from the eight equally
parsimonious trees obtained for the joined ISSR&SCoT data was almost tree like (Fig. 3).

We also made a more specific analysis by excluding the terminals of other species to study the evolutionary history of *S. nigrum* (Fig. 2). This was done in order to test hypothesis based on previous studies, e.g. [24, 107, 108]. The separate Phi test of the filtered diploid taxon set with the joined ISSR&SCoT data matrix found 222 informative sites, but did not find statistically significant evidence for recombination ($P = 0.5081$).

**Simultaneous analyses**

All performed analyses of the combined large matrix, even the one with only 100 replicates, obtained the same set of eight equally parsimonious trees (EPTs). The relationships of the terminals at species level on all of these trees were the same and can be presented in parenthetical notation as: *(S. dulcamara (S. americanum (S. chenopodioides *S. physalifoillum)) ((S. retroflexum *S. opacum) (S. nigrum (S. scabrum *S. villosum)))). The only difference obtained with the analysis of smaller reduced matrix was in the position of *S. opacum*—it was in the basal position in the clade including *S. nigrum, S. scabrum* and *S. villosum*, not as a sister of *S. retroflexum*.

**Discussion**

Our results could be challenged by larger analyses, both in terms of characters sampled and terminals included. The results obtained can be interpreted as a mere experimentation with the chosen methods. At this point we can present only tentative and preliminary interpretation of the genome evolution and origin of polyploid taxa. Despite of the limited taxonomic sampling there is a strong congruence with phylogenies obtained in previous molecular marker analyses using different datasets [39, 40] and broader sampling [6, 54–58, 62, 63]. In addition to the difficult interpretation of the observed morphological variation, genome relationships among the diploid and polyploid taxa are still a matter of debate. Despite of the great

**Fig. 1** NeighborNet planar graphs of the combined ISSR and SCoT binary data matrix based on the uncorrected $p$-distance for sect. *Solanum* and *S. dulcamara* as outgroup. NNet for 395 polymorphic loci (ISSR and SCoT) and 49 accessions, based on uncorrected $p$-distances, filtered with 0.95 threshold level. Abbreviations are listed in Table 1. Relevant splits separating the diploids and polyploids are highlighted. Red color splits separating *S. americanum, S. chenopodioides* and *S. physalifoillum* var. *nitidibaccatum*; blue splits separating *S. nigrum, S. scabrum* and *S. villosum*; green splits separating *S. retroflexum* and *S. opacum*. Further details are discussed in the text.
efforts to resynthesize polyploids using the combination of different crosses of species, e.g. [107, 109, 110], the exact origin and phylogeny of even the most widespread and common species like *S. nigrum*, have remained ambiguous. The first hybridization studies with early seed protein electrophoresis experiments [111] provided the preliminary hypothesis of the origin and the idea of complex reticulating events as reasonable scenarios of evolutionary history for some polyploids. It now seems clear that the complexity of this widespread and important group is based on polyploids.

On the origin of *S. nigrum*: taxonomic implications

The first studies on genomic relationships started with the identification of the possible progenitors of *S. nigrum*, which was referred as an autoallohexaploid [112] or autohexaploid [107, 113] and in other studies as an allohexaploid [25] species. The confusion around the origin of *S. nigrum* could be attributed to its meiotic chromosomal behavior. The phenomenon that most species of the section show regular bivalent formation at meiosis or somatic chromosome doubling led early cytological workers to contradictory conclusions [24, 113–115]. Under a standard evolutionary model, the split graph reveals a reticulate evolutionary history of this species. In previous molecular studies based on different methods involving variable species a close relationship between *S. scabrum*, *S. nigrum* and *S. villosum* have been proposed [44, 58, 62], but contra Olet [40]. However, in a later parsimony analysis by Manoko [7] many of the species relationships were left as unresolved in the jackknife consensus tree. Although, it might be possible that the names are misapplied there, since what is referred by Manoko et al. [6] as *S. americanum* is probably *S. ptycanthum* from North America, and what is identified as *S. nodiflorum* is might be *S. americanum*.

In our simultaneous analysis *S. villosum* turned out to be a sister species of *S. scabrum*. The split graph constructed

---

**Fig. 2** NNet specific analysis of evolutionary history of *S. nigrum* with other terminals excluded, inferred from the combined ISSR and SCoT data based on uncorrected p-distances, with 0.95 threshold level. The inlet, a ‘splits rose’ illustrates the competing relationships and supposed reticulate history of *S. nigrum*. Major splits highlighted with blue and red color indicate parallel edges structure in a ‘box form’ indicating incongruence in the collected splits (bipartitions).
from multi-locus binary data supports *S. nigrum*, *S. scabrum* and *S. villosum* as closely related. This might indicate that the tetraploid *S. villosum* contributed to both of the mentioned hexaploid taxa. The strong split (Fig. 1, highlighted in blue) combining the three taxa might indicate that they all share the same diploid ancestor. The terminals of these four species in exclusion of other taxa were represented in a separate analysis (Fig. 2) because they were assumed to share complex reticulate history. In both split graphs (Figs. 1 and 2) the tetraploid *S. villosum* was linked only with one diploid species *S. americanum*, suggesting its autotetraploid origin. These results confirm the fluorescent chromosome banding study of Sultana and Alam [116], where *S. villosum* was shown to be an ancient tetraploid of *S. americanum* s.l., which in the course of its evolution had established regular meiosis (bivalent formation). Rao [112] showed that from the two of the three sets of *S. nigrum* chromosomes are homologous with each other, and possibly originated from *S. villosum*. The remaining set differs from these two. As presented here *S. nigrum* possibly originated from diploid species *S. americanum* s.l. and tetraploid *S. villosum* through amphiploidy of sterile triploid progeny as previously presented [27, 107, 108, 117–120]. Whether it can be coined as an autoallopolyploid depends on the degree of differentiation of the *S. villosum* genome as compared to the remaining set possibly originated from *S. americanum* s.l. If this is not the case, and only minute changes have been rearranged it could be regarded as a segmental allopolyploid. This needs to be verified by a precise molecular cytogenetic analysis. It is also evident that *S. americanum* s.l. provided its genome to *S. scabrum* too, as suggested by Ganaphati and Rao [121] based on the vigor of sterile hybrids derived from crossings. The crossing experiments carried out by Heiser et al. [122] also led to the same conclusions, alongside with other recent studies based on molecular markers [7, 39, 41, 44].

