

Comparative Analysis of Codon Usage Bias in Chloroplast Genomes of eight Argentina Species of Rosaceae and Their Phylogeny

Guang-Nan Zhang

Academy of Agriculture and Forestry Sciences of Qinghai University

Li-Zhen Ling

Liupanshui Normal University

Jie Tian

Academy of Agriculture and Forestry Sciences of Qinghai University

Shu-Dong Zhang

`sdchang@foxmail.com`

Liupanshui Normal University

Research Article

Keywords: Chloroplast Genome, Codon Usage Bias, Argentina, Phylogeny

Posted Date: September 19th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7501003/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

The genus *Argentina* (Rosaceae) represents a taxonomically complex group with significant ethnobotanical value, yet the morphological convergence among genus remains poorly resolved. Codon usage bias (CUB), an important genomic feature has not been systematically investigated in this genus, limiting our understanding of its chloroplast evolution and adaptive mechanisms.

Results

In this study, we sequenced and annotated the complete chloroplast genomes of three key *Argentina* species (*A. stenophylla*, *A. anserina*, and *A. phanerophlebia*) and reconstructed a phylogeny using 21 taxa (18 *Argentina* and three *Potentilla* outgroups). We further conducted a comparative analysis of synonymous codon usage bias in eight representative species. The chloroplast genomes exhibited a typical quadripartite structure with high conservation in size (155,148–155,972 bp), gene order, and content (129 genes). Phylogenetic analysis strongly supported the monophyly of *Argentina* and identified four major clades. Codon usage analysis revealed a weak overall bias and a consistent preference for A/T-ending codons. Neutrality, ENC-, and PR2-plots indicated that natural selection plays a dominant role in shaping CUB, overriding mutational pressure. We identified 16 optimal codons, only one of which (GGA, encoding Gly) was shared across all species, suggesting lineage-specific codon adaptation.

Conclusions

This study provides the first comprehensive insight into codon usage patterns in *Argentina* chloroplast genomes, enhancing our understanding of phylogenetic relationships and molecular evolution in this genus, with implications for future genetic research and codon optimization strategies.

Background

The genus *Argentina* Hill (Rosaceae: Rosoideae, Potentilleae), which comprises approximately 75 species primarily distributed across the Sino-Himalayan region and Malay Archipelago, represents a taxonomically complex group with significant ethnobotanical value [1–3]. Species such as *A. anserina* (L.) Rydb. produce starch-rich rhizomes traditionally used for nutritional supplementation and anemia treatment, while *A. lineata* (Trevis.) Soják and *A. leuconota* (D.Don) Soják are known for their hemostatic and detoxifying properties, respectively (Flora Reipublicae Popularis Sinicae, <http://frps.eflora.cn/>) [1]. Despite these functional attributes, molecular studies aimed at resolving the phylogeny of the genus have been hindered by taxonomic ambiguities resulting from morphological convergence with allied genera such as *Potentilla* and *Sibbaldia*. Initial generic delineation by Rydberg [4], based on stylar

insertion, was later refined by Soják [2] using stipule morphology. However, molecular analyses employing chloroplast markers (*matK*, *trnL-F*) and nuclear loci (ITS) have produced conflicting topologies regarding the monophyly of the genus and the transfer of *Sibbaldia micropetala* to *A. micropetala* [5, 6]. These inconsistencies highlight the need for phylogenomic approaches utilizing whole chloroplast genomes to clarify evolutionary relationships within Potentilleae [7].

Chloroplast genomes, characterized by their conserved quadripartite structure—consisting of Large and Small Single-Copy regions (LSC, SSC) flanked by Inverted Repeats (IRs)—and moderate nucleotide substitution rates, provide a robust foundation for phylogenetic reconstruction [8, 9]. Beyond structural conservation, synonymous codon usage bias (SCUB) in chloroplast protein-coding genes serves as a key indicator of molecular evolution, reflecting the interplay of mutational pressure, natural selection, and genetic drift [9]. Key metrics such as Relative Synonymous Codon Usage (RSCU), which quantifies amino acid-specific codon preferences [10], the Effective Number of Codons (ENC), measuring deviation from random codon usage [11], and neutrality plots (GC_{12} vs. GC_3), which discriminate between mutation-selection equilibrium [12], reveal lineage-specific adaptations in translational efficiency and nucleotide composition. Although SCUB analysis has been applied in some Rosaceae species [13, 14], systematic studies within Potentilleae, particularly in *Argentina*, are lacking. This presents an opportunity to elucidate codon optimization mechanisms driving chloroplast evolution in this understudied clade.

To address these gaps, this study integrates chloroplast genomics with comparative codon usage analysis. We sequenced and annotated the complete chloroplast genomes of three taxonomically pivotal *Argentina* species—*A. stenophylla*, *A. anserina*, and *A. phanerophlebia*—and reconstructed a phylogeny using 21 taxa, including 18 *Argentina* species and three *Potentilla* outgroups. Based on this phylogenetic framework, we conducted a comparative SCUB analysis of representative species from the major clades. This work provides new insights into the evolutionary mechanisms shaping organellar genome function in *Argentina*.

Results

Characteristics of the Chloroplast Genomes of Three Argentina Species

The complete chloroplast genomes of three *Argentina* species were assembled and annotated. Their sizes were similar, ranging from 155,148 bp (*A. anserina*) to 155,972 bp (*A. stenophylla*) (Additional file 1: Table S1). Each genome exhibited the typical quadripartite structure of chloroplast DNA, consisting of a large single-copy region (LSC: 85,068–85,619 bp), a small single-copy region (SSC: 18,461–18,766 bp), and a pair of inverted repeat regions (IRs: 25,702–25,992 bp) (Additional file 1: Table S1; Fig. 1). The overall GC content ranged from 36.74–37.15%. A total of 129 genes were identified in each plastome, including 84 protein-coding genes (PCGs), 37 tRNA genes, and 8 rRNA genes (Additional file 2: Table S2). These genes were categorized into four functional groups: self-replication, photosynthesis, other functions, and genes of unknown function. The protein-coding genes are primarily involved in photosynthesis, photosynthetic phosphorylation, and genetic expression (Fig. 1; Additional file 2: Table

S2). The order and arrangement of genes were highly conserved across the three species, indicating structural stability of the chloroplast genomes in *Argentina*.

