Research Article
Evolution of Genome Size in Duckweeds (Lemnaceae)

Wenqin Wang,1 Randall A. Kerstetter,1,2 and Todd P. Michael1,2

1 Rutgers, Department of Plant Biology and Pathology, The State University of New Jersey, The Waksman Institute of Microbiology, Piscataway, NJ 08854, USA
2 Monsanto Company, 800 North Lindbergh Boulevard, Creve Coeur, MO 63167, USA

Correspondence should be addressed to Todd P. Michael, todd.p.michael@monsanto.com

Received 4 February 2011; Revised 5 April 2011; Accepted 19 May 2011

1. Introduction

The Lemnaceae, commonly known as duckweeds, are the smallest, fastest-growing, and simplest of flowering plants. In this globally distributed aquatic monocot family (Figure 1(a)), there are 33 species representing five genera: Spirodela, Landoltia, Lemna, Wolffia, and Wolffia. Among them, Spirodela is the most ancestral, while Wolffia is the most derived [1]. The individual plants range in size from 1.5 cm long (Spirodela polyrhiza) to less than one millimeter (Wolffia globosa). Therefore, there is a successive reduction of morphological structures in parallel with evolutionary advancement within the family (Figure 1(b)). Duckweeds are not simply miniature versions of larger angiosperms; they represent a highly modified structural organization that resulted from the alteration, simplification, or loss of many morphological and anatomical features [2]. The biomass doubling time of the fastest-growing duckweeds in optimal growth conditions is less than 30 hours, nearly twice as fast as other “fast” growing flowering plants and more than double that of conventional crops [3].

Before the days of Arabidopsis, duckweeds, and more specifically Lemna, were an important model system for plant biology [4]. Since duckweeds are small, morphologically reduced (although with root and leaf-like structure), fast growing, easily cultivated under aseptic conditions (Figure 1(c)), transformable, crossable, and particularly suited to biochemical studies (direct contact with media), it is an ideal system for biological research [5]. Much of what we know about photoperiodic flowering responses comes from fundamental research conducted on Lemna by the preeminent plant biologist Dr. William Hillman at the Brookhaven National Laboratories [6]. Some of the current uses of Lemnaceae are a testimony to its scientific, commercial, and biomass utility: basic research and evolutionary model system [7], toxicity testing organism [8], biotech protein factories [9], wastewater remediation [10], high protein animal feed, carbon cycling [5], and biofuel potential candidates [11].

The advent of high-throughput sequencing technologies has enabled a new generation of model plant systems [12]. In an effort to initiate duckweed genomic research, we endeavoured to identify species with small genomes that would be ideal for sequencing. First, we queried the Kew plant genome database (http://data.kew.org/cvalues/) and found that there were only 6 duckweed accessions that
Figure 1: Duckweeds are small aquatic plants that are widely distributed in nature and amenable to culturing in the lab. (a) Duckweeds growing in the Raritan Canal River, Piscataway, NJ, USA. This population of duckweed includes *Wolffia*, *Spirodela*, and *Lemna*. (b) The relative size of *Spirodela*, *Landoltia*, *Lemna*, *Wolffiella*, and *Wolffia* in the order of phylogeny as compared to an American Quarter. (c) Sterile *Spirodela polyrhiza* grown in the Schenk and Hildebrandt basal salt medium.

had been measured by the Feulgen method [13, 14]. DNA content of single species from each genus was determined and showed obvious difference. Due to it being laborious and time consuming, the popularity of Feulgen technique has waned. Feulgen has been largely replaced by flow cytometry (FCM) [15], a faster, easier, and more accurate method and the current preferred technique for genome size estimations and DNA ploidy analyses in plants [16].

In order to find the smallest duckweed genome for sequencing and also explore previous observations about genome complexity in duckweeds, we estimated the genome size of all of the five duckweed genera using FCM. These genome size measurements will form the foundation for future work in sequencing duckweed genome, and enabling duckweeds as a model and applied system.

