Ploidy and domestication are associated with genome size variation in Palms

by BEE F. GUNN, THIERRY BEULÉ, LUC BAUDOUIN, PASCAL ILBERT, CHRISTOPHE DUPERRAY, MICHAEL CRISP, AUGUSTE ISSALI, JL KONAN AND ALAIN RIVAL

which was submitted as “Genome size variation in Attaleinae (Arecaceae) with emphasis on coconuts (Cocos nucifera L.)” Manuscript received 12 April 2015; revision accepted 24 August 2015, currently available online on the site of *The American Journal of Botany*

http://www.amjbot.org/content/early/2015/10/02/ajb.1500164.full.pdf+html

Final title: Ploidy and domestication are associated with genome size variation in Palms

GENOME SIZE VARIATION IN ATTALEINAE (ARECACEAE) WITH EMPHASIS ON COCONUTS (*COCOS NUCIFERA* L.) [1]

BEE F. GUNN[2,7], LUC BAUDOUIN[3], THIERRY BEULÉ[4], PASCAL ILBERT[4], CHRISTOPHE DUPERRAY[5], MICHAEL CRISP[2], AUGUSTE ISSALI[6], JEAN-LOUIS KONAN[6] AND ALAIN RIVAL[4]

[2] Research School of Biology, Division of Evolution, Ecology and Genetics, The Australian National University, Canberra, ACT 0200, Australia

[3]CIRAD, UMR AGAP, 34398 Montpellier, France

[4] CIRAD, UMR DIADE, 34398 Montpellier, France

[5]Montpellier Rio Imaging, IRB, INSERM U1040, CHU de Montpellier, 34295 Montpellier, France.

[6] Station de Recherche Marc Delorme, Port Bouët. % 3 $ E L G M D Q & { W H G ¶ , Y R

- *Premise of the study:* Genome size is a major biological parameter that is correlated with several biological traits and affected by intense selection pressures such as domestication. Genome size variations among related species of palms is of evolutionary significance and further knowledge of genome size will provide crucial information needed for planning of whole genome sequencing and accurate annotations. In addition, large genomes tend to contain more repeated sequences, which makes assembly more difficult. In this paper, we studied the genome size of *Cocos nucifera* L. and its variations then we compared it to the values estimated for related palms of the Attaleinae subtribe.
- *Methods* We used flow cytometric analysis of isolated nuclei from young palm leaf material to estimate genome sizes of 23 coconut cultivars (Talls, Dwarfs and hybrids) worldwide and 17 palm species from Attaleinae. Ancestral genome size reconstruction was based on maximum likelihood phylogeny of Attaleinae from seven *WRKY* loci.
- *Key Results* The coconut genome is large and shows intraspecific variation associated with domestication. Variation among Tall coconuts was highly significantly higher than amongst Dwarfs. Comparison of Attaleinae genomes showed moderate variation across genera, except for *Jubaeopsis caffra*, *Voanioala gerardii, Beccariophoenix alfredii* and *Allagoptera caudescens* for which polyploidy led to increased genome sizes.
- *Conclusions* Our results contribute to understanding of the relationship between domestication and genome size in long-lived tree crops and they provide important information for implementation of whole genome

sequencing of the coconut and other domesticated plants. Polyploidy evolved independently in two clades within Attaleinae.

- **Key words:** Attaleinae; C-value; Cocos nucifera; domestication; flow cytometry, evolution; nuclear DNA content; polyploidy; minimum generation time; holoploid

INTRODUCTION

Polyploidy is an important process in the evolution of plants with far reaching effects from molecular to ecological levels and it contributes to reproductive isolation, as novel gene expressions led to divergence and potentially to speciation (Adams and Wendel, 2005; Comai, 2005). Polyploidy is known to occur among 80% of angiosperms (Masterton, 1994) and it is also common in domesticated plants. Indeed it is detectable in major crops such as cereals (wheat and rye), maize, cotton, potato, banana, sugar cane and coffee (Gaut and Doebley, 1997; Wendel and Cronn, 2003; Heslop-Harrison and Schwarzacher, 2007). More, polyploidy adds complexity when identifying the wild ancestors of a domesticated plant (Olsen and Wendel, 2013). Understanding the impacts of ploidy levels on the genome size is informative since gene duplications can play an important role in epigenetic gene silencing or expression and also provide protection against harmful viruses and transposons (Pichersky, 1990).

Detection of ploidy levels using flow cytometric methods provides a practical tool for plant breeders interested in polyploidy because they may be exploited for desirable phenotypic traits for horticultural purposes (Parris et al., 2010) or for plant conservation biologists as polyploidy may also be a hindrance to reproduction because of sterility of polyploids.

The C-value is equivalent to genome size in diploid species although it is always greater than the genome size(s) in polyploids (Bennett, Bhandol, and Leitch, 2000). Indeed, a diploid plant has two genomes, after gametic fertilization, whereas a polyploid has more than two genomes as a result of either autopolyploidization or allopolyploidization following hybridization (Stebbins, 1959). The C-value (holoploid genome size) of a species corresponds to the DNA amount in its unreplicated haploid

or gametic nucleus (pollen or sperm), regardless of its ploidy level (Swift, 1950; Greilhuber et al., 2005) and it is measured in picograms (pg) or base pairs (bp).

The genome size of a species has major effects on the growth, meiotic and mitotic cycles and on the expansion of cells. Cellular DNA content or nucleotypic changes therefore affect W K H L Q G L Y L G X D O ¶ V P R U S K R O R J L F D O D Q G S K development as well as adaptations to its environment (Price and Baranova, 1976; Bennett, 1998; Hardie and Hebert, 2003; Knight, Molinari, and Petrov, 2005). Large variation in C-values may have consequences or costs to the organisms and several studies have shown that C-values are often associated with ecological constraints in plants (Bennett, 1987; Knight, Molinari, and Petrov, 2005), temporal shifts in phenology such as the early flowering of *Fritillaria* sp. (2C = 96.5 ±254.8) (Grime and Mowforth, 1982) or sensitivity to ionizing radiations and climatic changes in plants and possibly also in animals (Sparrow and Miksche, 1961; Sparrow and Sparrow, 1965; Sparrow, Schwemmer, and Bottino, 1971).

