Wolffia, a minimalist plant and synthetic biology chassis

Eric Lam[1] and Todd P. Michael[2]

[1]Department of Plant Biology, Rutgers the State University of New Jersey, New Brunswick, NJ 08901, USA; [2]The Plant Molecular and Cellular Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

Eric Lam: 0000-0001-8462-9794
Todd P. Michael: 0000-0001-6272-2875

Correspondence:
eric.lam@rutgers.edu (E. Lam)
tmichael@salk.edu (T.P. Michael)

KEY WORDS: Duckweed, synthetic biology, minimalist plant, reduced morphology, diurnal regulation

GLOSSARY:

Chassis: Skeletal framework on which parts and/or components can be fastened and connected to perform specific functions.

Synthetic biology: Multidisciplinary science that seeks to create or redesign biological parts, pathways or systems to produce a desired functional output.

Artificial chromosomes: DNA constructs produced in laboratories that contain all the necessary parts and components in order for them to perform critical functions of natural chromosomes in a cell.

Top-down engineering: An approach to create synthetic chromosomes by starting with a natural, functional chromosome and then systematically deleting and replacing different portions with synthetic DNA constructs using a variety of strategies.

Bottom-up engineering: An approach to create synthetic chromosomes by synthesizing and assembling the various parts of a chromosome, such as centromere and telomeres, in bacteria or yeast platforms before mobilizing the construct into a recipient cell.

Duckweed: Common name of plants in the Lemnaceae family that are represented in five genera: *Spirodela*, *Landoltia*, *Lemna*, *Wolffia* and *Wolffiella*; there are a total of 36 different recognized species across the five genera.

***Wolffia*:** One of the five duckweed genera commonly known as Watermeal or Khai Nam and traditionally consumed in Southeast Asia for its high protein content.

Frond: Main body of a duckweed plant that is akin to a fusion between a leaf and stem.

Mother frond: Since duckweed predominantly grows through a yeast-like budding process, this is the original frond with respect to the new fronds.

Daughter frond: New frond generated by a mother frond and initially connected to the mother frond in duckweed via the specialized tissue called a stipe.

Meristem: A region of actively dividing cells forming new daughter fronds from the mother frond.

Pocket: An internal region of a duckweed frond, analogous to a bract or specialized leaf, where the meristem is located and where daughter fronds normally emerge.

Stipe: An internode-like structure that connects the daughter frond to the mother frond at the meristem.

Telomere: Arrays of repeat sequences typically 7 base pairs (bp) long with the sequence AAACCCT that are found on the ends of chromosomes to protect against exonuclease degradation, chromosomal fusions, as well as eliciting DNA damage response.

Centromere: Region of a chromosome to which the microtubules of the spindle attach, via the kinetochore, during cell division.

CENH3: The histone H3 variant that functionally defines a centromere in eukaryotic genomes.

Holocentric chromosomes: Small and distributed centromeres enabling microtubules to bind over the length of the chromosome; associated with species that display rapidly evolving chromosomes.

Ribosomal DNA (rDNA): Highly repetitive gene clusters in eukaryotic genomes responsible for making ribosomal RNA that is essential for protein synthesis.

Transposable elements (TEs): DNA sequence found in genomes that are remnants of ancient viruses that actively move or "jump" around the genome resulting in evolutionary innovations and genome bloating in plants.

Time-of-Day (TOD) control: The internal circadian clock integrates external environmental signals such as light/dark and temperature cycles to partition, or gate, internal biology such as gene expression, biochemistry and metabolism to specific times over the day.

ABSTRACT

A highly simplified species for genome engineering would facilitate rational design of a synthetic plant. A candidate species is the aquatic, non-grass monocot wolffia (*Wolffia australiana)* in the Lemnaceae family. Commonly known as Watermeal, wolffia is a rootless ball of several thousand cells the size of a pinhead and the fastest growing plant known on Earth. Its extreme morphological reduction is coupled to transposon mediated streamlining of its transcriptome, which represents a core set of non-redundant protein coding genes. Despite its body plan and transcriptome being highly specialized for continuous growth, wolffia retains cell types relevant to higher plants. Systems level studies with this species could enable the creation of a defined biological chassis for synthetic plant construction.

Designer genomes

Designer genomes are powerful platforms for fundamental research into the intricacies of life as well as to facilitate novel applications [1]. The first organism with a totally synthetic genome was reported in 2010 with the one megabase (Mb) prokaryotic *Mycoplasma mycoides* genome [2]. This advance was then extended to recode the codons of the molecular biology workhorse, *E. coli*, to enable the expansion and creation of a new amino acid that can be incorporated into proteins [3]. Progress has also been made to replace all 16 chromosomes, with 12 Mb of DNA total, in the yeast *Saccharomyces cerevisiae* leading to the first synthetic eukaryote [4]. Along the way of this combinatorial approach to assemble multiple synthetic chromosomes into a single yeast, strains with partial synthetic chromosomes have already yielded valuable insights and useful gene variants from directed evolution, demonstrating the power of this approach [5,6].

In more complex eukaryotes, such as plants and animals, the challenge of creating synthetic organisms is much greater, due to their larger genomes and higher complexity. Nevertheless, significant progress has been made in the past decade for construction of HACs (Human **Artificial Chromosomes**) and PACs (Plant **Artificial Chromosomes**), aided by improvements in synthesis and assembly of designer genome segments in bacteria- and yeast-based platforms [7,8]. For plants, the epigenetic nature of how functional **centromeres** are defined has hampered a **bottom-up engineering** approach to synthesize PAC from its component parts, although more recent innovations in *Drosophila melanogaster* deploying LacO arrays and LacI-**CENH3** fusion protein may enable the targeted recruitment of necessary factors to create a functional centromere in eukaryotes [9–11]. In addition, a **top-down engineering** approach utilizing **telomere** truncation via transgene insertion has shown good potential to replace whole arms of plant chromosomes by large insertions using a combination of site-specific recombinases to drive serial integrations of large synthetic fragments [12].

To facilitate the complex endeavor of creating a synthetic plant genome in the near future, we believe it would be advantageous to identify the simplest model plant possible in terms of having a minimal number of genes and lowest complexity in body plan, tissue types, and growth control pathways. These characteristics would help in determining the minimum number of genes required for a photoautotrophic plant, much like what has been achieved for bacteria using the synthetic genome approach [13]. Here we propose the plant analog of *S. cerevisiae*, the duckweed *Wolffia australiana*, as a next generation model plant based on its small size (~1mm), short doubling time (<24 hrs), limited cell number (~4,000), aquatic growth (for simple manipulations), small genome size (375 Mb), fewest annotated genes (15k), transformability, and presence of anatomical analogs for key plant tissues (Figure 1) [14]. Our recent studies with this aquatic monocot indicate that this minimalist plant could be envisioned as a **chassis** platform where component parts can eventually be rapidly added in or taken out with predictable

outcomes.