2. Materials and Methods

2.1. Plant Materials. 115 accessions of 23 duckweed species representing all 5 genera were measured in this study (Table S1 in Supplementary Materials available online at doi:10.1155/2011/570319) (They provide details of sample collection and results of nuclear DNA content measurements for *Lemnaceae*). Elias Landolt collected most of the duckweed accessions described in this work over the past 50 years (Landolt Duckweed Collection) [2]. Accessions were either obtained directly from Elias Landolt, BIOLEX (NC, USA) or The University of Toronto Culture Collection of Algae and Cyanobacteria (UTCC). Currently, the Landolt Duckweed Collection has been moved to Rutgers University. Additional lines were collected from lakes and wastewater
ponds by TPM and WW (NJ, USA). Plants were grown aseptically for 2 weeks with 1/2 full concentration of Schen and Hildebrandt Basal Salt mixture (Sigma, USA) liquid culture medium under short day growth condition (8 h light and 16 h darkness with constant temperature 23°C). We bar-coded all the determined and undetermined species by identification of polymorphisms of chloroplast atpF-atpH noncoding spacer [17].

2.2. Isolation and Staining of Nuclei. To estimate nuclear DNA contents with flow cytometry (FCM), sample tissue nuclei were stained with propidium iodide (PI) [18]. Briefly, 10 mg of fresh duckweed tissue and the same amount of the internal standard were chopped simultaneously with new razor blades and isolation buffer in a plastic Petri dish [19]. Isolates were filtered through a 30-µm nylon mesh into an Eppendorf tube. The suspensions of nuclei were stained with 50 µg mL⁻¹ PI mixed with 50 µg mL⁻¹ RNase (R4875, Sigma). The samples were incubated on ice for a few minutes before estimation by FCM.

2.3. Analysis of Nuclear DNA Content by FCM. PI-stained nuclei were analyzed for DNA content with a Coulter Cytomics FC500 Flow Cytometer (Beckman Coulter, Inc., Miami, Florida, USA). In all experiments, the fluorescence of at least 3000 G1-phase nuclei was measured. DNA content of each target sample was calculated by comparing its mean nuclear fluorescence with that of an internal standard (Figure 2(a)). We utilized internal controls that closely match the duckweed genome sizes being measured to ensure accuracy. The internal standard is a Brachypodium distachyon line, (Bd21, 300 Mbp) [16], Arabidopsis thaliana Columbia., (At, 147 Mbp) [20], and Physcomitrella patens ssp patens, (Pp, 480 Mbp) [21]. The numbers in bracket were generated by our flow cytometry equipment and our methods. Therefore, the validated genome sizes are not exactly the same but very close to cited references. Both duckweed and internal standards have very little secondary compounds, which will interfere with quantitative DNA staining. The absolute DNA content of a sample is calculated based on the values of the G1 peak means:

\[
\text{Sample 1C DNA content} = \left( \frac{\text{sample G1 peak mean}}{\text{standard G1 peak mean}} \right) \times \text{standard 1C DNA content (Mbp)}.
\]

At least, three independent biological replicates for each sample were analyzed on different days to estimate the mean DNA content. The transformation factor from pg to Mbp is: 1 pg = 978 Mbp [22].

2.4. Statistical Analysis. Data on intraspecies variation of genome size were analysed by ANOVA: single factor test. To test whether genome size variation was correlated with geographic location or altitude of populations, the Spearman correlation coefficient (r) was used.

3. Results

3.1. Intra- and Interspecies Variations of Genome Sizes. The genome sizes of 115 accessions from 23 species representing 5 genera were estimated by FCM (Table S1). The DNA content estimates varied nearly thirteen-fold, ranging from 150 Mbp in Spirodela polyrhiza to 1,881 Mbp in Wolffia arrhiza. We superimposed the estimated 1C-value on a phylogenetic tree for Lemnaceae based on combination of morphological, flavonoid, allozyme, and DNA sequence analysis [1] and found that there is a continuous increase of DNA content in order of Spirodela, Landoltia, Lemna, Wolffia, and Wolffia, which correlates well with the morphological reduction within the family (Figures 3(a) and 3(b)).