Chromosome numbers (2n), C-values and ploidy levels are tightly linked and remain constant for most species; nevertheless, there are exceptions for which variations do occur. Intraspecific variation in C-values is not rare in plants despite the absence of change in chromosome number of the species; for example, domesticated crops such as *Zea mays* (2n = 20) show 37% variation among various cultivar lines (Laurie and Bennett, 1985) and *Poa annua* (2n = 28) showed a 100% variation rate (Grime, 1983). The switchgrass, *Panicum virgatum* L. is a North American native perennial cultivated for pastures, rangelands and fuel biomass. Cytological studies reveal that it presents a series of karyotypes ranging from diploid (2n = 18) to dodecaploid (2n = 12C = 108) (Church, 1940; Riley and Vogel, 1982).

In Angiosperms, C-values range from 0.1 to 127.4 pg (Bennett, Bhandol, and Leitch, 2000), each value being characteristic of a given species. The palm family (Arecaceae) is among the most diverse, with C-values ranging from 0.9 to 30 pg (Angiosperm 1C-values database (http://data.kew.org/cvalues/)). Within the Cocoseae tribe *Voanioala gerardii* J. Dransf., a polyploid (1C-value = 30 pg; n = ca. 300) shows the highest C-value.

Cocos nucifera L. (Arecaceae) has 16 chromosomes (Nambiar and Swaminathan, 1960; Abraham and Mathew, 1963) and is the only species of its genus. The coconut palm is cultivated globally on over 12 million hectares in the humid tropics. It is best regarded as a semi-domesticated species, a complex of local populations with all degrees of dependency upon humans, from nil to complete (Sauer, 1971). Although Harries (1978) G L V W L Q J X L V K H V ³ G R P H V W L F D W H G ´ D Q G this distinction refers to an ancient domestication event but acknowledges that both types are indifferently cultivated nowadays. Wild populations do exist but only in a few locations (Foale, 2005) but some of them might be feral *i.e.* formerly cultivated population surviving spontaneously (Baudouin, Gunn, and Olsen, 2014).

At the other end of the range, Dwarf coconut can be regarded as the most completely domesticated type (Gunn, Baudouin, and Olsen, 2011). This coconut type is usually grown near human habitations and account for only 5% of coconuts globally (Bourdeix et al., 2010). Its self-pollinating floral biology enables the true to type propagation of desirable genotypes and the screening for rare off-types based on recognizable phenotypic markers such as fruit color and shape. It is precocious, maturing usually after four years. Dwarf coconut is especially appreciated for the water of its immature nuts and its slow growth makes harvesting relatively easy for most of its relatively short lifespan (ca. 35 years) (Bourdeix et al., 2010). Finally, it is

dependent on human protection because it is a poor competitor in natural stands or in mixed plantings due to its short lifespan and to its reduced vigor.

The Tall coconut Z K L F K L V P R U H I U H T X H Q W O D F N P R V W R I W features found in Dwarf coconut. It is predominantly cross-pollinated and thus highly heterozygous. Tall coconuts are fast growing; they become reproductively mature later, usually after seven years and they live for 70 years or more (Bourdeix et al., 2010). Besides Talls and Dwarfs, relatively rare types are observed, among them Semi-Talls, which are self-pollinating like Dwarfs but relatively more robust. The FRPSDFW'ZDUIUHSUHVHQWHGEWKH*Leka* Dwarf from the South Pacific is not related to the other Dwarfs. It is cross-pollinating, and as vigorous as a Tall and owes its small size to a marked reduction in internode length and in the distance between leaflets (Lebrun et al., 2005) .

To date, genome size has been estimated for only 3% of total palm species, principally based on Feulgen-microdensitometry methods (Greilhuber, 1986; Röser, Johnson, and Hanson, 1997). Flow cytometry has become the predominant method for ploidy studies and determination of absolute DNA contents of cells, due to its high sample throughput and relative ease of sample preparation (Dolezel and Bartos, 2005; Dolezel, Greilhuber, and Suda, 2007). Intraspecific genome size has been shown to vary between cultivars and wild progenitors in Angiosperms (Greilhuber, 2005), and such subtle changes may be detected only when using flow cytometry. Karyotyping analyses does not allow for the detection of infraspecific genome size differences because the number of chromosomes is unlikely to vary and when Feulgen-microdensitometry method is used, the presence of tannins in root tissue may interfere with the Feulgen dye then causing errors in the measurement of nuclear DNA amounts (Greilhuber, 1986).

Determination of the genome sizes of cultivated coconuts and ploidy level are essential prerequisites for the sequencing of the coconut genome. This will provide precise calculation for the optimal depth of reads required and accurate assembly and annotations of the coconut genome. Genome sequences have been recently generated and made publicly available for two palm species of major economic importance, namely the date palm (Al-Dous et al., 2011) and the oil palm (Singh et al., 2013). For the coconut palm, future genome sequencing will be of paramount interest for the identification of genes responsible for disease resistance and characters of agro-ecological interest such as drought or salt tolerance (Fan et al., 2013). The integration of gene discovery and Marker Assisted Breeding will pave the way for the generation of new coconut cultivars, which will be better adapted to changing agro-climatic conditions.