Phylogenetic analyses

We reconstructed the phylogeny of *Argentina* using complete chloroplast genome sequences from 18 species, with three *Potentilla* species as outgroups (Fig. 2). The results strongly supported the monophyly of the genus *Argentina* (PP = 1, BS = 100). Within the genus, four distinct clades were identified. *A. anserina* formed a basal sister taxon to *A. smithiana* (PP = 1, BS = 100), and these two species were most closely related to *A. micropetala*. *A. phanerophlebia* constituted a separate branch. The remaining species were divided into two major clades: six *Argentina* species formed one well-supported branch, while *A. stenophylla* clustered closely with *A. microphylla*, *A. stenophylla* var. *emergens*, *A. taliensis*, *A. festiva*, *A. lineata*, *A. polyphylla*, and *A. parvula*. For subsequent codon usage analysis, two representative species from each clade were selected, resulting in a total of eight species: *A. anserina*, *A. smithiana*, *A. phanerophlebia*, *A. micropetala*, *A. cardotiana*, *A. leuconota*, *A. stenophylla*, and *A. parvula*.

Analysis of Codon Usage Bias Parameters

We analyzed 53 protein-coding sequences from the chloroplast genomes of the eight *Argentina* species (Additional file 3: Table S3; Fig. 2). The overall GC content (GCall) was similar among species, ranging between 37.80% and 38.00%. The GC content at each codon position was below 47%, with the trend GC1 (45.98–46.15%) > GC2 (37.78–37.92%) > GC3 (29.43–29.92%). Nucleotide composition analysis at the third codon position revealed a marked bias: T3 (46.93–47.25%) > A3 (42.86–43.17%) > G3 (18.04–18.34%) > C3 (16.66–16.97%), indicating a pronounced preference for A and T endings. The Effective Number of Codons (ENC) values ranged from 35 to 60 across all species (Additional file 3: Table S3; Fig. 2), suggesting weak codon usage bias.

Correlation analysis revealed that GCall was significantly correlated with GC1, GC2, and GC3 ($P < 0.01$; correlation coefficients: 0.46–0.82). GC1 and GC2 were also significantly correlated ($P < 0.05$), but no significant correlation was observed between GC1 and GC3 (Fig. 3), indicating that mutational bias was most pronounced at the third codon position. ENC showed a significant positive correlation with GC3 in all species ($P < 0.01$) and a negative correlation with GC2 in *A. cardotiana*. No significant correlation was detected between ENC and GC1 or GCall (Fig. 3). The Codon Adaptation Index (CAI) values all exceeded 0.40, further supporting low codon bias and moderate gene expression levels. CAI was significantly correlated with GC1 in all species, but not with GCall in *A. micropetala*, *A. parvula*, or *A. phanerophlebia* (Fig. 3).

Optimal Codon Analysis

A total of 29 codons with RSCU > 1 were identified across the eight *Argentina* species, indicating high-frequency usage (Additional file 4: Table S4). Among these, UUA (Leu) exhibited the highest RSCU value, followed by GCU (Ala). Notably, 55.2% of these preferred codons ended in U, 41.4% in A, and only one (UUG, Leu) ended in G, further supporting a pronounced bias toward A/U-ending codons in the chloroplast genomes of *Argentina*. This pattern suggests a shared codon usage strategy among species within the genus.

Based on established criteria (RSCU > 1 and $\Delta\text{RSCU} \geq 0.08$), 16 codons were identified as optimal (Fig. 4). However, the number of optimal codons varied considerably among species: *A. anserina* and *A. cardotiana* each contained 12; *A. smithiana* and *A. stenophylla* possessed 6 and 7; *A. leuconota* and *A. parvula* had 5; *A. micropetala*; while *A. phanerophlebia* contained 4 and 2, respectively. Remarkably, only one codon, GGA (Gly), was common to all eight species, underscoring substantial interspecific divergence in codon use preference.

Clustering analysis based on RSCU values revealed relationships distinct from those in the phylogenetic tree. *A. stenophylla* formed the earliest diverging branch, followed by *A. anserina* and *A. smithiana*. *A. cardotiana* and *A. leuconota* clustered together as a sister clade to another group comprising *A. micropetala* and *A. parvula* (Fig. 4). This discordance suggested that codon usage bias in *Argentina* may be influenced by lineage-specific evolutionary pressures beyond those reflected in phylogenetic structure.

Neutrality Plot Analysis

Neutrality analysis, which examines the relationship between GC12 and GC3, serves to evaluate the relative contributions of mutation pressure and natural selection to codon usage bias. In the genus *Argentina*, GC12 values ranged from 0.32 to 0.54, while GC3 values varied between 0.18 and 0.37 (Fig. 5). The correlation between GC12 and GC3 was weak and non-significant, with correlation coefficients ranging from 0.0289 to 0.0421, indicating distinct evolutionary constraints acting on these nucleotide positions. Regression analysis further revealed that mutation pressure accounted for only 24.76–29.59% of the variation in codon usage, whereas natural selection explained 70.41–75.24%. These results demonstrate that natural selection played a dominant role in shaping codon preference in the chloroplast genomes of *Argentina*.

ENC-Plot Analysis

To further assess the influence of compositional constraints on codon usage, ENC-plot analysis was conducted. The majority of genes across all eight species fell below the expected ENC curve (Fig. 6), indicating that codon usage bias is influenced by both mutation pressure (particularly at the third codon position) and natural selection. The consistent distribution pattern among species reinforced the

conclusion that natural selection exerted a stronger influence than mutational bias, corroborating the findings from the neutrality plot analysis (Fig. 5).

PR2-Plot Analysis

According to the PR2, under pure mutational pressure, the frequencies of (A and T) and (G and C) at the third codon position are expected to be equal. However, deviations from this equilibrium reflect the influence of natural selection. In the eight *Argentina* species, gene points were asymmetrically distributed across the four quadrants, with a notable concentration in the lower right region ($T > A$ and $G > C$) (Fig. 7). This biased distribution indicates a preferential use of T and G at the third codon position and further supports the predominant role of natural selection in shaping codon usage patterns.

