Molecular characterization of cowpea [Vigna unguiculata (L.) Walp.] subspecies with SSR markers

Allonoumi J. C. Quenum
Université Cheikh Anta Diop de Dakar: Universite Cheikh Anta Diop de Dakar
Remy S. Pasquet
Institut de recherche pour le développement: Institut de recherche pour le developpement
Amy Bodian
Institut Sénégalais de Recherches Agricoles: Institut Senegalais de Recherches Agricoles
Daniel Fonceka
CIRAD
Yvette R. Djiboune
Institut Sénégalais de Recherches Agricoles: Institut Senegalais de Recherches Agricoles
Ndiaga Cisse
Institut Sénégalais de Recherches Agricoles: Institut Senegalais de Recherches Agricoles
Mame S. Mbaye
Université Cheikh Anta Diop de Dakar: Universite Cheikh Anta Diop de Dakar
Diaga Diouf
diaga.diouf@ucad.edu.sn

Universite Cheikh Anta Diop https://orcid.org/0000-0002-3262-1763

Research Article

Keywords: cowpea, wild relatives, SSR, phylogenetics

Posted Date: April 14th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2687516/v1

Version of Record: A version of this preprint was published at Genetic Resources and Crop Evolution on September 21st, 2023. See the published version at https://doi.org/10.1007/s10722-023-01738-x.

Abstract

Background Cowpea, *Vigna unguiculata*, is an important food legume in the tropics and subtropics. However, cowpea is a complex species with more than 10 subspecies which can hybridize and produce intermediate progenies. Partly because of the complex organization of the cowpea gene pool and the lack of adequate markers for these infraspecific units, cowpea breeders are not using the wild part of the cowpea gene pool.

Methods Here, we report the molecular characterization of 34 representative accessions with 61 SSR markers from coding regions.

Results If SSR failed to separate the closest groups, i.e. subsp. *alba*, subsp. *tenuis* and the perennial groups from subsp. *unguiculata*, a combination of few SSR markers can properly identify the main cowpea subspecies. Regarding the infraspecific phylogeny of cowpea, SSR markers confirm the special status of the annual subsp. *unguiculata* versus the different perennial subspecies. It confirms that subsp. *protracta* looks like the oldest subspecies, making the origin of the species in southern Africa likely. However, it unites in a single group, clearly separated from subsp. *unguiculata*, all the taxa of hybrid origin, i.e. subsp. *alba*, subsp. *tenuis*, subsp. *pubescens* and the BWA group of subsp. *unguiculata*.

Conclusion Although a limited number of markers was tested considering that several hundred of cowpea SSR are available, the present work shows that SSR markers can be used for molecular characterization of cowpea subspecies and can be very helpful for understanding its complex evolutionnary history.

Introduction

Cowpea, *Vigna unguiculata* (L.) Walp., is an important food legume in the tropics and subtropics, especially in sub-Saharan Africa where it is grown for its seeds as well as for its fodder [1].

Cowpea gene pool organization is fairly complex, with numerous subspecies, including some of hybrid origin [2]. This complexity may explain why genetic resources from the wild gene pool have never been used in breeding [1]. Indeed, if the different subspecies are morphologically well identified, there are numerous accessions which are intermediate or introgressed in various ways [2]. Therefore, it would be helpful to identify molecular markers that can confirm morphological identifications or detect, qualify, eventually quantify, introgressions in some plants or accessions. Since DNA barcoding is used for the species identification of an organism by comparing some DNA sequences with those from a database [3], a kind of barcoding of the different cowpea subspecies is definitely needed in order to help the cowpea research community.

However, if Internal transcribed spacer (ITS) and Maturase K *matK* are widely used in plant barcoding [3], both showed insufficient variability within cowpea and ITS variants can be encountered within a single cowpea plant [4]. In addition, compared to animals, polyploïdy, apomixy, and hybridization events make a double barcoding (chloroplastic and nuclear) a necessity in plants [5, 6]. This is especially the case with

cowpea since numerous chloroplast capture events were detected between cowpea infraspecific groups [2].