We are keen to know if the phenotypic differences such as dwarfism and fruit morphology observed between Dwarf and Tall cultivars and their different generation times (three vs seven years) are related to their genome size. In this study, we explored genome size variation using flow cytometry in 23 coconut genotypes from around the globe, including two Australian wild-sown coconuts. Our objectives were: 1) to determine the actual genome size of coconut, for which contradictory values were published; 2) to study possible intraspecific variations, and the impact of domestication on genome size; 3) to test whether genome size is less variable in Dwarf than in Tall coconut types and 4) to reconstruct ancestral genome sizes across the Attaleinae subtribe.

MATERIALS AND METHODS

Plant Material ² We sampled immature leaves from 23 adult palms originating from 23 coconut populations chosen to cover the genetic diversity of the genus (Appendix 1). Two of them were self-sown, putatively wild, populations from Australia (Mission Beach, lat. -17.869121°, long. 146.106338° and Lizard Island, lat. -14.667717°, long. 145.446729°). The other coconut types under study were traditional and advanced cultivars from the collection preserved at Marc Delorme Research Station (CNRA, Côte d'Ivoire). They include seven self-pollinating Dwarf cultivars, 15 cross- pollinating Tall cultivars, one cross-pollinating "compact Dwarf" cultivars and three population hybrids (one Tall × Tall and two Dwarf × Tall).

Fresh leaf material was collected from the unopened spear leaf of the palm whenever possible. In addition, we sampled leaf material for 17 species across 8 genera of the Cocoseae: *Attalea*, *Beccariophoenix*, *Butia*, *Elaies*, *Jubaeopsis*, *Lytocaryum*, *Allagoptera* and *Sygarus* from the living collections of the Royal Botanic Gardens in Sydney, Australia. We obtained genome size values for additional four species from the Angiosperm 1C-values database. We wrapped approx. 4 cm length of each leaf in moistened tissue paper and placed it into an envelope kept at 4°C to preserve it during transportation to the IRB laboratory in Montpellier, France.

Estimation of 2C-value ² To determine genome size, we first used razor blades to chop coconut and *Petunia hybrida* E. Vilm. leaves in order to extract nuclei. The *P. hybrida* Px PC6 (Vilmorin), 2C = 2.85pg was grown in the greenhouse and used as calibration standard following Coba de la Peña and Brown (2001). Approximately 1 cm^2 of fresh leaves were chopped in 500 µL of ice-cold Otto I buffer (Dolezel, Binarova, and Lucretti, 1989) with the following modifications: no spermine was added and we replaced β-mercaptoethanol with 10 mM sodium metabisulphite which was added immediately before use (Rival et al., 1997). The

lysate was then filtered through disposable filters using 20 µm nylon mesh (Partec CellTrics®) in order to isolate nuclei from cell debris and aggregates. Then 500 L of the filtrate were pipetted into a new disposable tube and 20 L of DAPI ¶-diamidino-2-phenylindole, dihydrochloride) fluorochrome solution (0.1mg mL^{-1}) were added, for a final DAPI concentration of 4 µg mL^{-1}. After homogenizing and stabilizing for 5 minutes at room temperature, the stained nuclei suspensions were analyzed.

We measured relative fluorescence intensities from stained nuclei using a Beckman-Coulter CyAN™ ADP flow cytometer (Beckman Coulter Inc., U.S.A.) with at least 500 nuclei analyzed per run. We repeated measurements of the G1 peaks (non-replicated phase of the cell cycle) for each coconut cultivar 3-5 times with internal standards and used the means (µ ± s.d.) in our assessment of the absolute YDOXHRIWKHFRFRQXW¶JHQRPHVL]HLHOGLQJJUDSKLFDORXWSXWVDV illustrated in Figure 1.

Data Analysis ² The first step of data analysis consisted in a visual examination of the cytometer plots (Fig. 1) in order to exclude unreliable runs (i.e. with low signal to noise ratio, mainly due to inadequate preservation of analyzed plant material).

Proportionality of G1 peak values with internal standard ² The proportionality of the G1 peak values between the coconut genotypes and the internal standard (*Petunia hybrida*) was checked through regression analyses in order to determine the correlation between the G1 peak values of the internal standard and studied coconut genotypes. The results from the regression analysis of G1 peak values for various coconuts against the internal standard (*Petunia hybrida*) were highly correlated (corrected $R^2 = 0.9997$ when the intercept was fixed to 0) thus confirming

their proportionality. The proportionality coefficient was 2.0921 ± 0.0041 (mean ± s.e.). This enabled the use of the ratio of the coconut G1 values to the internal standard to calculate the absolute genome size of the coconut ecotypes (see Appendix 1).

Genome size for each sample was estimated as $G_C = D_C/D_S*G_S$ where D_C is the G1 peak value of coconut, D_S is the G1 peak value of the standard, and G_S is the genome size of the standard (2.85 pg for *Petunia*). We examined variation in genome size among cultivars using ANOVA and we applied the F-test to determine the significance of the values. We tested for possible effects of domestication on genome size of *Cocos nucifera* by forming two groups: Tall (n = 16), and Dwarf (n = 7) again using ANOVA. We followed the same method to analyze variation between Indo-Atlantic and Pacific groups of geographical origin. Finally, we used boxplots to visualize changes in DNA amounts in Dwarf and Tall coconuts. Calculations and graphical representation were carried out using R software (Chambers et al., 1983; R Development CoreTeam, 2011).

Ploidy level ***²*** Ploidy in flow cytometric assays equates a constant DNA quantity (C-value) of the complete chromosome complement with respect to a published reference standard of known ploidy. We determined the ploidy level of the coconut from the positions of the G1 peaks in cytometry histograms. The presence of polyploidy is reflected in the position of the dominant G1 peak and the appearance of more than one non-reference dominant peak in a single sample apart from the internal standard.