Discussion

Chloroplast genomes have been widely employed in plant phylogenetic studies through various approaches, including restriction fragment/site comparisons, structural rearrangements, and sequence variation analyses [8]. Previous studies indicated that chloroplast genomes of *Argentina* species are highly conserved in terms of genome structure, sequence, gene order, and content. Consequently, structural homoplasy limits their utility as phylogenetic markers; however, several highly divergent regions have been identified that offer resolution at deeper nodes [8]. In this study, we newly sequenced and assembled the chloroplast genomes of three *Argentina* species and conducted a comprehensive phylogenetic analysis incorporating 18 species within the genus. Our results robustly support the monophyly of *Argentina* and its division into four major clades, consistent with recent phylogenomic studies based on complete chloroplast genomes [15]. These findings confirm that chloroplast genomic data provide a reliable phylogenetic framework for resolving deep evolutionary relationships within *Argentina*.

Synonymous mutations have traditionally been considered “silent,” as they do not alter the encoded amino acid. However, growing evidence indicates that such mutations can profoundly influence gene expression, protein function, and organismal fitness through mechanisms such as modulation of mRNA stability, translation efficiency, and co-translational folding. For instance, 75.9% of synonymous mutations in yeast significantly reduce fitness [16]. In this study, we systematically analyzed synonymous codon usage in the chloroplast genomes of eight *Argentina* species. We identified a weak codon usage bias and a pronounced preference for A/T-ending codons, particularly at the third position—a pattern commonly observed across land plants [13, 17]. Furthermore, correlation analyses indicated that gene expression levels (inferred via CAI) may be associated with GC content at the first codon position. Optimal codons are known to enhance translational efficiency through mechanisms such as matching tRNA abundance, accelerating elongation rates, and facilitating proper protein folding [18, 19]. A recent study in cucumber demonstrated that a synonymous mutation in the ACS2 gene affects translation efficiency via m6A RNA modification and altered mRNA secondary structure, ultimately influencing fruit length [20]. However, the functional impacts of codon usage may be modulated by

positional effects, tissue specificity, and global cellular constraints [21–24]. Notably, optimal codon identities varied considerably among the *Argentina* species studied, suggesting divergent evolutionary trajectories or species-specific regulatory adaptations.

Our analyses—including neutrality plots, ENC-plots, and PR2 analysis—consistently indicated that natural selection plays a dominant role in shaping codon usage bias in the chloroplast genomes of *Argentina*, outweighing the influence of mutation pressure. This finding aligns with previous reports in *Caragana* [17] and *Lilium* [25]. Species of *Argentina* often inhabit high-altitude environments (above 3,000 m), and their adaptations—such as compact morphology, pubescent leaves, and signatures of positive selection in chloroplast genes [15, 26]—may be reflected in codon usage patterns. It is well established that natural selection can favor specific synonymous codons to optimize translation efficiency and accuracy, often in accordance with genomic GC content [27]. We speculate that selection may favor A/U-ended codons to match the host's tRNA pool, thereby enhancing translational efficiency. Future studies quantifying tRNA abundance and calculating tRNA adaptation indices could provide deeper insights into the mechanisms of translational selection. Ultimately, understanding species-specific codon preferences may inform strategies for transgene optimization and synthetic biology applications in these taxa.

Conclusion

This study presents the first comprehensive analysis of codon usage bias in the chloroplast genomes of *Argentina* species, integrated with a robust phylogenomic framework. Our results confirm the monophyletic status of the genus and reveal four well-supported clades, consistent with recent phylogenetic studies. The chloroplast genomes are highly conserved in structure and gene content, yet exhibit significant synonymous codon usage variation. The overall weak codon bias and predominant A/T-ending preference reflect adaptive strategies likely influenced by natural selection aimed at optimizing translational efficiency. The identification of species-specific optimal codons underscores the diversity in regulatory mechanisms among *Argentina* species. These findings not only elucidate the evolutionary dynamics of chloroplast genomes in *Argentina* but also provide a foundation for future functional studies and genetic engineering applications, such as enhancing transgene expression through codon optimization. Further research into tRNA abundance and adaptation indices will be essential to fully unravel the mechanisms of translational selection in this genus.

Materials and Methods

Plant Materials, DNA Extraction, Sequencing, and Assembly

Fresh leaves from three *Argentina* species (*A. stenophylla*, *A. anserina*, and *A. phanerophlebia*) were collected from various provinces in China and rapidly dried using silica gel. Species identification was confirmed by Prof. Shu-Dong Zhang, and voucher specimens were deposited at the Herbarium of Liupanshui Normal University. Total genomic DNA was extracted using a modified CTAB method. DNA

libraries were constructed and subjected to paired-end sequencing (PE150) on the Illumina NovaSeq platform.

High-quality reads were used for de novo assembly of chloroplast genomes with SPAdes v3.15 [28]. The resulting contigs were visualized and circularized using Bandage [29]. Complete chloroplast genomes were annotated with the Plastid Genome Annotator (PGA) [30], using reference genomes of related Argentina species. Genome maps were generated using OGDRAW [31].

Phylogenetic analysis of Argentina

A total of 21 cp genomes, comprising 18 *Argentina* species and three *Potentilla* species as outgroups (see Table S1), were included in the phylogenetic analysis. Whole chloroplast genome sequences were aligned using MAFFT v6.833 [32]. Phylogenetic trees were reconstructed using both Bayesian Inference (BI) and Maximum Likelihood (ML) methods, following previously established protocols [33].

Codon Usage Analysis

Protein-coding sequences (CDS) longer than 300 bp were extracted from the chloroplast genomes of eight *Argentina* species. Only sequences with standard start (ATG) and stop (TAA, TAG, TGA) codons were retained. Nucleotide composition indices, including GC content at the first, second, and third codon positions (GC1, GC2, GC3) and overall GC content (GCall), were computed using EMBOSS-cusp (<https://www.bioinformatics.nl/cgi-bin/emboss/cusp>). Relative Synonymous Codon Usage (RSCU) and the Effective Number of Codons (ENC) were calculated using CodonW 1.4.2 [11, 34]. Codons with RSCU > 1 were considered preferentially used, and an ENC value ≤ 35 was indicative of significant codon usage bias [35–37].

Sources of Codon Usage Bias

Neutrality plots (GC12 vs. GC3) were constructed to evaluate the influence of mutation pressure versus natural selection [38]. A regression slope near 0 suggests dominant natural selection, whereas a slope near 1 indicates mutation-dominated bias [39].