Within cowpea, cpDNA-RFLPs are interesting markers that can characterize the different subspecies [2] but this technic is obsolete and tedious, limiting its use. While the interesting restriction site mutations can be converted into much more convenient SNP as Feleke et al. [7] did for the *BamHI* s13.3 mutation, more cpDNA mutations should be found after full sequencing of the chloroplast genome of the various subspecies.

However, regarding nuclear genome, there is no available marker for molecular characterization of cowpea subspecies. Recent molecular studies did not explore the cowpea genepool beyond the domesticated cowpea and its wild progenitor, i.e. subsp. *unguiculata* [e.g. 8–11]. The only work focusing on the wild gene pool is the one from Ogunkamni et al. [12] based on simple sequence repeats (SSR) but they did not try to charaterize the different subspecies. However, SSR were successfully used for the identification of close species in some complex taxonomic groups, e.g. *Psidium* (Myrtaceae), [13] *Rhododendron* [14] and the Mediterranean *Tamarix* [15].

Since several hundred of cowpea SSR are available, especially from functional regions [e.g. 16, 17], the objective of the present work is to prove that SSR can be used for molecular characterization of cowpea subspecies nuclear genome.

Materials And Methods

Plant materials

The plant materials consisted of 30 wild cowpea accessions provided by Meise Botanical Garden, Belgium (http://db.plantentuinmeise.be/RESEARCH/COLLECTIONS/LIVING/PHASEOLUS/index.html), 3 wild accessions from Senegal and the breeding line Melakh provided by ISRA. All the subspecies and taxonomic groups were represented, except subsp. *dekindtiana* (Harms) Verdc. *sensu stricto* from South Angola mountains, still absent from living collections (Table 1). Most of these accessions were already included in previous works [2, 7, 18] and MT and SP numbers used previously were kept instead of their equivalent four digits NI numbers from Meise Botanical Garden. *Vigna vexillata* (L.) A.Rich. NI 1014 was added as outgroup. Plants were grown in pots filled with sandy soil, without inoculation and watered with tap water twice a week.

DNA Isolation And Genotyping

DNA extraction, PCR, and electrophoresis methods followed exactly Sarr et al. protocol [8]. Considering the goal of our study, highly polymorphic SSR were discarded, especially the one showing polymorphism within subsp. *unguiculata* alone [e.g. 8, 19, 20]. A total of 61 SSR primers were selected and tested. The

SSR primers can be downloaded from the Cowpea Genomics Knowledge Base (CGKB) (http://cowpeagenomics.med.virginia.edu/CGKB).

Data analysis

Parsimony analysis was made with Paup* 4.0a169 [21]. The two most variable markers, i.e. SSR 6193 and 6220 were removed from the data set for this parsimony analysis.

Chromosomal Location Of The Ssr Markers And Map Construction

Each polymorphic SSR marker used in this study was blasted against the cowpea genome available in Phytozome (https://phytozome-next.jgi.doe.gov/). The markers were mapped on Munoz-Amatriain et al. [22] chromosomes based on their physical position using MapChart 2.3. [23].

Results

SSR Polymorphism

Out of the 61 SSR primers tested, 27 yielded amplified products across all cowpea subspecies. Some primers like SSR 6326 amplified subsp. *unguiculata* and accessions from close subspecies but not the accessions of subspecies far from subsp. *unguiculata*, which suggests mutations in the anchoring region. They were not included in the analysis.

Vigna vexillata was initially included as an outgroup but the primers did not amplify the DNA for half of the accessions. For the other half, the *V. vexillata* allele was different from all the *V. unguiculata* allele. The only exception was SSR 6209 which yielded an allele for NI 1014 that is similar to the allele of subsp. *baoulensis*. Therefore, NI 1014 was not included in the parsimony analysis and the tree was not rooted.

Finally, 18 SSR markers were polymorphic (average 3.83 alleles per locus). With the exception of the very variable SSR 6193 (8 alleles) and SSR 6620 (12 alleles), the number of alleles varied from 2 to 5 alleles for the polymorphic loci (Table 2). The 18 polymorphic SSR are distributed over 10 chromosomes (Fig. 1). Some markers are located in close vicinity (SSR 6193 and 6222, SSR 6225 and 6246, SSR 6274 and 6674) but, within these marker pairs, both markers behave very differently.