Duckweeds are historical plant models readily amenable to broader impact applications

Before arabidopsis (*Arabidopsis thaliana*) was the model plant of choice, species in the Lemnaceae family commonly known as **duckweed** were important for reductionist approaches to plant biology [15]. Flowering time, circadian biology, photosynthesis, and phytohormone biology were elucidated in different species of duckweed due to their small size, fast growth, and rapid uptake of labelled compounds from their growth media [16,17]. Recent publication of high-quality genomes for two clones of *W. australiana* revealed streamlined genomes with only ~15,000 conserved, non-redundant protein-coding genes [14]. Coupled to the fact that wolffia species readily flowers, makes crosses, is transformable, and requires simple growth conditions, we propose here that it will be an important addition to the model plant toolbox for next generation molecular studies and **synthetic biology**. While the duckweed family has attributes that facilitate their use as an excellent model plant platform, they also hold great promise to provide a new crop system that can augment traditional agriculture to provide sustainable bioproducts [17]. Moreover, as illustrated by the long-running Waksman Student Scholar program at Rutgers University of New Jersey (https://wssp.rutgers.edu/), duckweed is highly amenable to a classroom setting for hands-on educational programs that integrate learning and research applications for the benefit of society at large. In the past decade, the duckweed research community has also been generating a wide array of publicly accessible resources such as germplasms, community newsletter, and genomics datasets that include high quality reference genomes and transcriptomes [17]. The most morphologically reduced, smallest, and fastest growing duckweed genus ***Wolffia*** thus has the potential to become the yeast of reductionist plant biology for basic and applied research [18].

Reduced Wolffia morphology still represents anatomical analogs for key plant tissues

Wolffia is the fastest growing plant known on Earth, with some species doubling in under 24 hours [19]. *Wolffia* primarily grows through a yeast-like budding process where a **mother frond** gives rise to a **daughter frond** that stays attached until the daughter frond gives rise to a grand-daughter frond (Figure 1C). It has been hypothesized that if *Wolffia* had access to unlimited nutrients and CO_2, it could give rise to 10^{30} plants in 4 months, a volume roughly equivalent to Earth [19]. For this reason, Duckweed has attracted attention for its potential as a biofuel crop grown on wastewater, where it plays the dual role of phytoremediation due to its ability to withstand and accumulate high levels of contaminants [20].

In essence, *Wolffia* is a green ball of about 4,000 cells without any visible organs, lacking both roots and vasculature with only a photosynthetically active stem-leaf structure called a **frond**, the size of a pinhead (~1 mm). However, it has been theorized that the frond can be interpreted in relation to the anatomical features of a typical plant [21]. The most accepted interpretation to date posits that the frond is an axillary stem with bract-like region making up the **pocket** where the daughter frond is attached through an internode-like structure called a **stipe** (Figure 1B). Therefore, the progression of mother to daughter to grand-daughter can be viewed as similar to the stem and leaf structure of a plant rosette (Figure 1E,F).

While *Wolffia* only has a few morphologically-distinct tissues, it does contain critical cell types important to plant research. The **meristematic** cells are located close to the stipe and continuously give rise to new daughter fronds (Figure 1B-D). New daughter fronds start to form at the 18-cell stage and differentiate at the 30 cell stage giving rise to upper and lower epidermal as well as parenchyma cells. The parenchyma cells further differentiate to palisade and spongy cell types, the latter of which result in larger cells at the ventral portion of the frond with more vacuolar space and fewer chloroplast, leading to its lighter green color and providing

buoyancy as well as stability (Figure 1B). Because *Wolffia* is not completely submerged, it has 15-30 stomata on its dorsal epidermis. Several other cell types are hypothesized based on histochemical and morphological observations (starch, papillae, etc.), for which more advanced technologies such as single cell sequencing and micro-CT could help to elucidate their complexities [22,23].

Although *Wolffia* primarily grows in a yeast-like budding process, it also can flower in nature and flowering can be induced under laboratory settings to facilitate genetic studies [24,25]. In most duckweeds, a flowering frond has slightly different characteristics from a non-flowering frond in terms of size, air pockets, and number of stomata, suggesting that flowering is determined prior to the release of a particular daughter frond. Many *Wolffia* species make two flowers, while some just one, and they occupy most of the dorsal epidermal space when flowering occurs (Figure 1D). Each flower emerges from a separate pouch or cavity and makes one stamen and pistil; in general, one seed is produced per fruit and is desiccation-tolerant. *Wolffia* thus has all the basic tissues for a standard plant making it a good candidate as a ready canvas for angiosperm synthetic biology.

Wolffia has a streamlined plant genome ideal for synthetic biology

There are eleven known species in the *Wolffia* genus with genome sizes spanning an order of magnitude from 375 to 2,203 Mb and representing several ploidies [17,26,27]. The average haploid genome size is ~800 Mb with a base chromosome number of 20 (2n=2x=40). However, some species have triploid (2n=3x=60) and tetraploid (2n=4x=80) representatives, which may partially explain the span in genome sizes within the genus. Both sequenced clones of *W. australiana* are diploid and have a haploid genome size at ~375 Mb as estimated by both flow cytometry and k-mer analysis [14,26], which is similar in size to that in rice, while twice the size of those in the model plant arabidopsis and the Greater Duckweed *S. polyrhiza* (Table 1).

The first duckweed genome to be sequenced was that of spirodela since it has the smallest genome and occupies a basal position in the Lemnaceae family [28–31]. Subsequently, representative genomes for all five genera are either published or publicly accessible, including reference quality genomes for two clones of *W. australiana* (wa8730 and wa7733) [14,22,28,32]. The genomes of *W. australiana* are found to be collinear with those of the other duckweed genera (Figure 2A), despite a reduced number of conserved non-redundant protein-coding genes (~15,000), fewer **ribosomal DNA (rDNA)** arrays and dispersed **centromere** repeats [14]. The genome size difference between *W. australiana* and *S. polyrhiza* mainly results from the doubling of **transposable elements (TEs)** in the former (50% vs. 23% respectively; Table 1; Figure 2A). However, the solo to intact ratio (S:I) of TEs in *W. australiana*, which is an indicator of the relative level and timing of TE loss, is one of the highest across plants tested. The higher S:I ratio and TE removal in *W. australiana* could in part explain the fact that this species has the smallest genome in the *Wolffia* genus.