Evolution of 2C value in Attaleinae ***²*** We estimated the absolute genome size of the 17 species using flow cytometry (Appendix 2). In order to design the evolutionary tree of the *Attaleineae*, we used seven *WRKY* nuclear loci from Meerow

et al. (2009), concatenated to sequence length of 5.648 kb for 56 taxa across the Attaleinae available from Genbank. We conducted maximum likelihood analyses using PHYML software (Guindon and Gasceul, 2003) implemented through Geneious 6.1.7 (Biomatters Dev. Team 2013) with the following criteria: initial BioNJ tree, NNI topology search, GTR substitution model, discrete Gamma model, 4 categories, random seed and 100 bootstrap replicates.

We applied the maximum likelihood approach as described in Pagel (1999) for ancestral character reconstruction as implemented in the Mesquite software. The maximum likelihood trees (100) were imported into Mesquite Version 2.5 (Maddison and Maddison, 2008) and a character matrix of 2C-values for 18 taxa were appended to the DNA sequences. We traced the 2C-values sizes as continuous characters on to the ML tree in order to infer ancestral state likelihoods. We used *Bactris* and *Elaeis* as outgroups for the non-spiny Attaleinae.

RESULTS

Absolute genome size of the coconut

The overall mean of genome size was 5.963 pg, after exclusion of the hybrid genotypes. The residual standard deviation was 0.0641 pg. This represents the uncertainty due to the breadth of the peaks and to random fluctuations of the experimental conditions.

Ploidy level of coconut cultivars

The DNA histograms obtained for all the coconut cultivars under study clearly showed a single G1 peak, suggesting that all samples were only diploids (Fig. 1). G1 peaks occurred in the same position relative to the internal standard in all cases. Since the *Petunia hybrida* standard used has nearly half the DNA quantity of the coconuts, it is possible that if haploid cells were present in the coconut samples, their peaks may

have overlapped with the standard but leaf cells are somatic and do not undergo meiosis. Nevertheless, the possible presence of spontaneous haploids was checked in several samples without internal standards and it proved constantly negative.

Variation of genome size in coconut [2] We performed an analysis of variance (ANOVA) based on 16 Tall and 7 Dwarf coconut types (Table 1). On average, Tall and Dwarf coconuts differed in genome size ($F = 10.90$, $df = 1$, P value = 0.00163). There were also significant differences among Talls ($F = 10.45$, $df = 15$, P value = $2.68\ 10^{-11}$) but the studied Dwarfs were not significantly different ($F = 1.34$, $df = 6$, P value = 0.257). The estimated mean and confidence interval ($\alpha = 0.05$) of genome size were 6.00 [5.97 ±6.03] and 5.95 [5.74 ±6.16] in Dwarfs and Talls respectively. This takes into account both empirical errors and the estimated variance of genome size (in Talls). Although the genome size in Dwarf is superior to the *average* genome size of Talls, it remains within the range of Tall coconuts. It is also the case of the three additional individuals we sampled in population hybrids (one Tall × Tall, 2C = 6.13 pg and two Dwarf × Tall, 2C = 5.90 pg and 5.92 pg respectively).

Our results reveal limited (CV = 1.7%) but significant variation in genome size in coconut. These variations occur both in the Indo-Atlantic and in the Pacific genetic groups (respective means and confidence intervals 6.01 [5.79 ±6.25] and 5.90 [5.76 ±6.09]), but they could not be detected among Dwarfs.

Genome size in Attaleinae

Within the Attaleinae subtribe, the holoploid genome sizes were as follows: *Voanioala gerardii* = 60 pg (Johnson et al. 1989), *Allagoptera caudescens* (Mart.) Kunze = 10.70 pg, *Attalea* sp. = 4.02 ±4.34 pg, *Butia* sp. = 3.06 ±3.42 pg, *Beccariophoenix sp*. = 3.6 ±7.47 pg, *Cocos nucifera* = 5.966 ± 0.111pg, *Jubaeopsis caffra* Becc. = 20.98 pg, *Lytocaryum weddellianum* (H. Wendl.) Toledo = 3.72 pg and

309 *Syagrus sp.* = 3.9 ±6.9 pg. The genome size of *Beccariophoenix madagascariensis*
310 Jum. and H. Perrier was 3.6 pg whilst that of its sister taxon *Becc. alfredii* was almost
311 twice (7.47 pg) suggesting that the latter is a tetraploid.

312 ***Reconstruction of genome size (2Cx) evolution in Attaleinae*** [2] The most
313 recent common ancestor (TMRCA) is defined as the most recent lineage from which
314 two diverging lineages were descended. The inferred ancestral genome size of the
315 MRCA of the Attaleinae based on the maximum likelihood topology (second internal
316 node, Fig. 3) was 4.95 pg and it was 5.20 pg for the African/Malagasy and South
317 American clades. The genome size of the MRCA of *Beccariophoenix* and *Voaniaola*
318 + *Jubaeopsis* was 5.81 pg and the inferred genome size for TMRCA of *Voaniaola* +
319 *Jubaeopsis* was 6.12 pg. The inferred ancestral genome size for *Cocos nucifera* was
320 5.90 pg. The genome size of TMRCA of the *Cocos*/*Syagrus* clades was 4.97 pg and
321 for paraphylectic *Syagrus*, the genome size of the TMRCA of the two major clades
322 was 4.90 pg. The MRCA of *Attalea /*(*Allagoptera* + *Allagoptera* + *Parajubaea*)
323 clades was 4.86 pg (Fig. 3). Genome size amongst *Butia* appears to be the smallest
324 (3.06 pg) with inferred ancestral genome size leading to the MRCA of *Jubaea*
325 *chilensis* + *Butia* clade being 4.54 pg, showing a reduction in *Butia* but an increase in
326 the closely related *J. chilensis* (5.1 pg).

327 **DISCUSSION**

328 ***Genome size in coconut and its variations***

329 Our results indicate that the genome size of the coconut is 5.963 ± 0.111 pg or
330 5.757 Gbp. This value differs from the results obtained through Feulgen-
331 microdensitometry by Röser et al. (1997). In addition, the 4C value of *Cocos nucifera*
332 was reported inconsistently by these authors: indeed in Table 3 the value was 14.19
333 pg while in the Results and Discussion section it was 10.2 pg. Our result is somewhat

334 larger than in Zonneveld et al. (2005) and is consistent with the estimate published by
335 Sandoval et al. (2003) based on different cell phases.