Regarding SSR that could be used for molecular characterization, i.e. that show no variability within a subspecies or a group, 11 SSR are characterizing 6 subspecies or varieties (Table 2). A combination of SSR 6246, 6274, and 6920 almost characterizes subsp. *stenophylla* (and unfortunately SP 304). A combination of SSR 6209, 6212, 6274, and 6920 characterizes var. *protracta*. Var. *protracta* is the taxonomic group the most difficult to characterize.
A unique combination of three alleles from SSR 6246, 6274, and 7067 characterizes most accessions from subsp. *alba*, subsp. *tenuis*, subsp. *pubescens*, and the BWA group of var. *spontanea*, as well as

accession SP 141 from the IOCP group of var. *spontanea*.

Parsimony Analysis

The parsimony analysis (Fig. 2) yielded numerous trees with a length of 52. They differed in the position of MT 340 (with subsp. *pawekiae* or with var. *kgalagadiensis*), SP 167 and SP 304, and SP 219 and SP 582 (with subsp. *unguiculata*, with the subsp. *alba* - subsp. *pubescens* polytomy, or in a fourth clade). The tree here presented has a consistency index of 0.6346 and a homoplasy index of 0.3654.

Although this tree is not rooted, we can consider a basal polytomy with 3 clades. The first clade includes subsp. *baoulensis*, subsp. *letouzeyi*, subsp. *pawekiae*, subsp. *stenophylla*, var. *kgalagadiensis*, var. *protracta*, i.e. the main subspecies [2]. The second clade includes subsp. *pubescens*, subsp. *alba*, the BWA group, and subsp. *tenuis*, i.e. the subspecies of hybrid origin [2]. The third clade fits subsp. *unguiculata*, including two accessions from the IOCP group.

Discussion

The SSR tested are spread all over the genome. They are not concentrated in few chromosomes and are representative of the whole genome. The SSR tested can characterize all the main subspecies [2], i.e. subsp. *pawekiae*, subsp. *letouzeyi*, subsp. *baoulensis*, var. *protracta*, var. *kgalagadiensis*, and subsp. *stenophylla*, as well as the annual subsp. *unguiculata*, but they failed to characterize most of the subspecies and groups of hybrid origin [2], i.e. subsp. *alba*, subsp. *tenuis* as well as the the BWA group and the IOCP group of var. *spontanea*. There is no set of SSR yet for characterizing subsp. *tenuis* nor subsp. *alba*.

After Pasquet et al. [2] parsimony analysis of cowpea chloroplast, this is the first cowpea gene pool parsimony analysis based on nuclear DNA. The chloroplast DNA led to a seven clades polytomy while we have here a three clades polytomy. And if subsp. *unguiculata* forms a single clade in both analyses, there are major differences between both analyses.

Chloroplast DNA clades A, B, D, E, and the accessions not belonging to any clade are here pooled in the main clade, with the exception of subsp. *alba* accessions here included in the hybrid origin clade. Regarding the organization of the cowpea gene pool, this work confirms the opposition between the main subspecies and the subspecies of hybrid origin. With the exception of the paraphyletic subsp. *stenophylla* and var. *protracta*, all the main subspecies as well as the annual subsp. *unguiculata* are monophyletic. According to this nuclear phylogeny, var. *kgalagadiensis* could deserve a subspecies status.

Interestingly the split between the forest subspecies from the Mensensis group and the savannah subspecies from the Dekindtiana group does not appear in this analysis. The forest subspecies do not make a monophyletic group and neither the savannah subspecies. Instead of the forest versus savannah opposition, it seems that we have an opposition between the main subspecies with a keel twisted toward

left (with the exception of subsp. *letouzeyi*) and the subspecies which are showing a keel twisted toward right, i.e. subsp. *unguiculata* and the subspecies with an hybrid origin.