There are several genome features found in all plants such as **telomeres**, **centromeres**, and **rDNA arrays** that are unique in *Wolffia* and could provide an advantage for synthetic biology applications. The *W. australiana* genome has long (≥10 kb) telomere arrays (AAACCCT) at the end of its chromosomes compared to *S. polyrhiza* (<5 kb) [14], which may suggest a method to stabilize chromosomes. In addition, it may provide a clue as to the longevity of *Wolffia* fronds, which go into a stationary state and do not senesce when resources are limited. However, similar to *S. polyrhiza*, *W. australiana* has weak or no centromere arrays consistent with the centromeres being **holocentric**, dispersed, or not centered on repeat arrays [33]. In any case, the potentially small centromere size in duckweeds could facilitate chromosome engineering and may be used to easily break and fuse small chromosomes. Also similar between the *Wolffia*

and *Spirodela* genomes is the small size of their rDNA arrays, which are less than half the size of those in other plants of similar genome sizes [14,30]. In arabidopsis, the centromere arrays are megabases in length, which have only recently been resolved, and the rDNA arrays have yet to be resolved due to their size and sequence conservation between repeats [34].

Wolffia *contains a core non-redundant gene set focused on unrestricted growth*

In addition to a streamlined genome, Wa8730 and Wa7733 have a reduced number of predicted protein coding genes at 14,324 and 15,312 respectively, which is even fewer than spirodela (18,486) and the fewest predicted of any high quality plant genome resolved to date (Table 1) [17]. Despite the decrease in gene number, the *W. australiana* genomes are highly syntenic across the duckweed genera with over 97% in shared syntenic blocks as highlighted by a syntenic block with a tight linkage between core circadian clock genes conserved as far back as the bryophytes (Figure 2A) [35]. A second published Wa8730 genome suggested a higher gene count of 22,293 and proposed the evidence for a recent whole genome duplication (WGD) [36]. However, over 21% (6,890) of the predicted genes from this genome are not syntenic to those in the complete *S. polyrhiza* (sp9509) reference genome, suggesting those conclusions are incorrect due to deficient genome assembly that resulted in residual haplotypes and repeats (final genome size 456 Mb). In terms of gene gains/losses, however, their work largely confirmed our findings for orthogroup changes between *Wolffia* and the more basal genus of *Spirodela*.

The reduced gene number in *W. australiana* is apparently due to gene fractionation over time resulting in a non-redundant gene set with 90% of the orthogroups having less than 10 paralogous copies as compared to 80%, 60% and 50% in other higher plants such as spirodela (*S. polyrhiza*), arabidopsis (*A. thaliana*) and rice (*Oryza sativa*), respectively (Figure 2C) [14]. The reduction in gene redundancy compared to other plants should facilitate gene and pathway engineering. In addition, wolffia has the smallest coding space (exons only; 20 Mb) amongst plant genomes characterized to date (Table 1), which also will reduce the size of the target for synthesis, assembly and engineering. Moreover, the *W. australiana* proteome represents a core set of green lineage genes with over 98% contained in orthogroups (gene families) across a comprehensive survey of monocots, eudicots, bryophytes and algae (Table 1; Figure 2B) [14,37]. Together, these features of wolffia provide a streamlined set of core genes for higher plant synthetic biology.

Several simplified species have been suggested as model systems such as the alga *Chlamydomonas reinhardtii* as well as the bryophytes *Marchantia polymorpha* (liverwort) and *Physcomitrella patens* (moss) (Table 1) [38–40]. These model systems are quite powerful for fundamental biological questions that span the green lineage as well as with tools and resources that make them superior over wolffia for specific questions. However, wolffia represents the gene content and structure of higher plants compared to these models. In contrast to wolffia, marchantia, physcomitrella and chlamydomonas are less reflective of higher plant gene content with only 68%, 65% and 49% respectively of genes in higher plant orthogroups. Moreover, in these models, the families have a high level of redundancy with 57%, 56% and 45% respectively of their genes being found in orthogroups less than 10 paralogs per family (Table 1). The fact that wolffia better represents higher plants and has less redundancy in gene space make it ideal from an engineering standpoint.

As floating macrophytes, the strategy that the Lemnaceae have adopted to compete with other aquatic species, such as the ubiquitous algae, is apparently fast growth. By rapidly covering the surface of ponds and still water bodies, duckweeds can limit the amount of light from reaching other photosynthetic competitors in the lower water column. The morphological reduction and

gene loss thus appear to be coupled to optimize *W. australiana* for continuous unregulated growth [14]. There is a specific reduction or loss of genes associated with plant immunity, architecture, light signalling, and the circadian clock. For instance, *W. australiana* has lost all but one canonical nucleotide-binding, leucine-rich repeat (NLR) immune receptor, which suggests it either leverages a distinct disease resistance strategy or since it doubles so rapidly it may not require a robust immune response or systemic acquired resistance. There is a well known trade-off between growth and immune response in plants that is partially mediated through the plant brassinosteroid phytohormone pathway [41–45]. In addition, since Effector Triggered Immunity often results in programmed cell death, this may not be an optimal way for a minimalist plant to fight pathogens. Conversely, *W. australiana* has expanded specific gene families for wax and sphingolipid production [14], which may provide more of a mechanical protection against pathogen invasion with the former and additional lipid-mediated defense signaling for the latter [46].

Simplification of their body plan and streamlining of its associated developmental process appear to be additional strategies adopted by the more derived *Wolffia* species to boost their growth efficiency through miniaturization. Consistent with the lack of roots, diverse organs, and vasculature, *W. australiana* has lost or reduced architectural genes in the *WUSCHEL (WOX)*, *CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN* (*CASP*), and *CLAVATA3/ESR-RELATED* (*CLE*) families. Interestingly, many of these genes are lost through TE-mediated disruptions as compared to syntenic orthologs in *S. polyrhiza* [14]. Other genes that are apparently lost through TE-mediated disruption in wolffia include light signalling and circadian clock genes, which help to integrate environmental signals to control growth. One interpretation is that architecture and light signalling genes are intimately coupled and the reduction in morphology leads to parallel loss of genes involved in environmental control.