336 It has been proposed that genome size has a nucleotypic impact on a number
337 of life history traits including the minimal generation time (MGT), which is long in
338 the case of coconut (Bennett, 1987). However, other factors need to be considered
339 such as adaptation to environmental variations. In particular, families with small
340 genomes are more speciose (Knight, Molinari, and Petrov, 2005). This is the case of
341 Arecaceae, which represents a large family with relatively small genomes among
342 perennial plants (Zonneveld, Leitch, and Bennett, 2005). The influence of nucleotype
343 could however still hold at more restricted evolutionary scale: the coconut genome is
344 about 1.5 times larger than that of the African oil palm *Elaeis guineensis* Jacq. (3.76 ±
345 0.09 pg (Rival et al., 1997) which has a shorter MGT and a higher leaf emission rate.

346 We found that genome size varies significantly among coconuts. This
347 variation is limited (CV = 1.7%) and affects both Indo-Atlantic and Pacific groups.
348 The genome size of the self-pollinating Dwarfs is within the range of the Talls but
349 above average and uniform. This difference was not expected if we consider the
350 positive correlation of genome size with MGT and the negative correlation with
351 stomatal density. In fact, time to flowering in Talls is 4 to 5 years, and only 2 to 3 in
352 Dwarfs (Pillai et al., 1973). Stomatal density is on average 208 mm^{-2} (Talls) and 232
353 mm^{-2} (Dwarfs) (Rajagopal et al., 1990).

354 Plant domestication is an evolutionary process that involves artificial selection
355 and leads to population bottlenecks that can reduce the genetic diversity relative to the
356 wild progenitors through selection of preferred phenotypes (Doebley, Gaut, and Smith,
357 2006). Human selection may affect the patterns of the genome architecture of
358 domesticated plants (Olsen and Wendel, 2013). In the case of coconuts, phenotypic

traits were further influenced by consanguinity resulting from the shift from allogamy to autogamy (see (Miller and Gross, 2011)). This resulted in the expression of genetic load as shown by an increase in the rate of meiotic abnormalities in Dwarfs compared to the Talls, by the poor endosperm development and (at least partly) by reduced vegetative vigor in Dwarfs (Swaminathan and Nambiar, 1961). Considering their uniform and comparatively large genome, the phenotype of the domesticated Dwarfs cannot be accounted for by a nucleotypic influence. The most likely explanation is that Dwarfs were derived from a single Tall ancestor which happened to have a large genome and that this trait has not evolved since then. Coconuts (including Dwarfs) have a long generation time and the number of generations since the appearance of autogamy is probably less than 100.

Evolution of genome size in Attaleinae

The Attaleinae is monophyletic and includes all members of the Cocoseae except the spiny cocosoids (Bactridinae and Elaeidinae), (see (Dransfield et al., 2008). The Cocoseae tribe diverged from its closest relatives *Roystonea* /*Reinhardtia* ca. 55 ± 58 million years ago (mya). Its spiny and non-spiny members diverged about 46 mya (Gunn, 2004; Roncal et al., 2013). Most Attaleinae are diploid while *Allagoptera caudescens*, *Becc. alfredii* Rakotoarin *et al.*, *Jubaeopsis caffra* and *Voanioala* have undergone polyploidization events in the past and have retained their duplicated genomes. A study by Shapcott et al. (2007) on the genetic diversity of the diploid *Becc. madagascariensis* found highly inbred populations. Microsatellite data did not show differentiation between *Becc. alfredii*, and the northern *Becc. madagascariensis* population. It is possible that selfing within these northern populations led to polyploidy with subsequent dispersal by frugivores to new habitats resulting in

speciation. Including *Beccariophoenix alfredii*, a tetraploid shown in this current study, we found that polyploidy occurred at least four times within the Cocoseae.

Our phylogenetic analysis suggests that the ancestral genome size for the Attaleinae may have been small (ca. 4.80 pg). We observed some variability in genome size at the generic level but genome size within a given genus was broadly conserved except for *Syagrus glaucescens* and *S. romanzoffiana* (Fig. 3). The Attaleinae diversified in South America and for the highly speciose taxa such as *Syagrus, Attalea* and *Butia*. In general, their genome sizes are much smaller than the species poor Malagasy/African clade (*Beccariophoenix*, *Voanioala* and *Jubaeopsis*) it is possible that small genome size may play a role providing competitive advantages for these South American taxa to diversify into different biomes as small genome size has been shown to correlate with shorter minimum generation time (MGT), increased reproductive rate and reduced reproductive costs especially in perennial diploid monocots (Bennett, 1972; Midgley, 1991). Our study suggests a role for domestication in genome size evolution and revealed that polyploidy is relatively common within the Attaleinae and has evolved multiple times independently.

Towards coconut genome sequencing

Our research has implications for the future sequencing and annotation of coconut nuclear genome. To date, the genome of two economically important palms have been sequenced and published namely for *Phoenix dactylifera* (estimated 1C ~ 671Mb by Al-Mssallem et al. (2013)) and *Elaeis guineensis* (1C~1.8 Gb according to Singh et al. (2013)). There is also a draft genome sequence available for *E. oleifera* (Filho et al., 2015). Its long generation time and bulkiness make coconut breeding a lengthy process. Thus, marker assisted selection and genomic breeding are likely to accelerate genetic progress. Transcriptomes produced through Next Generation

Sequencing have already been published by (Fan et al., 2013) and (Huang et al., 2014). A preliminary draft coconut genome sequence was presented by Alsahiati et al. (2014) without prior estimation of genome size and variation among cultivars. The coconut genome is 4 and 1.6 times larger than the date palm and oil palm respectively, which requires a much deeper sequencing effort. In addition, a larger genome means that more repeated sequences are present thus causing increased difficulty for assembly. This difficulty can however be overcome by combining the extension of scaffold using paired-end generation of large sequences with the production of a high density linkage map. Whole genome sequencing will pave the way to a variety of approaches such as SNP discoveries from genome wide association studies (GWAS). The whole genome sequence of the coconut will provide us with insights into decoding the traits associated with fruit morphology and more importantly to enable the discovery of QTLs associated with disease resistance such as lethal yellowing through association studies and mapping. Comparative genomics involving oil palm and date palm genome sequence will help elucidate key cellular mechanisms amongst Arecaceae.