Regarding the subspecies of hybrid origin, they appear obviously between the main subspecies and subsp. *unguiculata*, along with the BWA and IOCP groups, although they have almost proper alleles. These almost proper alleles (from SSR 6246, 6274 and 7067) are almost grouping all these accessions in the middle cluster while such a grouping was not appearing in Pasquet [18] nor in Ogunkanmi et al. [12]. Although grouped by these SSR markers, these accessions belong to three different chloroplast clades [2]. Chloroplast clades C and F are fitting the present hybrid origin clade. Subsp. *alba* having a var. *kgalagadiensis* chloroplast but standing here far from var. *kgalagadiensis* seems to be a clear example of old chloroplast captures. This confirms the hybrid origin of subsp. *alba* and suggests that the male ancestor capturing the var. *kgalagadiensis* chloroplast could be subsp. *tenuis* (or a taxon close to subsp. *tenuis*) instead of subsp *unguiculata*.

As observed with cpDNA [2] few accessions from the subspecies of hybrid origin are not in their expected clade. Subsp. *tenuis* MT 340 is associated with var. *kgalagadiensis*. It has 3 alleles in common with var. *kgalagadiensis* and 3 alleles in common with the other subsp. *tenuis* accessions. Subsp. *tenuis* SP 304 is also misplaced due to its allele at SSR 6246 mainly encountered in var. *protracta* (SP 304 was collected in Port Saint Johns in South Africa, few km away from a var. *protracta* area). Similarly, SP 141 stands with subsp. *alba*, subsp. *tenuis* and subsp. *pubescens* due to its allele at SSR 6246. These accessions are from a geographic area where different subspecies do overlap and where numerous intermediate plants are encountered. These discrepancies are likely due to recent hybridizations or to Incomplete Lineage Sorting [2].

This work also confirms the special status of the annual subsp. *unguiculata*. In all the analyses subsp. *unguiculata* stands at the opposite of the different perennial subspecies. This can be explained by its annual status. More generations should likely produce more mutations, as observed previously with cpDNA [2]. This should contribute to its isolation in the different analyses.

Var. *protracta*, standing at the bottom of the clade including all the main subspecies in the parsimony analysis, and not as well grouped as the other main subspecies, appears as the oldest subspecies. Since the parsimony analysis tree is not rooted, we could also consider var. *protracta* as a pivot between the main subspecies and the group made of subsp. *unguiculata* and the subspecies of hybrid origin. This should be in agreement with the hypothesis of the species *Vigna unguiculata* originating in southern Africa [24].

Conclusion

Unfortunately, subsp. *dekindtiana sensu stricto* from southern Angola is still unavailable and the outgroup accession was too far, which hampers the reconstitution of the complex evolutionnary history of *V. unguiculata*. But this work can be considered as the first parsimony analysis attempt of the *V. unguiculata* nuclear genome.

Of course, a larger set of primers would need to be tested on a larger set of accessions but the SSR tested allowed to characterize subsp. *pubescens* and all the main subspecies [2]. We are still far from an ideal two-loci system, although multiplex-SSR analysis [e.g. 25] can partly overcome this problem.

However, we can already conclude that SSR markers associated with SNP derived from chloroplast restriction site mutations should be the perfect tool for cowpea subspecies molecular characterization. Such a tool should help understanding the complex evolutionary history of the cowpea gene pool as well as improving its taxonomy. Maybe more important, it should help breeders for accessing the greatest part of the cowpea gene pool diversity.

Declarations

Ethical Approval not applicable

Competing interests The authors declare no competing interests

Consent to Participate All authors consent to participate.

Consent for Publication All authors consent to the publication of this research.

Authors' Contributions

RSP, NC, and AB conceived and designed the work, RD, DF, and AM helped AJCQ in her lab work, AJCQ, RSP and DD analyzed the results and wrote the first draft, all authors contributed to the final manuscript.

Funding

Allonoumi J.C. Quenum was supported by Deutscher Akademischer Austauschdienst In Region Scholarship Programme -CERAAS Senegal 2017, grant 91689724

Availability of data and materials not applicable

Statements and declarations

Competing interests The authors declare no competing interests

Acknowlegements

We thank the German Academic Exchange Service (DAAD) for its in country/in region Fellowship support to the first author. We thank the Botanical Garden of Brussels for the supply of their accessions. We thank the Botanical Garden of Brussels for the supply of living accessions.