ENC plots were generated by comparing observed ENC values against expected values under neutral evolution ($ENC_{exp} = 2 + GC3 + 29/[GC3^2 + (1 - GC3)^2]$) [40]. Points distributed near the expected curve imply mutation-driven bias, whereas deviations below the curve suggest selective constraints [41].

Parity Rule 2 (PR2) analysis [42] was applied to examine strand-specific bias using the parameters $A/(A + T)$ and $G/(G + C)$ at the third codon position.

Optimal Codon Identification

Genes were ranked based on their ENC values, with the top and bottom 10% defined as high- and low-expression sets, respectively. The difference in RSCU ($\Delta RSCU = RSCU_{high} - RSCU_{low}$) was computed for each codon. Optimal codons were identified as those with RSCU > 1 and $\Delta RSCU \geq 0.08$ [43].

RSCU Clustering

RSCU values from the eight *Argentina* species were subjected to hierarchical clustering analysis using Euclidean distance.

Abbreviations

CDS coding DNA sequence

cp chloroplast

CUB codon usage bias

ENC the effective number of codons

IR inverted repeat

LSC large single-copy

PCGs protein-coding genes

RSCU relative synonymous codon usage

SSC small single-copy

Declarations

Clinical trial number

Not applicable

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets analyzed during the current study are available from the NCBI database under accession number PX057651 (*A. stenophylla*), OR863700 (*A. phanerophlebia*) and PX057649 (*A. anserina*).

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the Open Subject from Key Laboratory of Qinghai-Tibetan Plateau Biotechnology of Ministry of Education (2023-SYS-04), Science and Technology Project of Liupanshui, grant number 52020-2019-05-05, 52020[2022]PT-20, and 52020-2023-0-2-18.

Author Contributions

Conceptualization, S-D.Z.; methodology, L-Z.L., G-N. Z. and J.T.; formal analysis, G-N. Z., L-Z.L. and S-D.Z.; writing—original draft preparation, G-N. Z, L-Z.L. and S-D.Z.; writing—review and editing, S-D.Z. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

Not applicable

References

1. Feng T, Moore MJ, Sun Y, Meng A, Chu H, Li J, et al. A new species of *Argentina* (Rosaceae, Potentilleae) from Southeast Tibet, with reference to the taxonomic status of the genus. *Plant Syst Evol.* 2015;301(3):911–21.
2. J. S: *Argentina* Hill, a genus distinct from *Potentilla* (Rosaceae). *Thaiszia J Bot.* 2010; 20:91–7.
3. Kechaykin AA, Shmakov AI. A system of subtribe Potentillinae J. Presl (Rosaceae Juss.). *Turczaninowia.* 2016; 1910(4):114–128.
4. Rydberg PA. Rosaceae. *North Am Flora.* 1908;22(3):239–533.
5. Dobes C, Paule J. A comprehensive chloroplast DNA-based phylogeny of the genus *Potentilla* (Rosaceae): implications for its geographic origin, phylogeography and generic circumscription. *Mol Phylogenet Evol.* 2010;56(1):156–75.
6. Eriksson T, Lundberg M, T?Pel M, ?stensson P, Smedmark JEE. *Sibbaldia*: a molecular phylogenetic study of a remarkably polyphyletic genus in Rosaceae. *Plant Syst Evol.* 2015;301(1):171–84.
7. Li QQ, Khasbagan, Zhang ZP, Wen J, Yu Y. Plastid phylogenomics of the tribe potentilleae (Rosaceae). *Mol Phylogenet Evol.* 2024;190:107961.
8. Raubeson LA, Jansen RK. Chloroplast genomes of plants. *Plant diversity and evolution genotypic and phenotypic variation in higher plants* Edited by R J Henry Wallingford UK. CABI Publishing. 2005;42(1):45–68.
9. Green BR. Chloroplast genomes of photosynthetic eukaryotes. *Plant journal: cell Mol biology.* 2011;66(1):34–44.
10. Ikemura T. Codon usage and tRNA content in unicellular and multicellular organisms. *Mol Biol Evol.* 1985;2(1):13–34.

11. Wright F. The 'effective number of codons' used in a gene. *Gene*. 1990;87(1):23–9.
12. Sueoka N. Directional mutation pressure and neutral molecular evolution. *Proc Natl Acad Sci U S A*. 1988;85(8):2653–7.
13. Ling L, Zhang S, Yang T. Analysis of Codon Usage Bias in Chloroplast Genomes of *Dryas octopetala* var. *asiatica* (Rosaceae). *Genes*. 2024; 15(7).
14. Li G, Zhang L, Xue P, Zhu M. Comparative Analysis on the Codon Usage Pattern of the Chloroplast Genomes in *Malus* Species. *Biochem Genet*. 2023;61(3):1050–64.
15. Li QQ, Zhang ZP, Aogan, Wen J. Comparative chloroplast genomes of Argentina species: genome evolution and phylogenomic implications. *Front Plant Sci*. 2024;15:1349358.
16. Shen X, Song S, Li C, Zhang J. Synonymous mutations in representative yeast genes are mostly strongly non-neutral. *Nature*. 2022;606(7915):725–31.
17. Li X, Liu L, Ren Q, Zhang T, Hu N, Sun J, et al. Analysis of synonymous codon usage bias in the chloroplast genome of five *Caragana*. *BMC Plant Biol*. 2025;25(1):322.
18. Nedialkova DD, Leidel SA. Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. *Cell*. 2015;161(7):1606–18.
19. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Sci (New York NY)*. 2009;324(5924):218–23.
20. Xin T, Zhang Z, Zhang Y, Li X, Wang S, Wang G et al. Recessive epistasis of a synonymous mutation confers cucumber domestication through epitranscriptomic regulation. *Cell*. 2025.
21. Qian W, Yang JR, Pearson NM, Maclean C, Zhang J. Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genet*. 2012;8(3):e1002603.
22. Wu X, Xu M, Yang JR, Lu J. Genome-wide impact of codon usage bias on translation optimization in *Drosophila melanogaster*. *Nat Commun*. 2024;15(1):8329.
23. Krisko A, Copic T, Gabaldón T, Lehner B, Supek F. Inferring gene function from evolutionary change in signatures of translation efficiency. *Genome Biol*. 2014;15(3):R44.
24. Yu CH, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS, et al. Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Mol Cell*. 2015;59(5):744–54.
25. Guona D, Mingyue S, Jiale W, Jiamei Z, Binbin L, Yinglin L, et al. Codon bias and phylogenesis analysis of chloroplast genome in medicinal plants of *Lilium*. *Chin Traditional Herb Drugs*. 2024;55(11):3835–44.
26. Krner C, Paulsen J, Spehn EM. A definition of mountains and their bioclimatic belts for global comparisons of biodiversity data. *Alp Bot*. 2011;121(2):73–8.
27. Hershberg R, Petrov DA. General rules for optimal codon choice. *PLoS Genet*. 2009;5(7):e1000556.
28. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes de novo assembler. *Curr Protocols Bioinf*. 2020;70(1):e102.

29. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics*. 2015;31(20):3350–2.
30. Qu XJ, Moore MJ, Li DZ, Yi TS. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods*. 2019;15(1):1–12.
31. Stephan G, Pascal L, Ralph B. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nuclc Acids Res*. 2019(W1):W59–64.
32. Katoh K, Kuma K, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res*. 2005;33(2):511–8.
33. Zhang SD, Yan K, Ling LZ. Characterization and phylogenetic analyses of ten complete plastomes of *Spiraea* species. *BMC Genomics*. 2023;24(1):137.
34. Fuglsang A. Accounting for background nucleotide composition when measuring codon usage bias: brilliant idea, difficult in practice. *Mol Biol Evol*. 2006;23(7):1345–7.
35. Liu H, He R, Zhang H, Huang Y, Tian M, Zhang J. Analysis of synonymous codon usage in *Zea mays*. *Mol Biol Rep*. 2010;37(2):677–84.
36. Duret L. tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. *Trends Genet*. 2000;16(7):287–9.
37. Liao W. Deciphering Codon Usage Patterns in Genome of *Cucumis sativus* in Comparison with Nine Species of Cucurbitaceae. *Agronomy*. 2021; 11(11).
38. Kawabe A, Miyashita NT. Patterns of codon usage bias in three dicot and four monocot plant species. *Genes Genet Syst*. 2003;78(5):343–52.
39. Li L, Peng J, Wang D, Duan A. Chloroplast genome phylogeny and codon preference of *Docynia longiunguis*. *Chin J Biotechnol*. 2022;38(1):328–42.
40. Parvathy ST, Udayasuriyan V, Bhadana V. Codon usage bias. *Mol Biol Rep*. 2022;49(1):539–65.
41. Fuglsang A. Impact of bias discrepancy and amino acid usage on estimates of the effective number of codons used in a gene, and a test for selection on codon usage. *Gene*. 2008;410(1):82–8.
42. Sueoka N. Translation-coupled violation of Parity Rule 2 in human genes is not the cause of heterogeneity of the DNA G + C content of third codon position. *Gene*. 1999;238(1):53–8.
43. Xiang H, Zhang R, Butler RR 3rd, Liu T, Zhang L, Pombert JF, et al. Comparative Analysis of Codon Usage Bias Patterns in Microsporidian Genomes. *PLoS ONE*. 2015;10(6):e0129223.

Figures

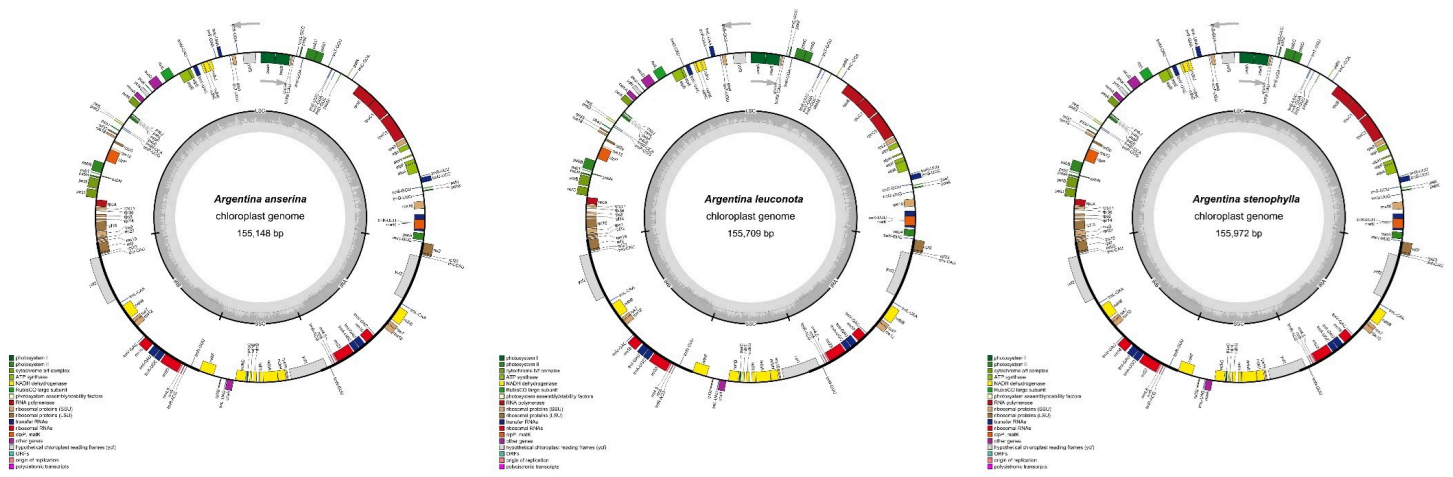


Figure 1

Chloroplast genome map of threet *Argentina* species. The grey inside circle indicates the GC level of every genomic position.