A key role of the circadian clock is to focus specific biological activities in the organism to the correct time of the day under an array of changing environmental conditions [47]. This apparently includes the regulation of plant innate immunity through rhythmic changes in the cellular redox state [48]. *W. australiana* has retained one third of the light signalling and circadian clock genes compared to other model plants, while maintaining a normal number of flowering time genes, and in parallel only 13% of its genes are expressed in a time-of-day (TOD) fashion compared with 50% in most other plants [14,49,50]. Two different mechanisms apparently contribute to this global decrease in TOD control: the loss of many genes such as those in the immune pathway that are normally under this regulation and fewer clock regulators such as those working through the Telobox (TBX) element [14]. The network of genes that are controlled in a TOD fashion through the TBX cis-element are highly conserved and validated as far back as the Lycophtye Isoetes [51]. For the latter, it is interesting to note that the TBX element still remains in the promoters of predicted orthologues in *W. australiana* to TBX regulated genes in other model plants. This could suggest that either the loss of TBX binding factors in *Wolffia* is relatively recent or that the TBX sites in these genes still contribute to their activity in the absence of diurnal regulation. Of the genes that do cycle in a TOD fashion, they are highly focused on genes associated with energy acquisition and carbohydrate metabolism. Therefore, it seems that as *W. australiana* streamlined its morphology and development, it gained uncontrolled/ungated growth due to the elimination of non-essential genes and feedback controls. This makes it an ideal open platform to build back fundamental pathways using defined components.

Wolffia australiana **as an optimal synthetic biology chassis for plants**

The ultimate opportunity for *W. australiana* to contribute to plant research and applications is represented by its features that are desirable for a synthetic biology chassis (Box 1). Its minimal

body plan, reduced non-redundant gene set, manageable genome components, and transformability make it ideal to functionally characterize genes through knock-outs, knock-ins, and ectopic control via regulatable promoters or protein modules. While some tools are currently available to leverage *Wolffia* as a synthetic biology chassis, there are still some areas of opportunity that require creative improvements for *Wolffia* to fully realize its potential (Box 1). For instance, bottom-up engineering requires the delivery of chromosome-sized pieces of DNA, which for building HACs has been accomplished through conjugation [8]. Plants have cell walls that hamper the addition of large pieces of DNA via conjugation, although duckweeds have very low levels of lignin and are very easy to protoplast [14,29,52,53]. Methods to make cell wall free protoplasts are quite mature in plants and coupled to the ability to regenerate plants directly from protoplasts can facilitate both bottom-up as well as top-down approaches [54].

Researchers and industry alike realized the potential of duckweeds as a bioreactor very early on and developed transformation protocols to introduce genes of interest for protein expression. As a result, there are robust methods for duckweed growth and transformation that have been continually updated and refined over the past 30 years. Generally in *Wolffia*, stable and transient transformation protocols are feasible with the former requiring a callus step [55–58]. Direct frond transformation, analogous to floral dipping in *Arabidopsis*, for stable integrations is desired, yet has only been achieved in *Lemna minor* and *Landoltia punctata* [59,60]. Fronds of *W. australiana* and other duckweeds can be readily transformed transiently using the biolistic method, which when coupled to CRISPR-Cas9 genome editing or gene expression systems, could enable the rapid knockout or the expression of genes in transformed cells *in planta* [61,62]. Also, new CRISPR editing based technologies are rapidly developing, such as gene drive systems, which can provide new opportunities for top-down genome engineering [63]. With further optimization of transformation and genetic methodologies in *W. australiana*, the high quality sequence assemblies for its genome should enable a path toward defining the minimal gene set in this plant and help pave the way for designing the genome of a synthetic plant (Box 1).

Since *Wolffia* is missing roots and vasculature, but these tissues exist in the rooted duckweed genera such as *Spirodela, Landoltia,* and *Lemna*, one can imagine building back those pathways using a defined suite of candidate genes in pathways of interest to understand how a root or vasculature are established as examples. In principle, since it is now possible to follow every cell with single cell sequencing technologies [64,65], and *Wolffia* only has ~4,000 cells, every modification at the genome level can be coupled with analyses at molecular resolution. Moreover, putative holocentric chromosomes of *Wolffia* provide an opportunity to build, break, and change the chromosomes, which could help usher in an era of bottom up-genome engineering in plants [9]. Ultimately, since all of the basic plant parts and pathways exist in *Wolffia*, it may be possible to rebuild it to a plant that looks much like *Arabidopsis* through introduction of different suites of genes for various structures and response pathways. This would be a tour-de-force in reductionist approach to ultimately defining the functions and roles for each gene in a plant, starting with a minimal synthetic plant genome as a chassis.

Acknowledgements

We thank Michael and Lam Labs for critical reading of this review. Thanks to Kelly Colt for designing parts of Figure 1. This work was supported by the Tang Genomics Fund to T.P.M. Duckweed research at the Lam laboratory is supported in part by a grant from the Department of Energy (DE-SC0018244), and a Multi-State Capacity project (#NJ12710) from the New Jersey Agricultural Experiment Station at Rutgers University.

REFERENCES:

1 Eisenstein, M. (2020) How to build a genome. *Nature* 578, 633–635
2 Gibson, D.G. *et al.* (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* 329, 52–56
3 Fredens, J. *et al.* (2019) Total synthesis of *Escherichia coli* with a recoded genome. *Nature* 569, 514–518
4 Mitchell, L.A. *et al.* (2017) Synthesis, debugging, and effects of synthetic chromosome consolidation: synVI and beyond. *Science* 355,
5 Dymond, J.S. *et al.* (2011) Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* 477, 471–476
6 Shen, Y. *et al.* (2016) SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. *Genome Res.* 26, 36–49
7 Richardson, S.M. *et al.* (2017) Design of a synthetic yeast genome. *Science* 355, 1040–1044
8 Brown, D.M. and Glass, J.I. (2020) Technology used to build and transfer mammalian chromosomes. *Exp. Cell Res.* 388, 111851
9 Dawe, R.K. and Kelly Dawe, R. Charting the path to fully synthetic plant chromosomes. , *Experimental Cell Research*, 390. (2020) , 111951
10 Palladino, J. *et al.* (2020) Targeted De Novo Centromere Formation in Drosophila Reveals Plasticity and Maintenance Potential of CENP-A Chromatin. *Dev. Cell* 53, 129
11 Mendiburo, M.J. *et al.* (2011) Drosophila CENH3 is sufficient for centromere formation. *Science* 334, 686–690
12 Birchler, J.A. *et al.* (2010) Engineered Minichromosomes in Plants. *CRC Crit. Rev. Plant Sci.* 29, 135–147
13 Hutchison, C.A., 3rd *et al.* (2016) Design and synthesis of a minimal bacterial genome. *Science* 351, aad6253
14 Michael, T.P. *et al.* (2020) Genome and time-of-day transcriptome of *Wolffia australiana* link morphological minimization with gene loss and less growth control. *Genome Res.* DOI: 10.1101/gr.266429.120
15 Hillman, W.S. and Culley, D.D. (1978) The Uses of Duckweed: The rapid growth, nutritional value, and high biomass productivity of these floating plants suggest their use in water treatment, as feed crops, and in energy-efficient farming. *Am. Sci.* 66, 442–451
16 Hillman, W.S. (1976) Calibrating duckweeds: light, clocks, metabolism, flowering. *Science* 193, 453–458
17 Acosta, K. *et al.* (2021) Return of the Lemnaceae: duckweed as a model plant system in the genomics and postgenomics era. *Plant Cell* 33, 3207–3234
18 Lam, E. *et al.* (2014) Duckweed in bloom: the 2nd International Conference on Duckweed Research and Applications heralds the return of a plant model for plant biology. *Plant Mol. Biol.* 84, 737–742
19 Sree, K.S. *et al.* (2015) How fast can angiosperms grow? Species and clonal diversity of growth rates in the genus *Wolffia* (Lemnaceae). *Acta Physiol. Plant* 37, 204
20 Fujita, M. *et al.* (1999) Nutrient removal and starch production through cultivation of *Wolffia arrhiza*. *J. Biosci. Bioeng.* 87, 194–198
21 Landolt, E. (1986) Biosystematic investigations in the family of duckweeds (Lemnaceae)(Vol. 2.) The family of Lemnaceae-a monographic study. vol. 1. *Veroff Geobot. Inst. ETH* 71, 1–563
22 Abramson, B.W. *et al.* 03-Jun-(2021) , Single nuclei transcriptome of the Lesser Duckweed *Lemna minuta* reveals cell trajectories for an entire plant. *bioRxiv*, 2021.06.03.446947

23 Jones, D.H. *et al.* (2020) Preparation, Scanning and Analysis of Duckweed Using X-Ray Computed Microtomography. *Front. Plant Sci.* 11, 617830

24 Maheshwari, S.C. and Chauhan, O.S. (1963) In vitro Control of Flowering in *Wolffia microscopica*. *Nature* 198, 99–100

25 Sree, K.S. *et al.* (2015) The duckweed *Wolffia microscopica*: A unique aquatic monocot. *Flora - Morphology, Distribution, Functional Ecology of Plants*, 210, 31–39

26 Wang, W. *et al.* (2011) Evolution of Genome Size in Duckweeds (Lemnaceae). *J. Bot.* 2011, Article ID 570319, 9 pages, doi:10.1155/2011/570319

27 Hoang, P.T.N. *et al.* (2019) Variation in genome size, cell and nucleus volume, chromosome number and rDNA loci among duckweeds. *Sci. Rep.* 9, 3234

28 Hoang, P.T.N. *et al.* (2020) Chromosome-scale genome assembly for the duckweed *Spirodela intermedia*, integrating cytogenetic maps, PacBio and Oxford Nanopore libraries. *Sci. Rep.* 10, 19230

29 Wang, W. *et al.* (2014) The *Spirodela polyrhiza* genome reveals insights into its neotenous reduction fast growth and aquatic lifestyle. *Nat. Commun.* 5, 3311

30 Michael, T.P. *et al.* (2017) Comprehensive definition of genome features in *Spirodela polyrhiza* by high-depth physical mapping and short-read DNA sequencing strategies. *Plant J.* 89, 617–635

31 Harkess, A. *et al.* (2021) Improved *Spirodela polyrhiza* genome and proteomic analyses reveal a conserved chromosomal structure with high abundance of chloroplastic proteins favoring energy production. *J. Exp. Bot.* 72, 2491–2500

32 Van Hoeck, A. *et al.* (2015) The first draft genome of the aquatic model plant *Lemna minor* opens the route for future stress physiology research and biotechnological applications. *Biotechnol. Biofuels* 8, 188

33 Mandrioli, M. and Manicardi, G.C. (2020) Holocentric chromosomes. *PLoS Genet.* 16, e1008918

34 Naish, M. *et al.* 31-May-(2021) , The genetic and epigenetic landscape of the Arabidopsis centromeres. *bioRxiv*, 2021.05.30.446350

35 Michael, T.P. Core circadian clock and light signaling genes brought into genetic linkage across the green lineage. *bioRxiv*, 02 Nov 2021, DOI: 10.1101/2021.11.02.466975

36 Park, H. *et al.* Genome of the world's smallest flowering plant, *Wolffia australiana*, helps explain its specialized physiology and unique morphology. *Communications Biology*, 4. (2021)

37 Van Bel, M. *et al.* (2018) PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res.* 46, D1190–D1196

38 Naramoto, S. *et al.* The bryophytes Physcomitrium patens and *Marchantia polymorpha* as model systems for studying evolutionary cell and developmental biology in plants. *The Plant Cell*. (2021)

39 Salomé, P.A. and Merchant, S.S. (2019) A Series of Fortunate Events: Introducing Chlamydomonas as a Reference Organism. *Plant Cell* 31, 1682–1707

40 Rensing, S.A. *et al.* (2020) The Moss *Physcomitrium* (*Physcomitrella*) *patens*: A Model Organism for Non-Seed Plants. *Plant Cell* 32, 1361–1376

41 Karasov, T.L. *et al.* (2017) Mechanisms to Mitigate the Trade-Off between Growth and Defense. *Plant Cell* 29, 666–680

42 Chae, E. *et al.* (2016) Cooperation and Conflict in the Plant Immune System. *PLoS Pathog.* 12, e1005452

43 Neuser, J. *et al.* (2019) HBI1 Mediates the Trade-off between Growth and Immunity through Its Impact on Apoplastic ROS Homeostasis. *Cell Rep.* 28, 1670–1678.e3

44 Wang, J. *et al.* (2021) Understanding the molecular mechanisms of trade-offs between plant growth and immunity. *Sci. China Life Sci.* 64, 234–241