References

1. Boukar O, Fatokun CA, Huynh BL et al (2016) Genomic Tools in Cowpea Breeding Programs: Status and Perspectives. Front Plant Sci 7:757. https://doi.org/10.3389/fpls.2016.00757
2. Pasquet RS, Feleke Y, Gepts P (2021) Cowpea [*Vigna unguiculata* (L.) Walp.] maternal lineages, chloroplast captures, and wild cowpea evolution. Genet Resour Crop Evol 68:2799–2812. https://doi.org/10.1007/s10722-021-01155-y
3. Antil S, Abraham JS, Sripoorna S et al (2022) DNA barcoding, an effective tool for species identification: a review. Mol Biol Rep. https://doi.org/10.1007/s11033-022-08015-7
4. Feleke Y (2006) Phylogenetic relationships in genus *Vigna* and cowpea revealed by DNA sequences from the chloroplast and nuclear genomes. Ph.D. dissertation, Kenyatta University, Kenya
5. Rieseberg LH, Wood TE, Baack EJ (2006) The nature of plant species. Nature 440:524–527. https://doi.org/10.1038/nature04402
6. Fazekas AJ, Kesanakurti PR, Burgess KS et al (2009) Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? Mol. Ecol Resour 9:130–139. https://doi.org/10.1111/j.1755-0998.2009.02652.x
7. Feleke Y, Pasquet RS, Gepts P (2006) Development of PCR-based chloroplast DNA markers to assess gene flow between wild and domesticated cowpea (*Vigna unguiculata*). Plant Syst Evol 262:75–87. https://doi.org/10.1007/s00606-006-0475-0
8. Sarr A, Bodian A, Gbedevi KM et al (2021) Genetic Diversity and Population Structure Analyses of Wild Relatives and Cultivated Cowpea (*Vigna unguiculata* (L.) Walp.) from Senegal Using Simple Sequence Repeat Markers. Plant Mol Biol Rep 39:112–124. https://doi.org/10.1007/s11105-020-01232-z
9. Huynh BL, Close TJ, Roberts PA et al (2013) Gene pools and the genetic architecture of domesticated cowpea. Plant Genome 6:1. https://doi.org/10.3835/plantgenome2013.03.0005
10. Fatokun C, Girma G, Abberton M et al (2018) Genetic diversity and population structure of a mini-core subset from the world cowpea (*Vigna unguiculata* (L.) Walp.) germplasm collection. Sci Rep 8:16035. https://doi.org/10.1038/s41598-018-34555-9
11. Ketema S, Tesfaye B, Keneni G et al (2020) DArTSeq SNP-based markers revealed high genetic diversity and structured population in Ethiopian cowpea [*Vigna unguiculata* (L.) Walp] germplasms. PLoS ONE 15:e0239122. https://doi.org/10.1371/journal.pone.0239122
12. Ogunkanmi LA, Ogundipe OT, Ng NQ et al (2008) Genetic diversity in wild relatives of cowpea (*Vigna unguiculata*) as revealed by simple sequence repeats (SSR) markers. J Food Agric Environ 6:263–268. https://hdl.handle.net/10568/90821
13. Tuler AC, Carrijo TT, Noia LR et al (2015) SSR markers: a tool for species identification in *Psidium* (Myrtaceae). Mol Biol Rep 42:1501–1513. https://doi.org/10.1007/s11033-015-3927-1
14. Wang J, Luo J, Ma YZ et al (2019) Nuclear simple sequence repeat markers are superior to DNA barcodes for identification of closely related *Rhododendron* species on the same mountain. J Syst Evol 57:278–286. https://doi.org/10.1111/jse.12460