Genes inside in the outer circle of genomic map are transcribed clockwise and vice versa. The different functional gene categories are shown in the different colors

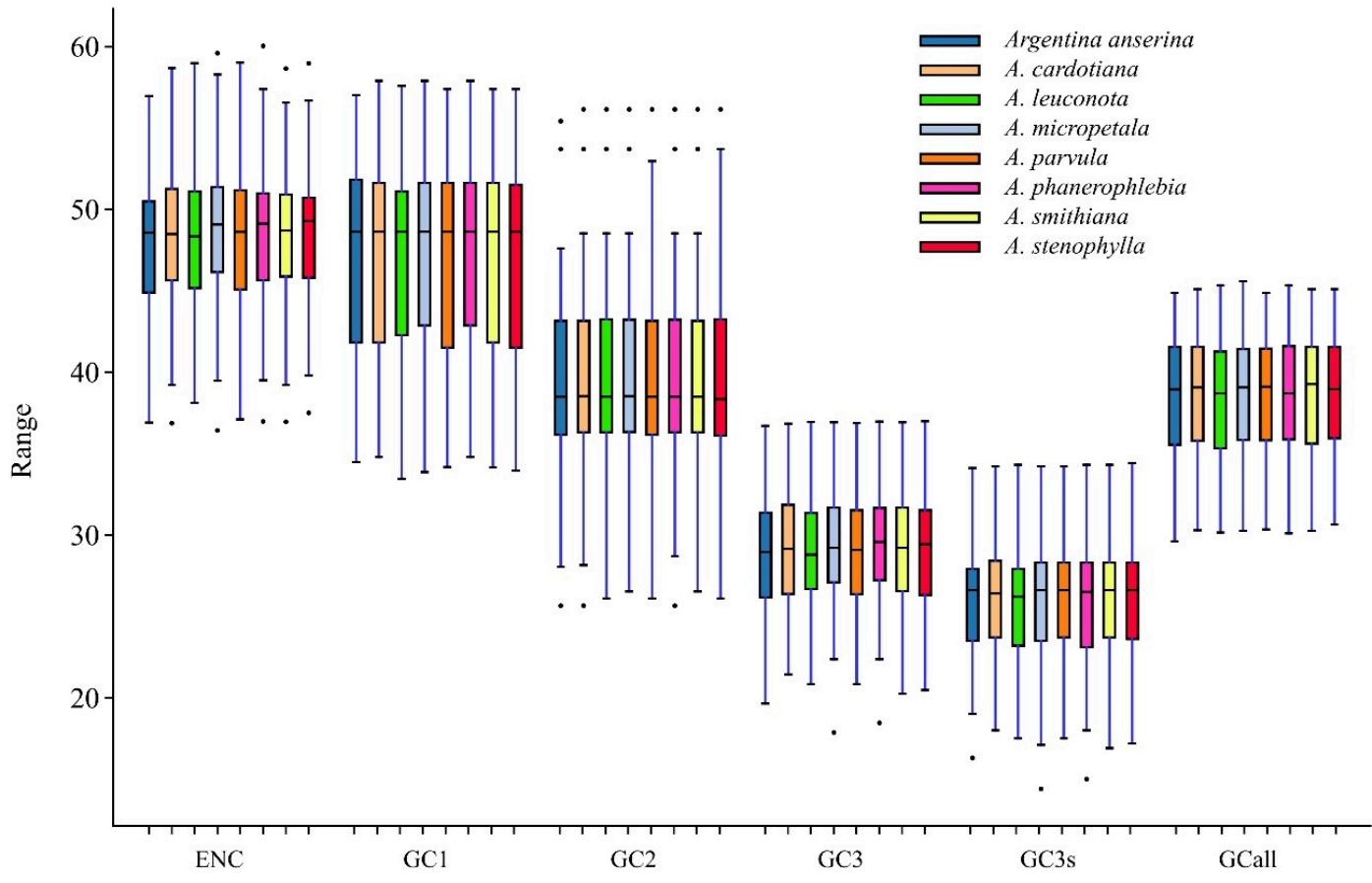


Figure 2

Codon usage bias related parameters of eight *Argentina* species

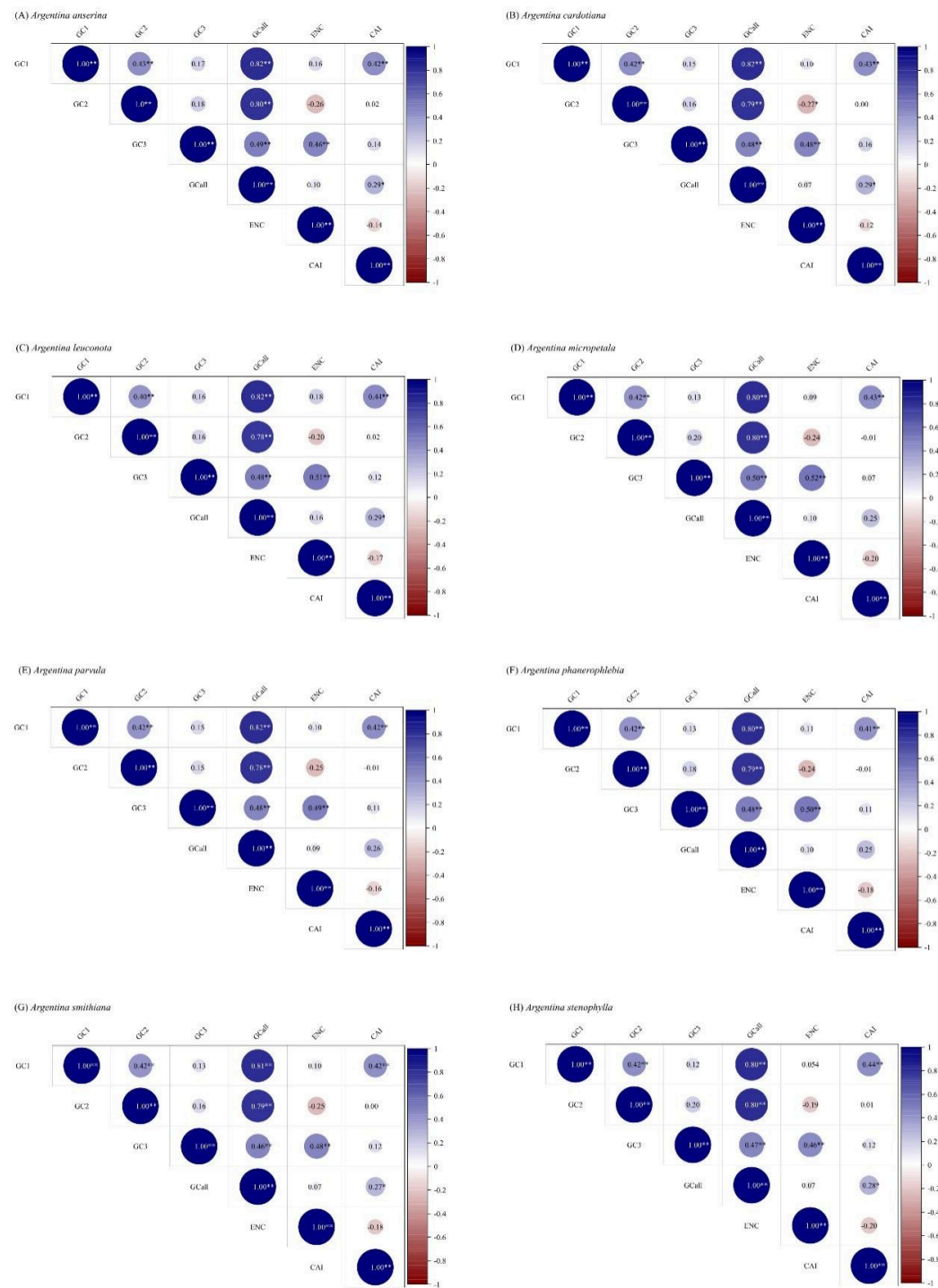


Figure 3

Correlation analysis between parameters of each gene of eight *Argentina* species