45 Lozano-Durán, R. *et al.* (2013) The transcriptional regulator BZR1 mediates trade-off between plant innate immunity and growth. *Elife* 2, e00983
46 Huby, E. *et al.* (2020) Sphingolipids: towards an integrated view of metabolism during the plant stress response. *New Phytol.* 225, 659–670
47 Michael, T.P. *et al.* (2003) Enhanced fitness conferred by naturally occurring variation in the circadian clock. *Science* 302, 1049–1053
48 Zhou, M. *et al.* (2015) Redox rhythm reinforces the circadian clock to gate immune response. *Nature* 523, 472–476
49 Filichkin, S.A. *et al.* (2011) Global profiling of rice and poplar transcriptomes highlights key conserved circadian-controlled pathways and cis-regulatory modules. *PLoS One* 6, e16907
50 Michael, T.P. *et al.* (2008) Network Discovery Pipeline Elucidates Conserved Time-of-Day–Specific cis-Regulatory Modules. *PLoS Genet.* 4, e14
51 Wickell, D. *et al.* 10-Jun-(2021) , Underwater CAM photosynthesis elucidated by Isoetes genome. *bioRxiv*, 2021.06.09.447806
52 Pagliuso, D. *et al.* (2020) High Saccharification, Low Lignin, and High Sustainability Potential Make Duckweeds Adequate as Bioenergy Feedstocks. *Bioenergy Res.* DOI: 10.1007/s12155-020-10211-x
53 Song, L.-Y. *et al.* (2020) Establishment and application of Transient expression system of protoplast in duckweed. *Xi Nan Nong Ye Xue Bao* 33, 1792–1798
54 Jeong, Y.Y. *et al.* (2021) Optimization of protoplast regeneration in the model plant *Arabidopsis thaliana*. *Plant Methods* 17, 21
55 Khvatkov, P. *et al.* (2021) *Wolffia arrhiza* as a promising producer of recombinant hirudin. *3 Biotech* 11, 209
56 Heenatigala, P.P.M. *et al.* (2020) Expression of LamB Vaccine Antigen in *Wolffia globosa* (Duck Weed) Against Fish Vibriosis. *Front. Immunol.* 11, 1857
57 Heenatigala, P.P.M. *et al.* (2018) Development of Efficient Protocols for Stable and Transient Gene Transformation for *Wolffia Globosa* Using Agrobacterium. *Front Chem* 6, 227
58 Khvatkov, P. *et al.* (2018) Development of *Wolffia arrhiza* as a Producer for Recombinant Human Granulocyte Colony-Stimulating Factor. *Front Chem* 6, 304
59 Yang, G.-L. *et al.* (2018) Frond transformation system mediated by *Agrobacterium tumefaciens* for *Lemna minor*. *Plant Mol. Biol.* 98, 319–331
60 Balaji, P. *et al.* (2016) Expression of anti-tumor necrosis factor alpha (TNFα) single-chain variable fragment (scFv) in *Spirodela punctata* plants transformed with *Agrobacterium tumefaciens*. *Biotechnol. Appl. Biochem.* 63, 354–361
61 Reinard, T. *et al.* (2020) Editing the Genome of *Wolffia australiana*. In *The Duckweed Genomes* (Cao, X. H. et al., eds), pp. 165–177, Springer International Publishing
62 Muranaka, T. *et al.* (2015) Characterisation of circadian rhythms of various duckweeds. *Plant Biol.* 17 Suppl 1, 66–74
63 Zhang, T. *et al.* (2021) Selective inheritance of target genes from only one parent of sexually reproduced F1 progeny in Arabidopsis. *Nat. Commun.* 12, 3854
64 Rich-Griffin, C. *et al.* (2020) Single-Cell Transcriptomics: A High-Resolution Avenue for Plant Functional Genomics. *Trends Plant Sci.* 25, 186–197
65 Rhee, S.Y. *et al.* (2019) Towards Building a Plant Cell Atlas. *Trends Plant Sci.* 24, 303–310

Table 1. *Wolffia* genome features compared to other model plants[a]

	Wolffia	Spirodela	Rice	Arabidopsis	Liverwort	Moss	Algae
Genome size (Mb)	375	158	364	135	225	518	111
Chromosome number (1n)	20	20	12	5	9	27	17
Gene (#)	14,508	18,486	52,424	35,386	19,138	32,926	17,741
Gene space: intron + exon (Mb)	64	73	121	64	71	102	94
Coding space: exon only (Mb)	**20**	**28**	58	33	23	36	39
Repeat content (%)	50	23	35	21	22	57	14
Genes in orthogroups (%)	**98**	**96**	78	88	68	65	49
Families with less than 10 paralogs (%)	**90**	**85**	57	66	57	56	45

[a]Genome stats for wolffia (*W. australiana*, Wa8730 clone), spirodela (*S. polyrhiza*, Sp9509 clone), rice (*Oryza sativa* var japonica), arabidopsis (*A. thaliana*, Columbia accession), liverwort (*Marchantia polymorpha*), moss (*Physcomitrella patens*) and algae (*Chlamydomonas reinhardtii*). The key metrics of duckweed genomes for this article are highlighted in **blue**.

Box1. *Wolffia australiana* as an optimal synthetic biology chassis.

Features in the organism of interest: About half of the current genome is repeats associated with TEs (transposable elements) and likely are dispensable; lowest paralog expansion amongst model plants examined to date = low functional redundancy expected; low structural complexity and elaboration of organ types; simple development optimized for fast growth

Minimal genome that is manipulatable: At present estimated at ~80 Mb (15 Mb promoter space + 65 Mb gene space) for all annotated protein encoding genes, which is about 7 times that of the yeast synthetic genome (12 Mb). With only the CDS (coding space), this goes down to 35 Mb, which is about 3 to 4 times the size of the yeast synthetic genome (Table 1). Further functional deletion studies using either telomere truncation or recombinase-mediated deletion strategies could potentially decrease this size further by identifying non-essential genes.

Core non-redundant genes: The majority (98%) of genes are in orthogroups conserved across a wide range of plant types, with only a small percentage in groups/families containing 10 or more members (Figure 2). This indicates there is likely minimal functional redundancy within an orthogroup, which would significantly facilitate gene function identification studies.

Easy to grow and manipulate: *Wolffia's* aquatic habitat coupled with small size enables flexible experimental design with economy in costs, time, and space. These characteristics also makes *Wolffia* extremely user-friendly in terms of high-throughput approaches for phenotyping using automated liquid handling methodologies.

Transformation/conjugation systems: Stable transformation of *Wolffia* species using Agrobacterium-based DNA delivery with plant calli have been demonstrated, as well as transient transfection using physical methods for plasmid introduction into intact plants. With the ability to produce protoplasts from *Wolffia*, conjugation systems via protoplast fusion should be feasible for the mobilization of synthetic chromosomes from other sources such as spheroplasts of yeast cells.