15. Terrones A, van der Bank M, Moreno J et al (2022) DNA barcodes and microsatellites: How they complement for species identification in the complex genus *Tamarix* (Tamaricaceae). J Syst Evol 60:1140–1157. https://doi.org/10.1111/jse.12830
16. Timko MP, Rushton PJ, Laudeman TW et al (2008) Sequencing and analysis of the gene-rich space of cowpea. BMC Genomics 9:103. https://doi.org/10.1186/1471-2164-9-103
17. Andargie M, Pasquet RS, Gowda BS et al (2014) Molecular mapping of QTLs for domestication-related traits in cowpea (*V. unguiculata* (L.) Walp). Euphytica 200:401–412. https://doi.org/10.1007/s10681-014-1170-9
18. Pasquet RS (1999) Genetic relationships among subspecies of *Vigna unguiculata* (L.) Walp. based on allozyme variation. Theor Appl Genet 98:1104–1119. https://doi.org10.1007/s001220051174
19. Li CD, Fatokun CA, Ubi B et al (2001) Determining genetic similarities and relationships among cowpea breeding lines and cultivars by microsatellite markers. Crop Sci 41:189–197. https://doi.org/10.2135/cropsci2001.411189x
20. Diouf D, Hilu KW (2005) Microsatellites and RAPD markers to study genetic relationships among cowpea breeding lines and local varieties in Senegal. Genet Resour Crop Evol 52:1057–1067. https://doi.org/10.1007/s10722-004-6107-z
21. Swofford D (2017) PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0. Sinauer, Sunderland, Massachusetts, USA
22. Munoz-Amatriain M, Mirebrahim H, Xu P et al (2017) Genome resources for climate-resilient cowpea, an essential crop for food security. Plant J 89:1042–1054. https://doi.org/10.1111/tpj.13404
23. Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78. https://doi.org/10.1093/jhered/93.1.77
24. Padulosi S (1993) Genetic diversity, taxonomy and ecogeographic survey of the wild relatives of cowpea (*Vigna unguiculata* (L.) Walpers). PhD dissertation, Universite catholique Louvain la Neuve, Belgium
25. Mitchell SE, Kresovich S, Jester CA, Hernandez CJ (1997) Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop Sci 37:617–624. https://doi.org/10.2135/cropsci1997.0011183X003700020051x

Tables

Tables 1-2 is available in the Supplementary Files section.

Figures

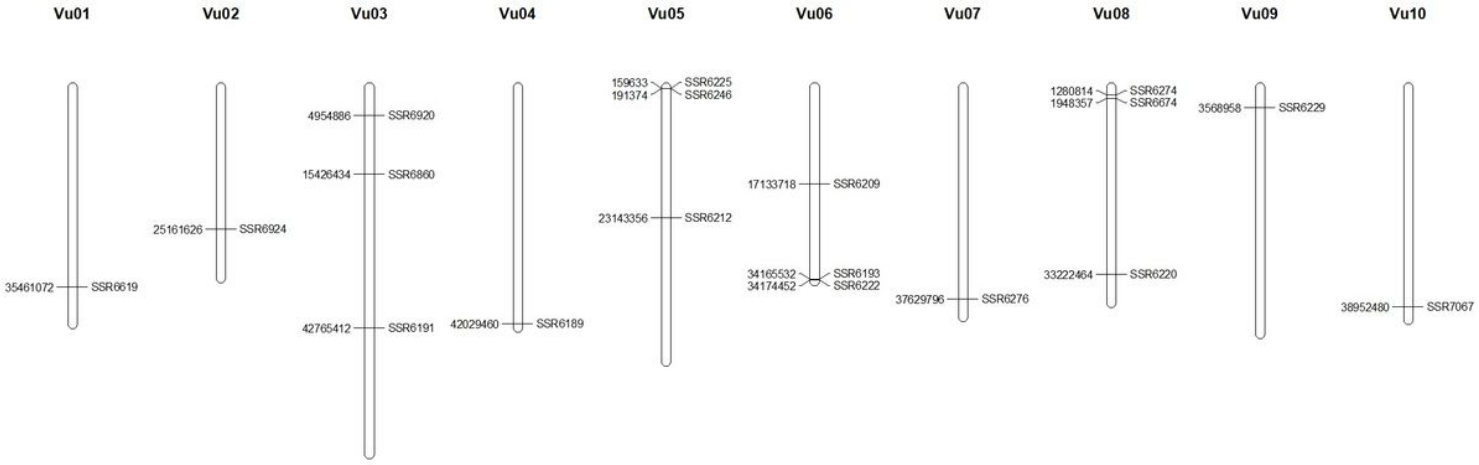

Figure 1

Distribution of the 18 polymorphic SSR on 10 of the cowpea chromosomes.

