Open Access

A reappraisal of the phylogenetic placement of the *Aquilegia* whole-genome duplication

Tao Shi[1,2]* and Jinming Chen[1,2]*

[1]Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China
Full list of author information is available at the end of the article
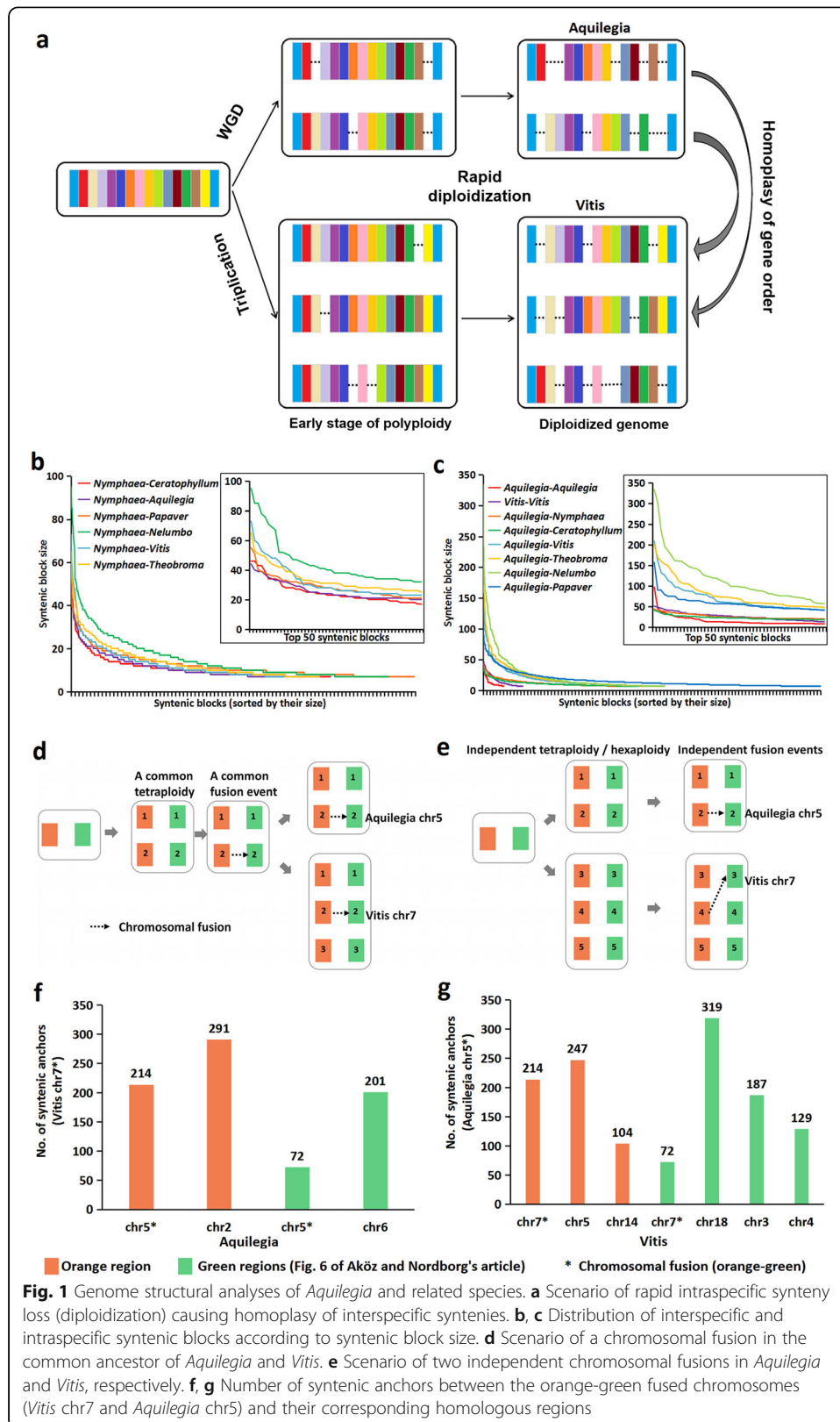
Abstract

The accurate placement of an ancient whole-genome duplication (WGD) in relation to the lineage divergence is important. Here, we re-investigated the *Aquilegia coerulea* WGD and found it is more likely lineage-specific rather than shared by all eudicots.