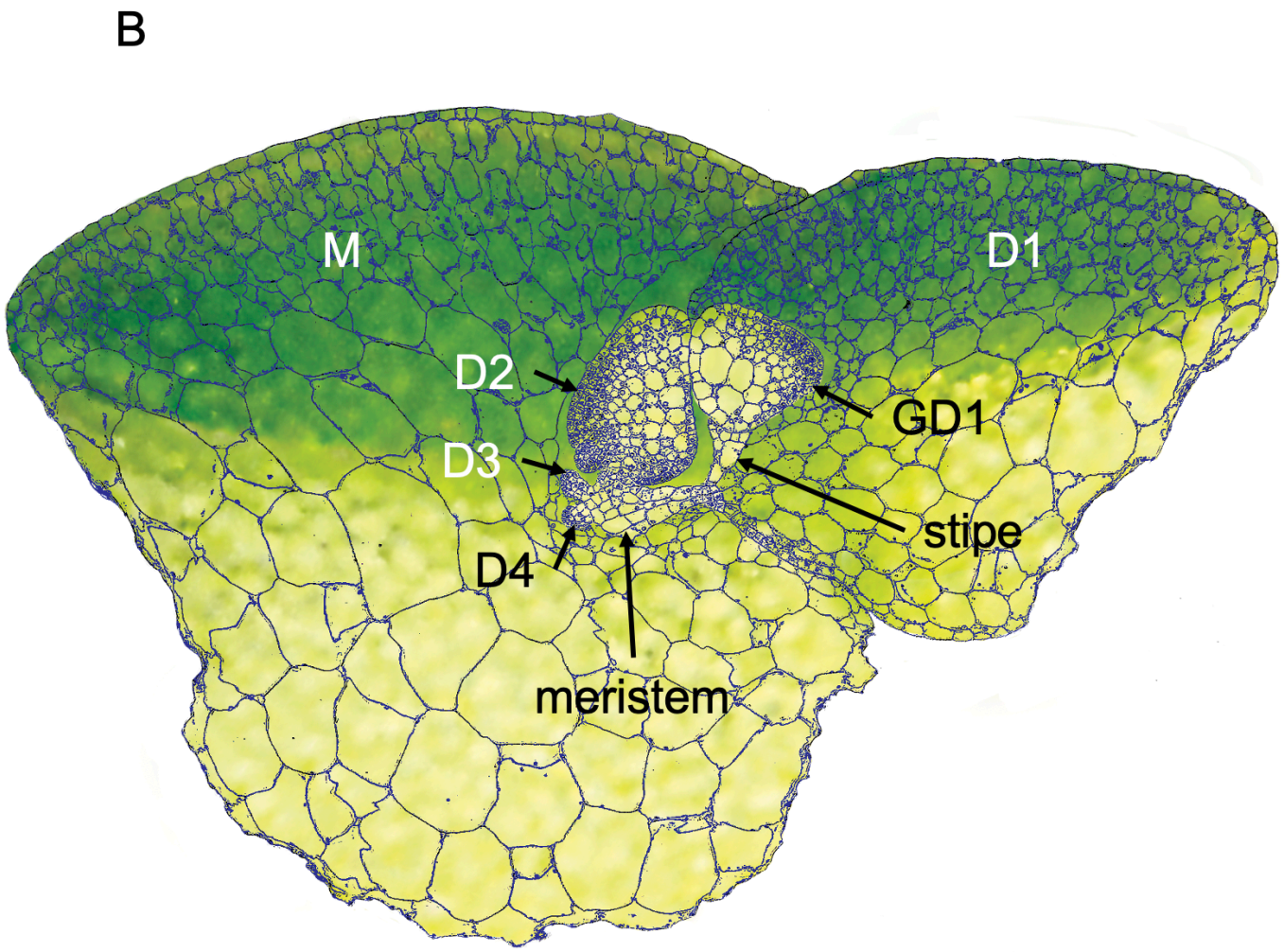
Simple anatomy provides a "Blank Canvas": The minimal cell types and body plan of *Wolffia* provide a basic foundation to build and test various assemblages of genes to query predicted functions for each component as well as the outputs and feedbacks in the pathway. For example, what is the minimum number of genes that would be required to rebuild a root in *Wolffia* from genes of rooted duckweeds such as *Spirodela* or *Lemna*?

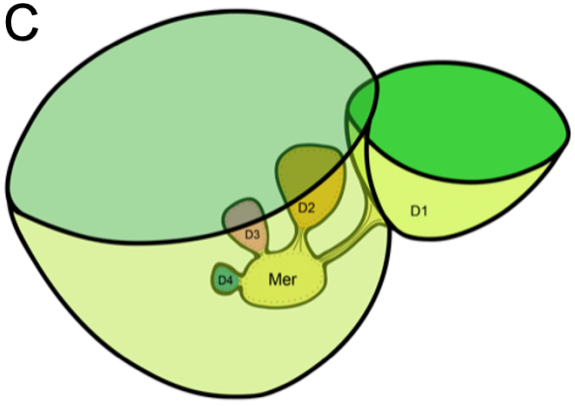
FIGURE LEGENDS:

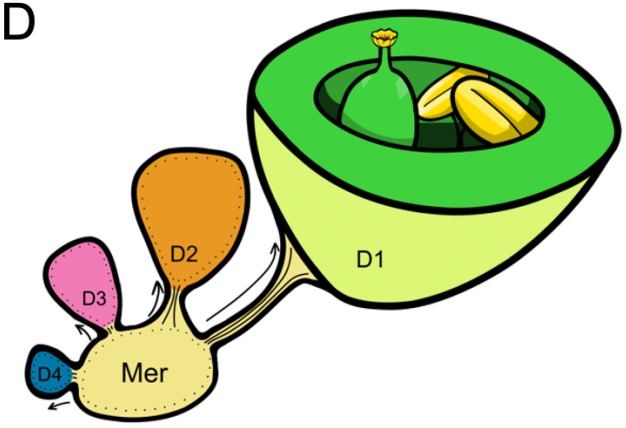
Figure 1. *Wolffia australiana* is essentially a ball of cells with no roots and few cell types. **(A)** *W. australiana* growing in a test tube and compared to a pencil for size reference. **(B)** Cross-section of *W. australiana* showing the mother frond (M; left) and the emerging daughter frond (D; right). The top cells are darker green due to their smaller size and higher number of chloroplasts. The bottom cells are larger with fewer chloroplasts and provide buoyancy. Cellular view of *W. australiana* highlighting the nested daughter (D) and granddaughter (GD) fronds that are developing from the meristematic tissue of the mother frond (M). The numbers (1,2,3,4) represent the successive generations of D and GD fronds. **(C)** Cartoon of the nested M and D fonds; and then **(D)** a close up of the meristem (mer) and how D fronds (D1-D4) emerge connected to an internode-like structure called a stipe (panel B; also arrows next to the stipe in panel D indicate direction of growth/extension). *Wolffia* also makes a flower structure including a stigma and stamen in a cavity on the top surface of the frond; and **(E)** if *Wolffia* fronds were to stay attached via the internode (stipe), the result may look much like the common plant architecture of a rosette that is found in a 15 day old arabidopsis plant **(F)**. **Photo credits:** panel A - Ryan Gutierrez and Philomena Chu; panel B - Ljudmilla Borisjuk; Drawing and photo composite by Kelly Colt.