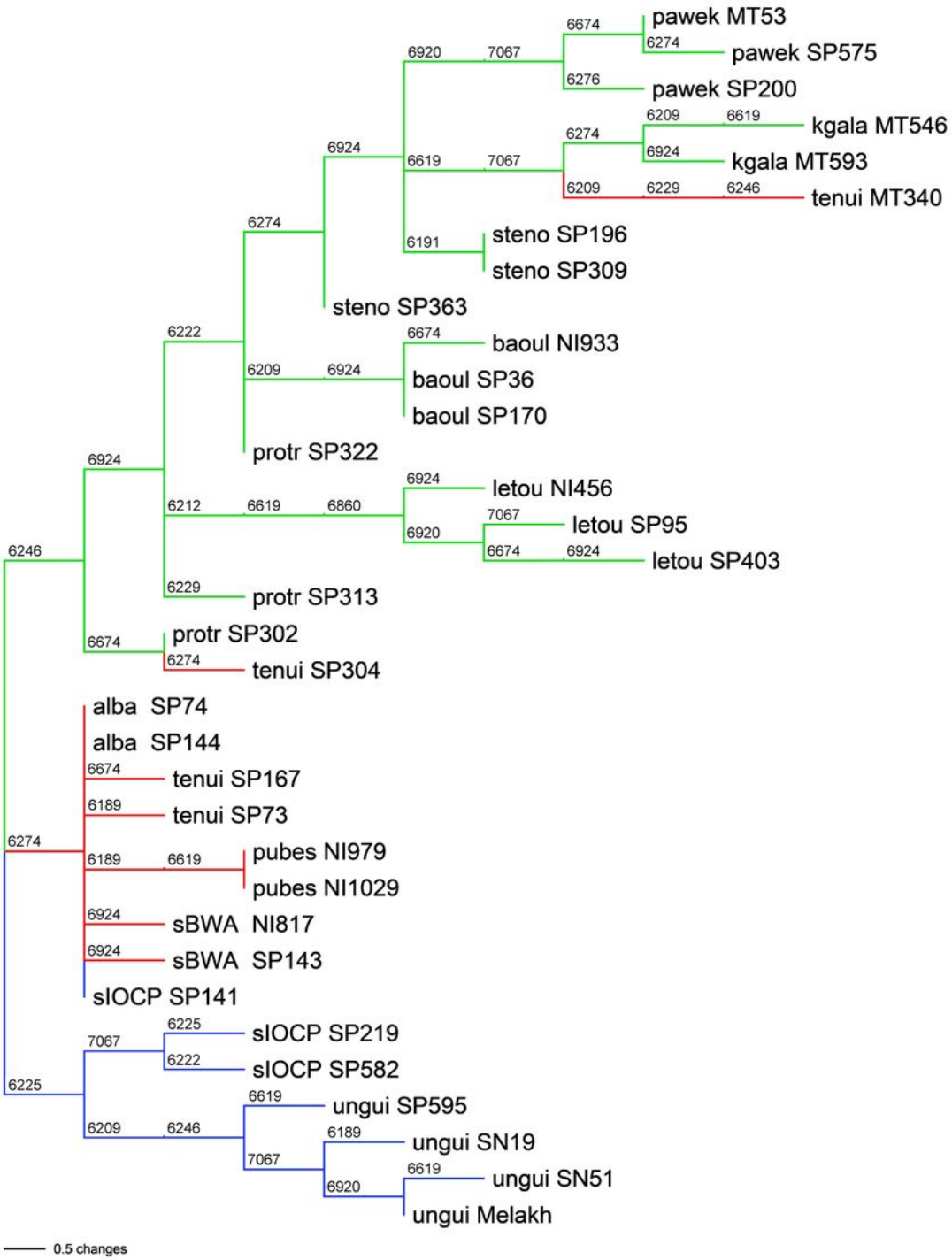

Figure 2

Parsimony analysis. SSR 6193 and SSR 6220 were not included in this parsimony analysis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Tables.docx