Whole-genome duplications (WGDs) are frequent and common in plants, contributing to the evolutionary novelty and adaptation to extreme environments [1, 2]. Aköz and Nordborg reported a common WGD in the ancestral eudicot: while this ancestral tetraploid is preserved in *Aquilegia* (a basal eudicot), the hexaploid in the ancestral core eudicot was formed by hybridization of this tetraploid and another diploid with a subsequent WGD [3]. Nevertheless, this is contradictory to the studies of *Nelumbo nucifera* [4, 5]. *Nelumbo*, another basal eudicot, having a much closer relationship with core eudicots [6], shows a slower synteny loss and substitution rate than *Aquilegia* and core eudicots when aligned to outgroup species including *Nymphaea colorata* [7], rice, and *Brachypodium distachyon* (Figs. 1b and 2e) [4], and thus, *Nelumbo* should preserve more of its traces. However, the *Ks* peak corresponding to the "shared WGD" is absent in *Nelumbo*, and only a lineage-specific WGD after the *Nelumbo-Macadamia* split was found [4] (Fig. 2a), which raises doubt about their hypothesis of the common tetraploid origin.

The authors' assertion that the *Aquilegia* WGD is shared by all eudicots is mainly based on clustering by gene order similarity or structural similarity within species genome and between species, which depends on synteny evolutionary rate. When the trait evolves with rate heterogeneity, the simple clustering by similarity can hardly reflect the true phylogeny. Therefore, the more rapid loss of intraspecific synteny in *Aquilegia* might cause homoplasy such that higher similarity in gene order of *Aquilegia-Vitis* than *Aquilegia-Aquilegia* can be observed (Fig. 1a). Indeed, *Aquilegia-Aquilegia* shows the most rapid decay of synteny with the smallest size and the fewest number of syntenic blocks, which is even smaller and fewer than older divergent pairs (*Aquilegia-Ceratophyllum* and