The separation of the other tetraploid taxa *S. retroflexum* with the hexaploid *S. opacum* in a well established group which is connected with parallel edges (Fig. 1 highlighted with green) to diploids *S. chenopodioides* and *S. physalifoillum* is also interesting. The separation of diploids with one clear split confirms the conclusion by Manoko et al.
uncorrected splits (exceeding the 0.95 threshold level) based on obtained NNet topologies could be the rearrangement of within and among the species.

Does not include all geographical regions of diversity from those obtained here. Our sampling is also restricted: it is unrealistic to expect to be able to build a reticulate phylogeny [130]. Accordingly, application of such multi-locus markers like ISSRs and SCoTs could produce incorrect genetic distances depending on the degree of genome rearrangements. As a result, genetic distances between hypothetical parental diploids and their derived allopolyploids estimated by PCR-based multi-locus banding patterns will increase. As in this case polyploids like S. nigrum and possibly other members of section Solanum have undergone substantial genomic rearrangements in their genome. Therefore, other methods (e.g. sequence based approaches) should be used for the detection of...
polyploid origin in the future studies. Phylogenetic analyses based on multiply sequences, combined with filtered supernetwork reconstruction have been used to detect incongruences and reticulations within larger datasets in different organisms up to class level [131]. Splits graphs and network based approaches have the potential to reveal reticulate history of black nightshades (sect. Solanum), but different methods should be chosen for further research, and the use of multi-locus markers should be avoided. An alternative to this would be the use of the complex nature of the rDNA locus (ITS, ETS) as suggested by Poczai and Hyvönen [132]. Accordingly, this region seems to be a warranted choice since it has been successfully utilized by Komarova et al. [126] together with NNet approaches to investigate the hybrid origin of Solanum sect. Petota.

Conclusions

Hybridization seems to be an important phenomenon in sect. Solanum and history of S. nigrum might be reticulate. Based on recent studies, e.g. [133], ancestors of Solanaceae and Solanum have gained one whole genome duplication (WGD) early in their evolutionary history, then the potato clade has gone through another [134]. S. nigrum represents a case where possibly ancient tetraploids and diploids formed a hexaploid species and these hybrids managed to exploit better their new habitat. In this case genetically differentiated diploids in a few combinations have created a series of genetically distinct polyploid populations. The gene flow between these new polyploids has not always been restricted, because different independent combinations have also arisen in the evolutionary history (e.g. S. opacum). The insufficient isolation that permitted further recombination between ancient polyploids and diploids have resulted in high level of genotypic and phenotypic polymorphism. This high level of novel genomic variability enabled species to succeed in their new environment (e.g. Africa, Asia or Australia). It is also interesting to note that both allo- and autopolyploidy, or the combination of these together with recurrent polyploidization, have also acted during the evolution and formation of S. nigrum and possibly in sect. Solanum too. It is possible that polyploids are formed over and over again subsequently from different parental genotypes and generate diverse array of species series. Hybridization among the polyploids has obviously also occurred and resulted in additional genetic variability. This feature of plants is not restricted to Solanum as it is well known that polyploids arose recurrently during flowering plant evolution [66].

In the future, intergenomic translocations—if present—has to be identified in this group as previously proposed in allopolyploid species of Nicotiana of the same family [127]. These studies involving chromosome painting and cytogenetic tools have potential to shed light on the complex processes that has acted during the evolutionary history of Solanum and help to elucidate the precise origin of polyploids. These kind of studies will be highly valuable for practical research programs (e.g. resistance genes and other important traits) aiming to utilize the genetic resources of this group. Despite of the complexity and possible multiple origins of the polyploids the group is still a rich source of unexploited new traits which should be utilized in the further improvement of the crop plants of Solanaceae. The difficult problem of completely unraveling the composite history of the section remains, although it is demonstrated that the application of recently developed phylogenetic network methods could provide the valuable first step towards the better understanding of its complex evolutionary history.

Acknowledgements We thank the staff of the Potato Research Center, University of Pannonia for their work; and several Botanical Gardens, collectors for the plant material. A special thank is due to the Staff of the Biotechnology Group, University of Pannonia for general assistance and facilities. This study was supported by the CIMO Research Fellowship Grant and by the MÖBCIMO Finnish Government’s Scholarship provided for the first author. The authors also thank Dr. Jennifer Edmonds for comments regarding this study.

References

2. Edmonds JM, Chweya JA (1997) Black nightshades. Solanum nigrum L. and related species. Promoting the conservation and use of underutilized and neglected crops, vol 15. International Plant Genetic Resources Institute, Rome, Italy


© Springer
30. Linnaeus C (1753) Species plantarum, vol 2. Stockholm
amplified fragment length polymorphism genotyping. PNAS 102:14694–14699
91. PICTcalc an online program for calculating Polymorphic Information Content (PIC). http://www.georgikon.hu/pic/english
95. SPLITSTREE v.4.6. http://www.splitstree.org