LITERATURE CITED

ABRAHAM, A., AND P. M. MATHEW. 1963. Cytology of coconut endosperm. *Annals of Botany* 27: 505–513.

ADAMS, K. L., AND J. F. WENDEL. 2005. Polyploidy and genome evolution in plants. *Current Opinions in Plant Biology* 8: 135–141.

AL-DOUS, E. K., G. BINU, M. E. AL-MAHMOUD, M. Y. AL-JABER, H. WANG, Y. M. SALAMEH, E. K. AL-AZWANI, et al. 2011. *De novo* genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nature Biotechnology* 29: 521–527.

AL-MSSALLEM, I. S., S. HU, X. ZHANG, Q. LIN, W. LIU, J. TAN, X. YU, et al. 2013. Genome sequence of the date palm Phoenix dactylifera L. *Nat Commun* 4.

ALSAIHATI, B., W. LIU, Q. LIN, AND I. S. ALMSSALLEM. 2014. Coconut genome *de novo* sequencing. Poster., Plant and Animal Genome, San Diego, California, USA.

BAUDOUIN, L., B. F. GUNN, AND K. OLSEN. 2014. The presence of coconut in southern Panama in pre-Columbian times: clearing up the confusion. *Ann Bot* 113: 1 - 5.

BENNETT, M. D. 1972. Nuclear DNA content and minimum genration time in herbaceous plants. *Proc. R. Soc. Lond.* 181: 109 - 135.

BENNETT, M. D. 1987. Variations in genomic form in plants and its ecological implications. *New Phytologist* 106: 177–200.

------. 1998. Plant genome values: How much do we know? *Proceedings of the National Academy of Sciences* 95: 2011–2016.

BENNETT, M. D., P. BHANDOL, AND I. J. LEITCH. 2000. Nuclear DNA amounts un Angiosperms and their modern uses - 807 new estimates. *Annals of Botany* 86: 859–909.

BOURDEIX, R., P. BATUGAL, J. OLIVER, AND M. GEORGE. 2010. Catalogue of conserved coconut germplasm.

CHAMBERS, J. M., W. S. CLEVELAND, B. KLEINER, AND P. A. TUKEY. 1983. Graphical methods for data analysis, 395 pp. Wadsworth International Group.

CHURCH, G. L. 1940. Cytotaxonomic studies in Gramineae: *Spartina*, *Andropogon*, and *Panicum*. *American Journal of Botany* 27: 263–271.

COBA DE LA PEÑA, T., AND S. C. BROWN. 2001. Flow cytometry. *In* C. Hawes AND B. Satiat-Jeunemaitre [eds.], Plant Cell Biology: A practical approach. Oxford University Press, Oxford.

COMAI, L. 2005. The advantages and disadvantages of being a polyploid. *Nature Reviews. Genetics* 6: 836–846.

DOEBLEY, J. F., B. S. GAUT, AND B. D. SMITH. 2006. The molecular genetics of crop domesticatiom. *Cell* 127: 1309–1321.

DOLEZEL, J., AND J. BARTOS. 2005. Plant DNA Flow Cytometry and estimation of nuclear genome size. *Annals of Botany* 95: 99–110.

DOLEZEL, J., P. BINAROVA, AND S. LUCRETTI. 1989. Analysis of nuclear DNA content in plant cells by flow cytometry. *Biologia Plantarum (PRAHA)* 31: 113–120

DOLEZEL, J., J. GREILHUBER, AND J. SUDA. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* 2: 2233–2244.

DRANSFIELD, J., N. W. UHL, C. B. ASMUSSEN-LANGE, W. J. BAKER, M. M. HARLEY, AND C. E. LEWIS. 2008. Genera Palmarum: The evolution and classification of palms., 397–454. Kew Publishing, Royal Botanic Gardens, Kew.

FAN, H., Y. XIAO, Y. YANG, W. XIA, A. MASON, Z. XIA, F. QIAO, et al. 2013. RNA-Seq analysis of *Cocos nucifera:* Trancsriptome sequencing and *De Novo* assembly for subsequent functional genomics approaches. *PloS One* 8: (3): e59997.

FILHO, J. A. F., F. P. LOBO, M. T. SOUZA JUNIOR, A. A. ALVES, AND E. F. FORMIGHIERI. 2015. Genome assembly of Amazonian Elaeis oleifera genotype and comparative analysis with other oil palm genomes. Poster., Plant and Animal Genome XXIII, San Diego, California, USA.

FOALE, M. 2005. An introduction to the coconut palm. *Coconut Genetic Resources*: 1.

GAUT, B. S., AND J. F. DOEBLEY. 1997. DNA sequence evidence for the segmental allotetraploid origin of maize. *Proceedings of the National Academy of Sciences* 94: 6809–6814.

GREILHUBER, J. 1986. Severely distorted Feulgen-DNA amounts in Pinus (Coniferophytina) after nonadditive fixations as a result of meristematic self-tanning with vacuole contents. *Canadian Journal of Genetics and Cytology* 28: 409-415.

GREILHUBER, J. 2005. Intraspecific variation in genome size in Angiosperms: Identifying its existence. *Annals of Botany* 95: 91—98.