Fig. 1 Genome structural analyses of *Aquilegia* and related species. **a** Scenario of rapid intraspecific synteny loss (diploidization) causing homoplasy of interspecific syntenies. **b**, **c** Distribution of interspecific and intraspecific syntenic blocks according to syntenic block size. **d** Scenario of a chromosomal fusion in the common ancestor of *Aquilegia* and *Vitis*. **e** Scenario of two independent chromosomal fusions in *Aquilegia* and *Vitis*, respectively. **f**, **g** Number of syntenic anchors between the orange-green fused chromosomes (*Vitis* chr7 and *Aquilegia* chr5) and their corresponding homologous regions
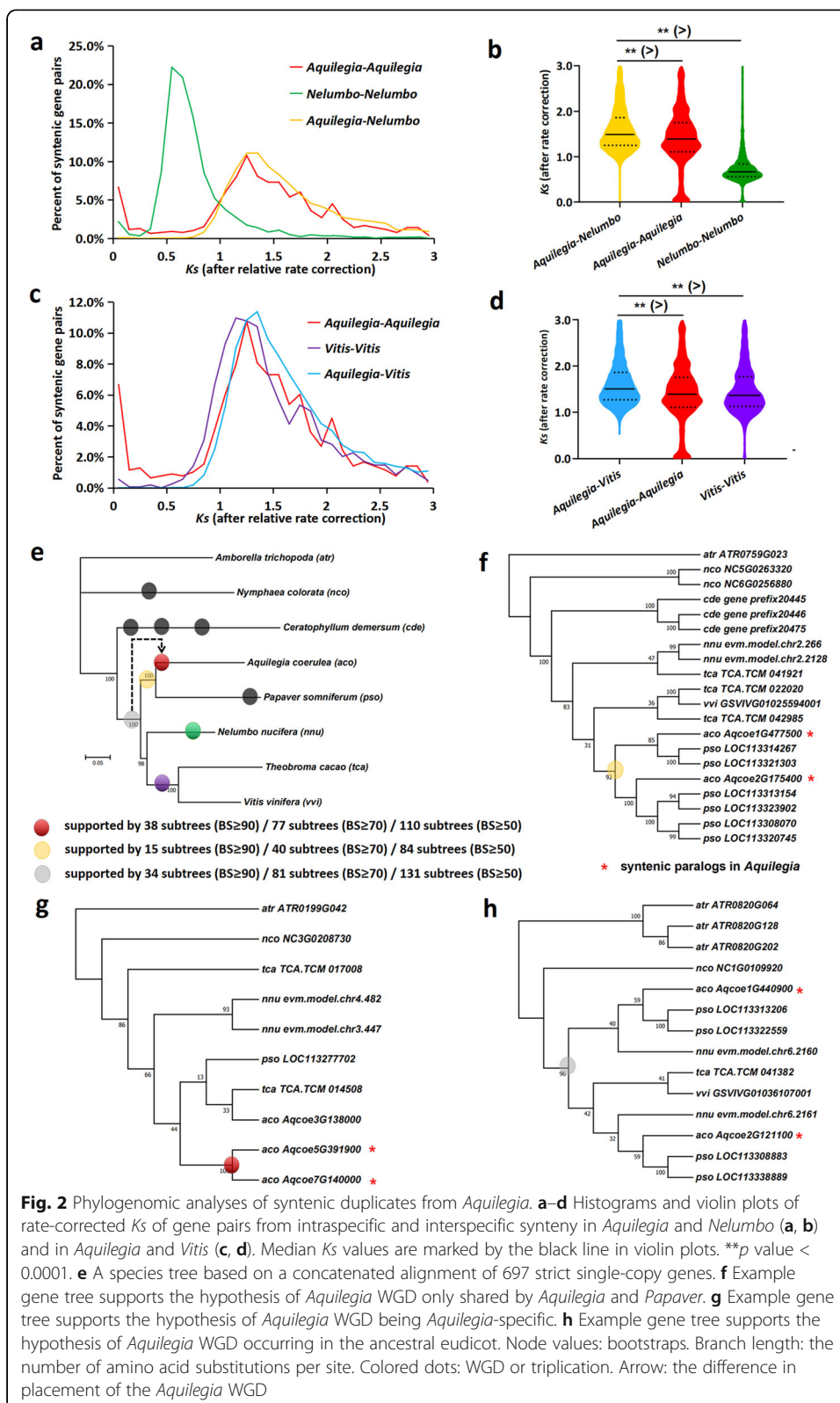
Fig. 2 Phylogenomic analyses of syntenic duplicates from *Aquilegia*. **a–d** Histograms and violin plots of rate-corrected *Ks* of gene pairs from intraspecific and interspecific synteny in *Aquilegia* and *Nelumbo* (**a**, **b**) and in *Aquilegia* and *Vitis* (**c**, **d**). Median *Ks* values are marked by the black line in violin plots. *******p* value < 0.0001. **e** A species tree based on a concatenated alignment of 697 strict single-copy genes. **f** Example gene tree supports the hypothesis of *Aquilegia* WGD only shared by *Aquilegia* and *Papaver*. **g** Example gene tree supports the hypothesis of *Aquilegia* WGD being *Aquilegia*-specific. **h** Example gene tree supports the hypothesis of *Aquilegia* WGD occurring in the ancestral eudicot. Node values: bootstraps. Branch length: the number of amino acid substitutions per site. Colored dots: WGD or triplication. Arrow: the difference in placement of the *Aquilegia* WGD