In the genus Spirodela, we measured genome size for 34 accessions and found that the 1C DNA content only varies from 150 to 167 Mbp (Figure 3(a); Table S1). The analysis of variance (ANOVA: single factor test) revealed that there was not a significant difference in Spirodela polyrhiza genome sizes (P > 0.05). Similarly, the 1C DNA content for 19 accessions of Landoltia punctata from 372 to 397 Mbp did not show significant variation (Figure 3(a); Table S1). In the genus Wolffia, the genome sizes range from 623 Mbp to 973 Mbp (Figure 3(a)), which is almost as 4–6 times large as Spirodela polyrhiza. Like Spirodela polyrhiza and Landoltia punctata, there are no obvious intraspecific genome size variations in Wolffia hyalina and Wolffia lingulata. In the genus Wolffia, we measured 11 species and found that they have the largest genome sizes on average among the duckweed family (Figure 3(a)). 5.3-fold difference was observed from Wolffia australiana (357 Mbp) to Wolffia arrhiza (1,881 Mbp).

In the genus Lemna, 7 species were investigated. There is a large amount of genome size variation in this genus. Lemna validiviana has the smallest genome size (323 Mbp), while Lemna aqueinocitadlis has the biggest (760 Mbp). Surprisingly, intraspecific genome-size fluctuations are also impressive. For Lemna minor, 26 accessions have genome sizes ranging from 356 to 604 Mbp with up to 69.6% of the intraspecific DNA content variance. We confirmed the intraspecific difference of them by randomly choosing 2 Lemna minor with simultaneous measurement of both accessions (26.0% difference between 6591 Lm and 7436 Lm, Figure 2(b)). Statistical analyses revealed significant differences among the Lemna minor accessions (P < 0.01). As well, Lemna aqueinocitadlis (424–760 Mbp, 79.2%) (Figure 2(c)), Lemna trisulca (446–709 Mbp, 59.0%), and Lemna japonica (426–600 Mbp, 40.8%) all show intraspecific difference, indicating a drastically uneven evolution of intraspecific genome expansion in Lemna.

3.2. 1C-Value and Latitude, Longitude, and Altitude. To investigate whether there is a correlation between genome-size variations and the geographic distribution in the duckweed, we compared genome size estimates with the latitude, longitude, and altitude of recorded collection. However, genome size variation was not correlated with
latitude by Pearson coefficient ($r$-value: *Spirodela* = −0.05, *Landoltia* = 0.17, *Lemma* = −0.07, *Wolffiella* = −0.17, *Wolffia* = 0.34) (Figure 4(b)), nor with longitude ($r$-value: *Spirodela* = 0.17, *Landoltia* = 0.04, *Lemma* = 0.26, and *Wolffia* = −0.41) (Figure 4(c)) except *Wolffiella* with a high $r$-value −0.86 possibly due to limited accessions ($n = 8$). No correlation was found between C-values and altitude, either ($r$-value: *Spirodela* = 0.13, *Landoltia* = −0.25, *Lemma* = −0.33, *Wolffiella* = −0.41, and *Wolffia* = 0.13) (Figure 4(d)). It is interesting we found that most of *Spirodela, Landoltia, Wolffiella,* and *Wolffia* were collected from a similar geographic range between 0° to 45° and preferred to localize above 600 m to 1200 m of altitude. In contrast, most of *Lemma* species were collected between 30° to 60° and preferred to distribute below 600 m. However, this most likely represents a sampling bias and could also explain the absence of a relationship between genome size and the environment in duckweed.