GREILHUBER, J., J. DOLEZEL, M. A. LYSAK, AND M. D. BENNETT. 2005. The origin, evolution and proposed stabilization of the terms "Genome Size" nd "C-value" to describe nuclear DNA contents. *Annals of Botany* 95: 255–260.

GRIME, J. P. 1983. Prediction of weed and crop response to climate based upon measurements of nuclear DNA content, 4. National Vegetable Research Station: The Association of Applied Biologists, Warwick.

GRIME, J. P., AND M. A. MOWFORTH. 1982. Variation in genome size - an ecological interpretation. *Nature* 299: 151—153.

GUINDON, S., AND O. GASCEUL. 2003. PhyML: A simple., fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52: 696–704.

GUNN, B. F. 2004. The phyogeny of the Cocoeae (Arecaceae) with emphasis on Cocos nucifera. *Annals of the Missouri Botanical Garden* 91: 505–522.

GUNN, B. F., L. BAUDOUIN, AND K. M. OLSEN. 2011. Independent origins of cultivated coconut (Cocos nucifera L.) in the Old World Tropics. *PloS One* 6: e21143.

HARDIE, D. C., AND P. D. HEBERT. 2003. The nucleotypic effects of cellular DNA content in cartilaginous and ray-finned fishes. *Genome* 46: 683-706.

HARRIES, H. 1978. The evolution, dissemination and classification of Cocos nucifera. L. *Botanical Review* 44: 265-319.

HESLOP-HARRISON, J. S., AND T. SCHWARZACHER. 2007. Domestication, genomics and the future for banana. *Annals of Botany* 100: 1073–1084.

HUANG, Y.-Y., C.-P. LEE, J. L. FU, B. C.-H. CHANG, A. J. M. MATZKE, AND M. MATZKE. 2014. De Novo Transcriptome Sequence Assembly from Coconut Leaves and Seeds with a Focus on Factors Involved in RNA-Directed DNA Methylation. *G3: Genes|Genomes|Genetics* 4: 2147-2157.

KNIGHT, C. A., N. A. MOLINARI, AND D. PETROV. 2005. Large genome constraint hypothesis: Evolution, ecology and phenotype. *Annals of Botany* 95: 177–190.

LAURIE, D. A., AND M. D. BENNETT. 1985. Nuclear DNA content in the genera Zea and Sorghum. Intergeneric, interspecific and intraspecific variation. *Heredity* 55: 307—313.

LEBRUN, P., A. BERGER, T. HODGKIN, AND L. BAUDOUIN. 2005. Biochemical and molecular methods for characterizing coconut diversity. *In* P. Batugal, R. Rao, AND J. Oliver [eds.], Coconut Genetic Resources, 225-247.

MADDISON, W. P., AND D. R. MADDISON. 2008. Mesquite, a modular system for evolutionary analysis. Version 2.5. Computer program distributed by authors, http://mesquiteproject.org [accessed 22 June 2014].

MASTERTON, J. 1994. Stomatal size in fossil plants: Evidence for polyploidy in majority of Angiosperms. *Science* 264: 421–424.

MEEROW, A., L. NOBLICK, J. W. BORRONE, T. COUVREUR, M. MAURO-HERRERA, W. HAHN, D. N. KUHN, et al. 2009. Phylogenetic analyses of seven WRKY genes across the palm subtribe Attaleinae (Arecaceae) identifies Syagrus as sister group of the coconut. *PloS One* 4: (10):e7353.

MIDGLEY, J. J. 1991. Ecological aspects of the rise of angiosperms: a challenge to the reproductive superiority hypotheses. *Biological Journal of the Linnean Soiety* 44: 81–91.

MILLER, A. J., AND B. L. GROSS. 2011. From forest to field: Perennial fruit crop domestication. *American Journal of Botany* 98: 1389-1414.

NAMBIAR, M. C., AND M. S. SWAMINATHAN. 1960. Chromosome morphology, microsporogenesis and pollen fertility in some varieties of coconut. *Indian Journal of Genetics and Plant Breeding* 20: 200–211.

OLSEN, K. M., AND J. F. WENDEL. 2013. A bountiful harvest: genomic insights into crop domestication phenotypes. *Annual review of plant biology* 64: 47-70.

PAGEL, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology* 48: 612–622.

PARRIS, K., J., T. G. RANNEY, H. T. KNAP, AND V. BAIRD, W. 2010. Ploidy levels, relative genome sizes, and base pair composition in Magnolia. *Journal of American Society of Horticultural Science* 135: 533–547.

PICHERSKY, E. 1990. Nomad DNA - A model for movement and duplication of DNA sequences in plant genomes. *Plant Molecular Biology* 15: 437–448.

PILLAI, R., R. NAIR, C. MATHEW, K. BAVAPPA, AND A. RAMADASAN. 1973. Studies on photoperiodic response reaction in coconut. *Journal of Plantation Crops* 1: 89 - 92.

PRICE, J. P., AND M. BARANOVA. 1976. Evolution of DNA content in higher plants. *Botanical Review* 42: 27–52.

R DEVELOPMENT CORETEAM. 2011. R: A language and environment for statistical computing. Computer program distributed by R Foundation for Statistical Computing, website: http://www.R-project.org [accessed on June 22, 2014].

RAJAGOPAL, V., B. KASTURI, K. V. KUMAR, AND S. R. VOLETI. 1990. Screening of coconut genotypes for drought tolerance. *Oléagineux* 45: 215–223.

RILEY, R. D., AND K. P. VOGEL. 1982. Chromosome numbers of released cultivars of switchgrass, indiangrass, big bluestem, and sand blue-stem. *Crop Science* 22: 1082–1083.