Aquilegia-Nymphaea) (Fig. 1c) [7]. Rapid convergent loss of the alternative copies, particularly housekeeping genes, by selection against redundancy after WGDs was found in angiosperms [8], and this is also congruent with more rapid syntenic gene loss that occurred soon after a WGD in teleost fish [9]. Therefore, their placement of a WGD simply by gene order similarity is flawed because of homoplasy caused by the heterogeneous rates of synteny loss. Meanwhile, the authors claimed that a fusion of ancestral eudicot chromosomes "green" and "orange" was observed in both *Aquilegia* chr5 and *Vitis* chr7, which is derived from a fusion in their tetraploid ancestor (Fig. 1d). Yet, they did not exclude an alternative possibility of two independent green-orange fusions in *Aquilegia* and *Vitis*, respectively (Fig. 1e). If their assumption is true, the fused green-orange chromosomes are expected to share closer homology (more syntenic anchors) to each other than the other green or orange regions. On the contrary, we found that the green and orange regions in the fused *Vitis* chr7 and *Aquilegia* chr6 share more anchors with non-fused green or orange regions (Fig. 1f, g). Therefore, the scenario of two independent chromosomal fusions is more likely.

Here, we showed that the *Aquilegia* WGD is more likely lineage-specific rather than common to all eudicots by two different phylogenomic approaches. (1) Ks values of syntenic paralogs and orthologs were measured to infer the sequential order of WGD and species divergence. Ks of each lineage were corrected using the relative rate (*Aquilegia* to *Nelumbo* to *Vitis* = 1:0.750:0.970) based on Ks branch length ratios of 1425 single-copy ortholog groups [10]. For comparison, $Ks > 3$ were removed to prevent substitutional saturation [11]. The Ks distances of syntenic orthologs for both *Aquilegia-Nelumbo* and *Aquilegia-Vitis* are significantly longer than intraspecific paralogs in *Aquilegia* (Mann-Whitney U test, both p value < 0.0001), suggesting this WGD exclusively occurred in *Aquilegia* (Fig. 2a–d). (2) To circumscribe *Aquilegia* syntenic duplicates in relation to the species divergence, 697 gene trees of ortholog groups from key taxa containing the *Aquilegia* syntenic paralogs were reconciled with the species tree using Notung2.9 (www.cs.cmu.edu/~durand/Notung) [12–14]. The species tree was constructed from 176 strict single-copy genes by OrthoFinder (www.stevekellylab.com/software/orthofinder) and IQ-TREE (www.iqtree.org) (Fig. 2e). When we applied bootstrap thresholds of 90, 70, and 50, we found the hypothesis that *Aquilegia* syntenic duplications occurred after the divergence of *Aquilegia* and the ancestor of *Nelumbo* and core eudicots is supported by 53, 117, and 194 subtrees, respectively (red and yellow dots in Fig 2e–g), whereas hypothesis that duplication occurred in the ancestral eudicot is supported by 34, 81, and 131 subtrees, respectively (gray dot in Fig. 2e, h). Due to the fact that *Aquilegia* WGD and *Aquilegia-Papaver* split are closely spaced in time, it is more difficult to resolve their sequential order. However, more subtrees support *Aquilegia* WGD being independent (red dot in Fig. 2e, g) than being shared by *Aquilegia* and *Papaver* (yellow dot in Fig. 2e, f), which is in line with a 2:2 syntenic relationship between *Aquilegia* and *Papaver* [15].

Authors' contributions
T.S. and J.-M.C. wrote the paper. The authors read and approved the final manuscript.

Author details
[1]Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China. [2]Center of Conservation Biology, Core Botanical Gardens, Chinese Academy of Sciences, Wuhan 430074, China.

References

1. Landis JB, Soltis DE, Li Z, Marx HE, Barker MS, Tank DC, Soltis PS. Impact of whole-genome duplication events on diversification rates in angiosperms. Am J Bot. 2018;105:348–63.
2. Van de Peer Y, Mizrachi E, Marchal K. The evolutionary significance of polyploidy. Nat Rev Genet. 2017;18:411–24.
3. Akoz G, Nordborg M. The Aquilegia genome reveals a hybrid origin of core eudicots. Genome Biol. 2019;20:256.
4. Shi T, Rahmani RS, Gugger PF, Wang M, Li H, Zhang Y, Li Z, Wang Q, Van de Peer Y, Marchal K, Chen J. Distinct expression and methylation patterns for genes with different fates following a single whole-genome duplication in flowering plants. Mol Biol Evol. 2020;37:2394–2413.
5. Ming R, VanBuren R, Liu Y, Yang M, Han Y, Li LT, Zhang Q, Kim MJ, Schatz MC, Campbell M, et al. Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). Genome Biol. 2013;14:R41.
6. One Thousand Plant Transcriptomes I. One thousand plant transcriptomes and the phylogenomics of green plants. Nature. 2019;574:679–85.
7. Zhang L, Chen F, Zhang X, Li Z, Zhao Y, Lohaus R, Chang X, Dong W, Ho SYW, Liu X, et al. The water lily genome and the early evolution of flowering plants. Nature. 2020;577:79–84.
8. De Smet R, Adams KL, Vandepoele K, Van Montagu MC, Maere S, Van de Peer Y. Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants. Proc Natl Acad Sci U S A. 2013;110:2898–903.
9. Inoue J, Sato Y, Sinclair R, Tsukamoto K, Nishida M. Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. Proc Natl Acad Sci U S A. 2015;112:14918–23.
10. Shi T, Huang H, Barker MS. Ancient genome duplications during the evolution of kiwifruit (Actinidia) and related Ericales. Ann Bot. 2010;106:497–504.
11. Tiley GP, Barker MS, Burleigh JG. Assessing the performance of Ks plots for detecting ancient whole genome duplications. Genome Biol Evol. 2018;10:2882–98.
12. Argout X, Salse J, Aury JM, Guiltinan MJ, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova SN, et al. The genome of Theobroma cacao. Nat Genet. 2011;43:101–8.
13. Yang Y, Sun P, Lv L, Wang D, Ru D, Li Y, Ma T, Zhang L, Shen X, Meng F, et al. Prickly waterlily and rigid hornwort genomes shed light on early angiosperm evolution. Nat Plants. 2020;6:215–22.
14. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449:463–7.
15. Guo L, Winzer T, Yang X, Li Y, Ning Z, He Z, Teodor R, Lu Y, Bowser TA, Graham IA, Ye K. The opium poppy genome and morphinan production. Science. 2018;362:343–7.

Publisher's Note