### 4. Discussion

#### 4.1. Genome Evolution in Duckweeds

In the phylogeny of *Lemnaceae*, there is a strong relationship observed between genome size evolution and morphological progression. We found that the ancestral genus *Spirodela* has the smallest genome size, while the most advanced genus *Wolffia* contains biggest genome size (Figure 3; Table S1), which correlates with the morphological reduction rather than organism
Figure 3: Genome size variation across the duckweeds. Estimated 1C-value superimposed on a phylogenetic tree for *Lemnaceae* based on combination of morphological, flavonoid, allozyme, and DNA sequence analysis [1]. The species in black were what we tested, and the species in the grey were the ones we did not examine in this experiment. In the bracket is the number of different accessions we tested. (b) Average genome sizes (y-axis) of duckweed species negatively parallel with degree of primitivity (x-axis). Duckweed species are arranged on the x-axis from lower to higher evolutionary status, which deduced from primitive and derived morphological traits [13].

complexity within the family. This result is consistent with Geber’s finding, which showed that there was a relationship between DNA content and degree of primitivity [23].

Genome doubling has been a pervasive force in plant evolution, which has occurred repeatedly [24]. Even the smaller genome of *Arabidopsis thaliana* has been impacted by genome duplication [25]. Cytological variation by counting the chromosomes was extensively investigated within duckweed. They concluded that polyploidy ($2n = 20, 30, 40, 50, 60$, and $80$) is the main intrapopulational variation [2], which means polyploidization was very active and occurred in the duckweeds for multiple rounds in the past.
After polyploidization, transposable element mobility, insertions, deletion, and epigenome restructuring contribute to the successful development of a new species and also genome size changes [26]. Changes in genome structure could lead to differential gene loss, extensive changes in gene expression [27], and have immediate effects on the phenotype and fitness of an individual [28]. It is likely polyploidy might drive the divergence during duckweed evolution.

4.2. Geographic Distribution and Genome Size Variation. It was suggested that variation in DNA content has adaptive significance and is correlated with the environmental traits of species [29]. The environmental conditions of plants are to a large extent determined by latitude, longitude, and altitude. Previous studies have indicated a positive correlation between genome size and latitude (associated with the length of sun light with the growing season and the temperature) and also altitude (associated with the temperatures) among plant species. For example, the increase of DNA content corresponded with the increasing latitude found in the Pinaceae family [30] and with increasing altitude observed in Zea mays [31]. Duckweeds are distributed broadly around the world (Figure 4(a)). Our result shows that there is no significant overall correlation of genome size with latitude, longitude, and altitude (Figure 4). The same result was found in Vicia faba [32], Sesleria albicans [33], and Asteraceae [34]. A summary revealed that these relationships were not straightforward and not clear. Five studies (Picea sitchensis, Berberis, Poaceae, and Fabaceae, Tropical versus temperate grasses, 329 tropical versus 527 temperate plants) found positive, seven (Arachis duranensis, Festuca arundinacea, North American cultivars of Zea mays, 162 British plants, 23 Arctic plants, 22 North American Zea mays, and 11 North American Zea mays) found negative, and five (Allium cepa, Dactylis glomerata, and Helianthus) found nonsignificant correlations between genome size and latitude. Additionally, nine were positive, eight were negative, and six were not statistically significant between genome size and altitude [35]. But
the different environmental distribution of the *Lemna* genus (30° to 60° of latitude and below 600 m of altitude) with the other four duckweed genera (0° to 45° of latitude and 600 m to 1200 m of altitude) might explain the large intraspecific genome size variation.