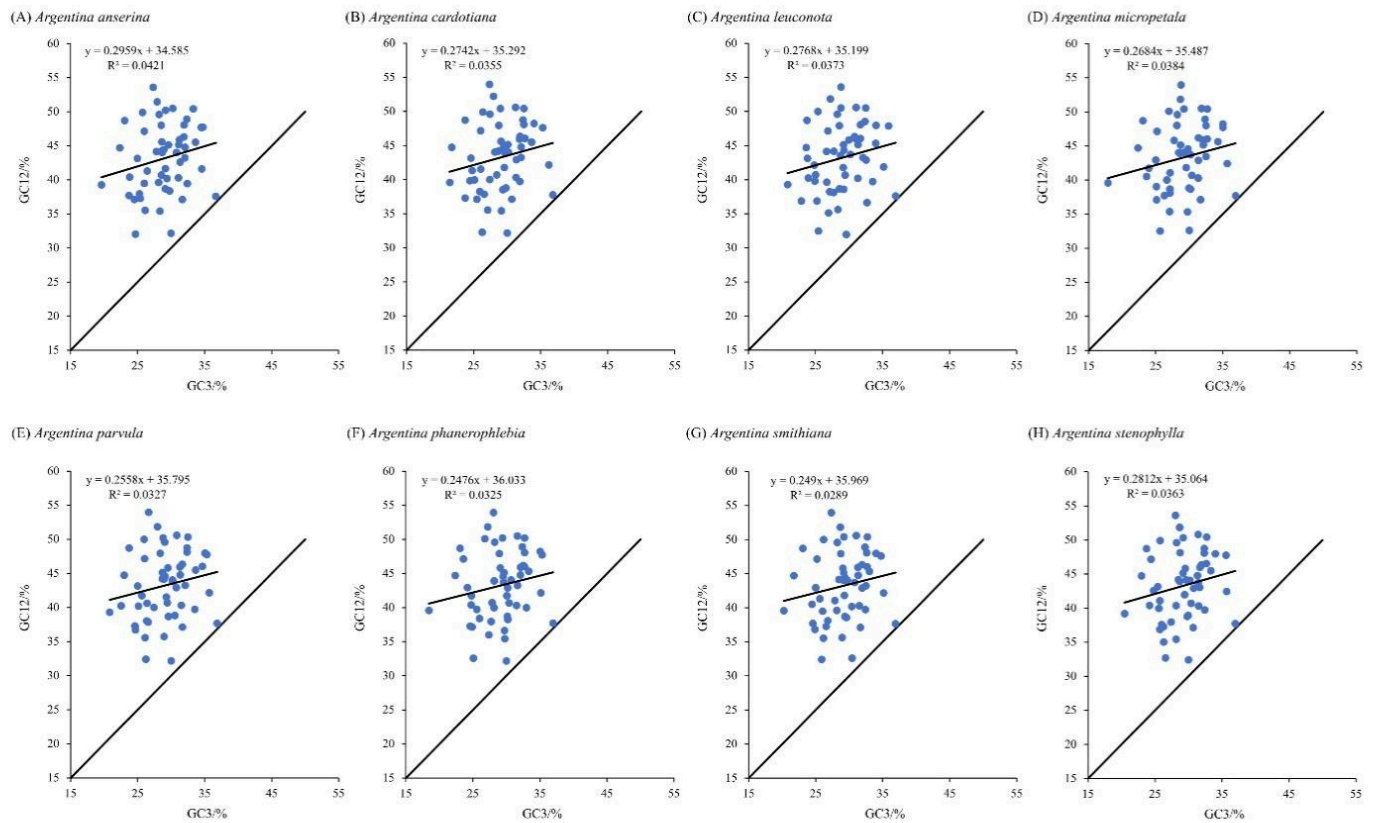


Figure 5

Neutrality plot of eight *Argentinas* species

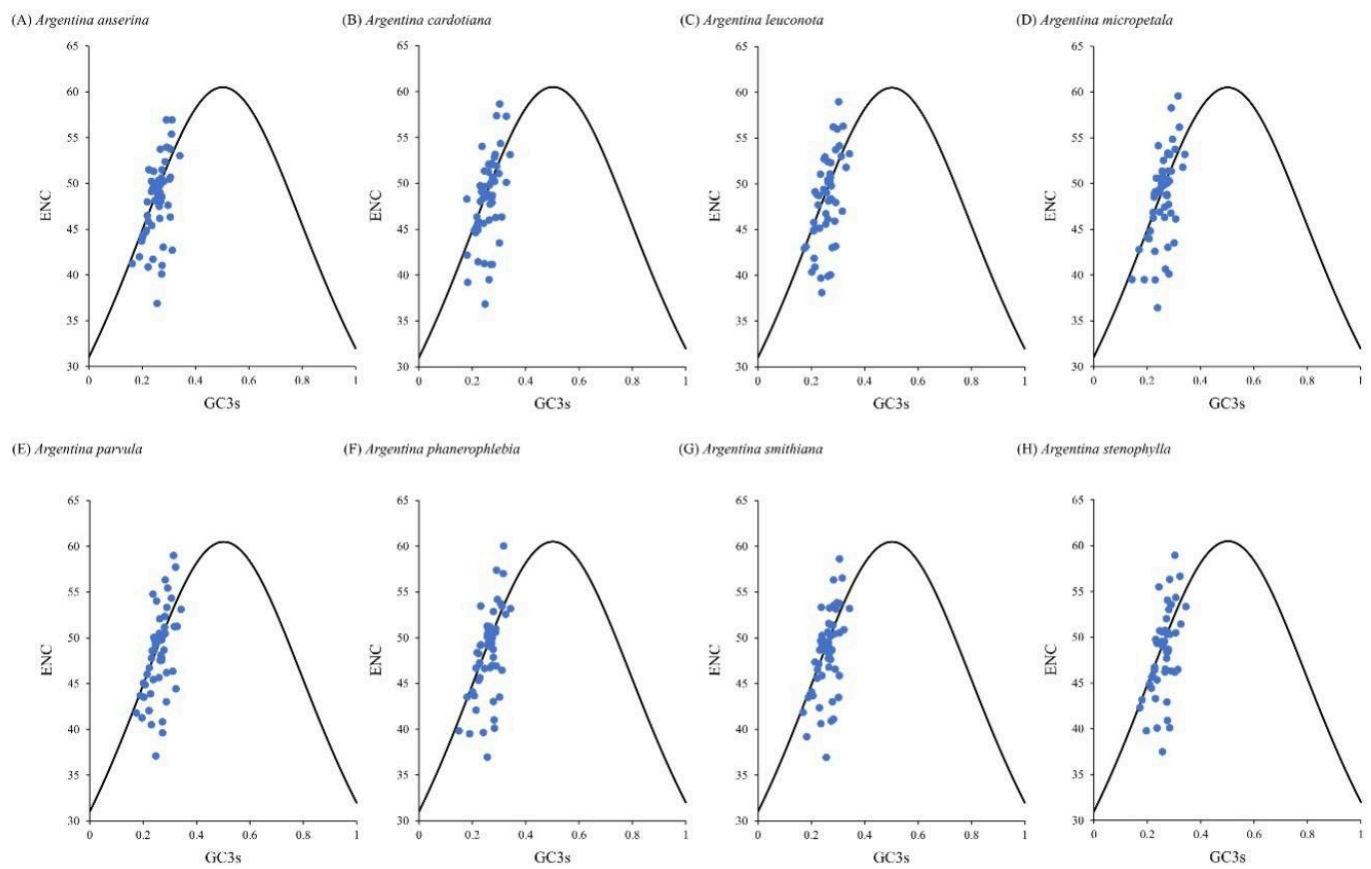


Figure 6

ENC-plot of eight *Argentina* species

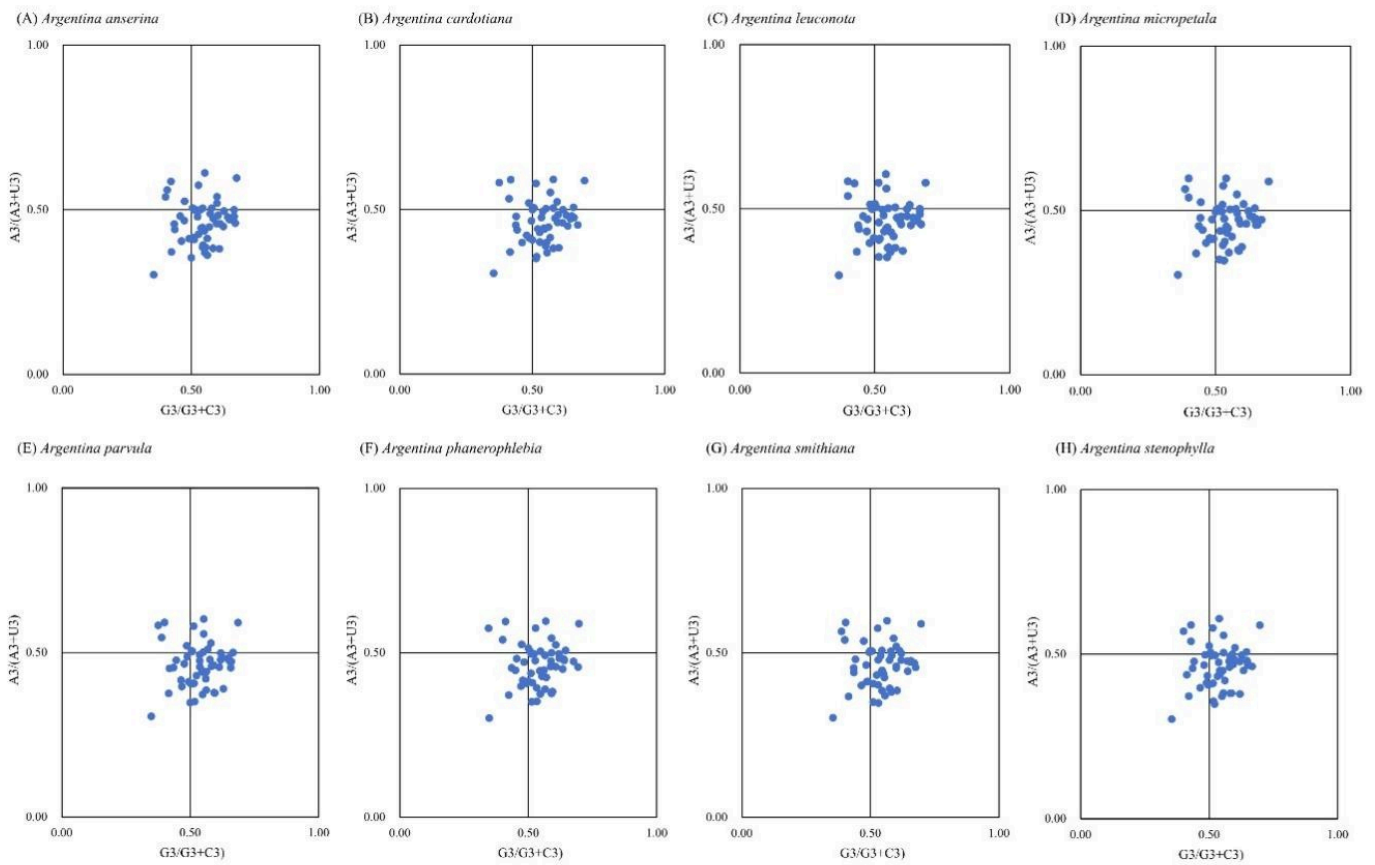


Figure 7

PR2-plot of eight *Argentina* species

Supplementary Files

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

- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)
- [TableS4.xlsx](#)
- [TableS5.xlsx](#)