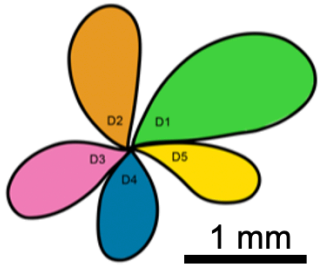
Figure 2. *Wolffia* represents a conserved core set of non-redundant eudicot genes. (A) A synthetic block conserved back to bryophytes across all three duckweed genera containing *PSEUDO RESPONSE REGULATOR 5* (*PRR5*; blue) and *LATE ELONGATED HYPOCOTYL* (*LHY*; red). Genera: *Spirodela* (sp9509), *Lemna* (lm5633), and *Wolffia* (wa8730). Forward genes (blue), reverse genes (green), and repeats (orange). **(B)** *Wolffia* (green) genes are found in 98% of the core eudicot families, compared to other species such as *Spirodela* (orange; 93%), *Arabidopsis* (brown; 88%) and *Oryza sativa* (rice, red, 78%). **(C)** *Wolffia* (Wa, green) has fewer genes (10%) in large families (greater than 10 paralogues) compared to *Spirodela* (*Sp*, orange), *Arabidopsis* (*At*, brown) and *O. sativa* (*Os*, red).

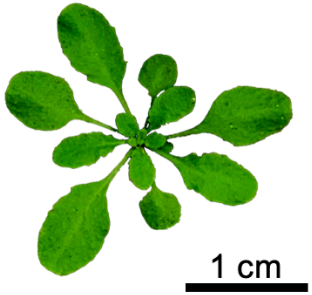

A
1388-3/H
HARD

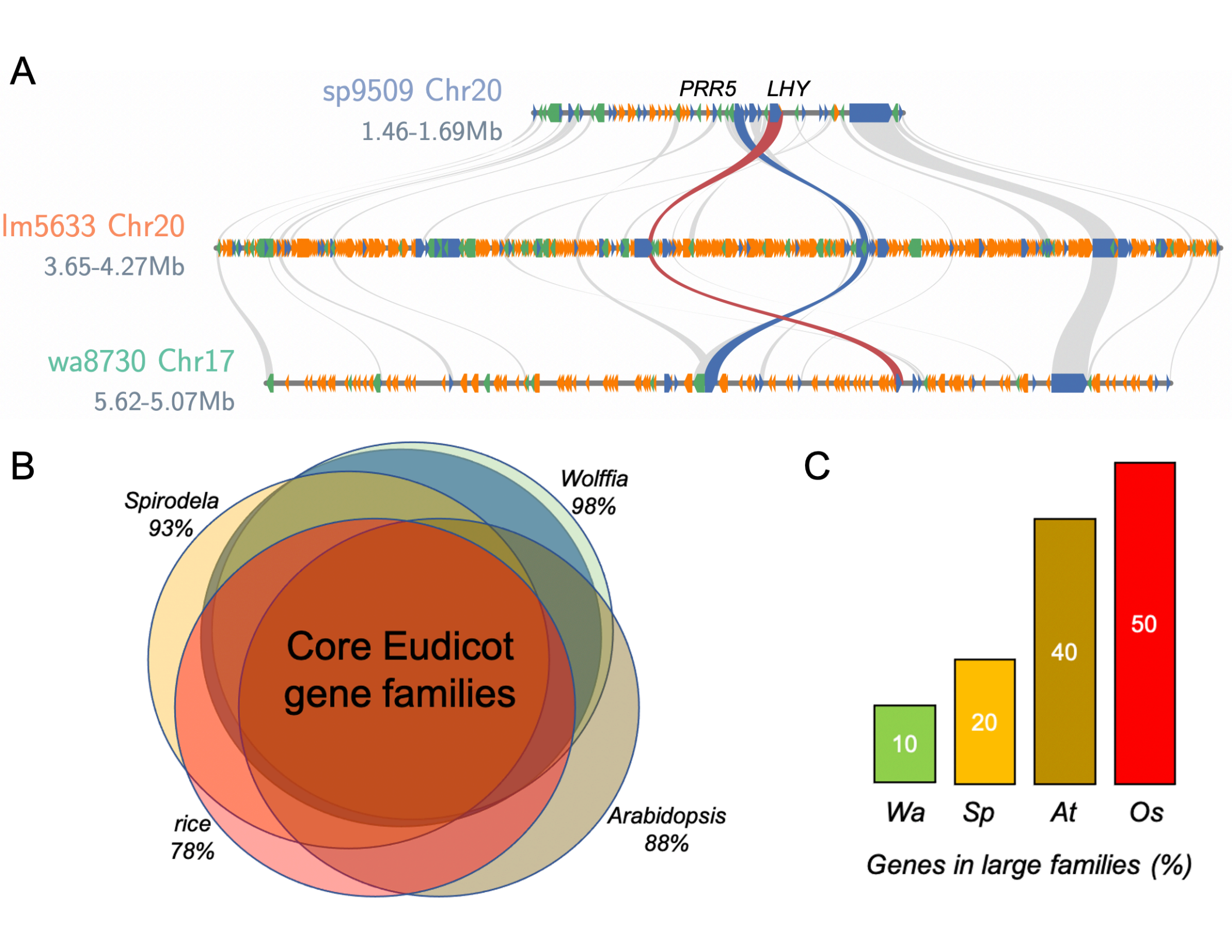

B
M
D1
D2
GD1
D3
stipe
D4
meristem


C
D1
D2
D3
D4
Mer


D
D1
D2
D3
D4
Mer


E
D1
D2
D3
D4
D5
1 mm


F
1 cm


A
sp9509 Chr20
1.46-1.69Mb
PRR5
LHY
lm5633 Chr20
3.65-4.27Mb
wa8730 Chr17
5.62-5.07Mb
B
Spirodela
93%
Wolffia
98%
Core Eudicot
gene families
rice
78%
Arabidopsis
88%
C
10
20
40
50
Wa
Sp
At
Os
Genes in large families (%)