### 4.3. Intraspecific Variation in Genome Size

Intraspecific genome consistency has been reported in *Allium cepa* [36], *Glycine max* [37], and *Capsicum* and *Gossypium* [38]. We also found a similar result for *Spirodela polyrhiza*, *Landoltia punctata*, *Wolfiella hyalina*, *Wolfiella lingulata*, and *Wolfia australiana*, which do not have statistical intraspecific differences in genome size (Table S1). One explanation is that these species have a mechanism to maintain genome size constancy, for example, by intraspecific stabilizing selection on genome size [39]. On the other hand, we found obvious intraspecific variation in *Lemna minor*, *Lemnaceae aequinoctialis*, *Lemna trisulca*, and *Lemna japonica*. Some artifacts of intraspecific variation in genome size have been noted, such as environmentally induced variations, secondary compounds and fluorescence staining inhibitor, and erroneously determined species [15, 40]. However, our experiments are not complicated by these factors. We developed an easy bar-coding method to correctly identify duckweed species, which allowed us to correct any mismapped duckweed in the collection [17]. As cytosolic components may change in response to changes in the environment, we grew the duckweed plants under identical conditions. We used internal standardization such as *Brachypodium distachyon*, *Arabidopsis thaliana*, and *Physcomitrella patens* that were prepared simultaneously and under the same experimental conditions as the duckweed accessions. Both duckweed and the internal standard have very little secondary compounds, which may affect genome size estimates. Additionally, we performed biological replicate on different days to eliminate instrument bias. In addition, intraspecific differences were independently confirmed by simultaneously measuring two accessions of the same species by FCM (Figures 2(b) and 2(c)).

The intraspecific variation may result from different numbers of repeated sequences, including satellite DNA [41], transposable elements [42], and ribosomal genes [43]. Large-scale polymorphism of heterochromatic repeats exist in the DNA of *Arabidopsis thaliana* and could account for about 50% of the variance among the *Arabidopsis thaliana* accession [44]. In addition, the amount of rDNA accounts for the differences in genome size between closely related lines of *Linum usitatissimum* (flax) [43]. The activity of transposable elements (TE) potentially multiply 20~100 times (~0.1~1 Mbp) in a single generation [45]. For example, the BARE-1 TE is positively correlated with genome size within wild barley (*Hordeum spontaneum*) in response to sharp microclimatic divergence [42]. Deletions and insertions (INDELs) are most likely not candidates for genome size differences in duckweed. In *Drosophila melanogaster*, genome loss is only less than 1 bp per generation [46], indicating a small contribution to genome-size variation. However, in the fast growing duckweeds, which only need 2~5 days for each generation, one could imagine it is more likely that TE have higher rate than other flowering plants to influence genome size within and between species.

### 5. Conclusion

This is the first extensive analysis of genome sizes in duckweeds and examination of genome size variations across a range of taxonomic levels. We showed that duckweeds, in general, have remarkable smaller genome size compared with other flowering plants. The smallest genome size of *Spirodela polyrhiza*, combined with its sterile and controllable culture, fast growing, and promising application in research, suggest that this species may be good candidates for ongoing whole-genome sequencing projects and a model experimental tool. The 150 Mbp *Spirodela polyrhiza* genome is being sequenced by the DOE-JGI community-sequencing program (CSP), which will address challenges in alternative energy, bioremediation, and global carbon cycling. Also, the availability of a DNA C-values database of duckweeds and a consensus higher-level phylogenetic tree has opened the way for exploring the general processes underlying the evolution of genomes. Obvious intraspecific variation in duckweeds will also provide nice material to study the mechanism of within-species and between-species variation in genome size. However, the main force driving the intraspecific variance and how the genome size affects the phenotype still requires more research.

### Acknowledgments

We thank Elias Landolt from Institut für Integrative Biologie (Zürich, Switzerland), Lynn Dickey, Nirmala Rajbhandari, and Peaches Staton from BIOLEX (NC, USA), and Judy Acreman UTCC (Toronto, Canada) to generously provided duckweed accessions. We are grateful to Theresa Hyejeong Choi from EOHSI (Rutgers University) for assistance with flow cytometry. We would like to thank Justina Marcinow, Marina Ermakova and Yiheng Yan for experimental help. We thank Julia Trumbull for maintaining the Landolt Duckweed Collection at Rutgers. This work was supported by an Energy Fellowship and a Waksman Fellowship to W. Wang, and startup funds from the Waksman Institute of Microbiology to T. P. Michael.

### References