RIVAL, A., T. BEULE, P. BARRE, S. HAMON, Y. DUVAL, AND M. NOIROT. 1997. Comparative flow cytometric estimation of nuclear DNA content in oil

palm (*Elaeis guineensis* Jacq) tissue cultures and seed-derived plants. *Plant Cell Reports*: 884–887.

RONCAL, J., F. KAHN, B. MILLAN, T. COUVREUR, AND J.-C. PINTAUD. 2013. Cenozoic colonization and diversification patterns of tropical American palms: evidence from Astrocaryum (Arecaceae). *Botanical Journal of the Linnean Society* 171: 120–139.

RÖSER, M., M. A. T. JOHNSON, AND L. HANSON. 1997. Nuclear DNA amounts in palms (Arecaceae). *Botanica Acta* 110: 79–89.

SANDOVAL, A., V. HOCHER, AND J.-L. VERDEIL. 2003. Flow cytometric analysis of the cell cycle in different coconut palm (*Cocos nucifera* L.) tissues cultured in vitro. *Plant Cell Reports* 22: 25–31.

SHAPCOTT, A., M. RAKOTOARINIVO, R. J. SMITH, G. LYSAOVA, M. F. FAY, AND J. DRANSFIELD. 2007. Can we bring Madagascar's critically endangered palms back from the brink? Genetics, ecology and conservation of the critically endangered palm *Beccariophoenix madagascariensis*. *Botanical Journal of the Linnean Society* 154: 589–608.

SINGH, R., M. ONG-ABDULLAH, E.-T. LOW, M. A. A. MANAF, AND R. ROSLI. 2013. Oil palm genome sequence reveals divergence of interfertile species in Old and New Worlds. *Nature* 500: 335–339.

SPARROW, A. H., AND J. P. MIKSCHE. 1961. Correlations of nuclear volume and DNA content with higer tolerance to chronic radiation. *Science* 134: 282–283.

SPARROW, A. H., S. S. SCHWEMMER, AND P. J. BOTTINO. 1971. The effects of external gamma radiation from radioactive fallout of plants with special reference to crop production. *Radiation Botany* 11: 85–118.

SPARROW, R. C., AND A. H. SPARROW. 1965. Relative radiosensitivities of woody and herbaceous spermatophytes. *Science* 147: 1449–1451.

STEBBINS, G. L. 1959. The role of hybridization in evolution. *Proceedings of the American Philosophical Society* 103: 231–251.

SWAMINATHAN, M. S., AND M. C. NAMBIAR. 1961. Cytology and origin of the Dwarf Coconut Palm. *Nature* 192: 85–86.

SWIFT, H. 1950. The constancy of desoxyribose nucleic acid in plant nuclei. *Proceedings of the National Academy of Sciences* 36: 643–654.

WENDEL, J. F., AND R. C. CRONN. 2003. Polyploidy and the evolutionary history of cotton, Advances in Agronomy, vol. 78, 139–186. Academic Press.

ZONNEVELD, B. J. M., I. J. LEITCH, AND M. D. BENNETT. 2005. First nuclear DNA amounts in more than 300 Angiosperms. *Annals of Botany* 96: 229–244.

607
608 **Table 1** ANOVA of estimated DNA contents (pg).

	Df	Sum Sq	Mean Sq	F value	Pr (>F)
Between types	1	0.04511	0.04511	10.90	0.001680 **
Within Dwarf type	6	0.03318	0.00553	1.34	0.2568
Within Tall type	15	0.64875	0.04325	10.45	2.683×10^{-11}
Residuals	56	0.23183	0.00414		

609

610 **Appendix 1.** Absolute genome sizes (pg) estimated for *Cocos nucifera* L. cultivars

611 sampled with *Petunia hybrida* internal standard, from flow cytometry.

Cultivar	Internat. abbrev.	Habit	N	Abs. genome size/pg (mean±sd)	Origin	Collection Locality
Andaman Ordinary Tall	ADOT	Tall	4	6.02 ± 0.09	Andaman Island	Sta. MD_L03A13
Brazil Green Dwarf	BGD	Dwarf	4	5.94 ± 0.03	Brazil	Sta. MD_L13A28
Catigan Green Dwarf	CATD	Dwarf	4	6.04 ± 0.04	Philippines	Sta. MD_L05A15
Cameroon Kribi Tall	CKT	Tall	2	5.87 ± 0.20	Cameroon	Sta. MD_L12A09
Cameroon Red Dwarf	CRD	Tall	3	6.02 ± 0.02	Cameroon	Sta. MD_L06A13
Gazelle Peninsular Tall	GPT	Tall	3	5.89 ± 0.08	Papua New Guinea	Sta. MD_L08A12
Ghana Yellow Dwarf	GYD	Dwarf	3	5.96 ± 0.03	Ghana	Sta. MD_L02A30
Lizard Island Tall	LIZ	Tall (ws)	4	5.89 ± 0.05	Australia	ANBG_BG753A
Laccadive Micro Tall	LMT	Tall	3	6.13 ± 0.00	Laccadives Archipelago	Sta. MD_L08A18
Mission Beach	MISB	Tall (ws)	2	5.87 ± 0.00	Australia	RBG SYD_20101370
Malayan Tall	MLT	Tall	4	5.79 ± 0.06	Malaysia	Sta.MD_L03A18
Malayan Yellow Dwarf	MYD	Dwarf	2	5.94 ± 0.02	Malaysia	RBG SYD_903153
Mozambique Tall	MZT	Tall	3	6.19 ± 0.04	Mozambique	Sta. MD_L03A13
Niu Leka Dwarf	NLAD	Compact Dwarf	4	5.94 ± 0.06	Fiji	Sta. MD_L08A09
Pilipog Green Dwarf	PILD	Dwarf	6	6.01 ± 0.08	Philippines	Sta. MD_L35A28
Panama Tall	PNT	Tall	4	6.01 ± 0.03	Panama	Sta. MD_L03A12
Solomon Island Tall	SIT	Tall	3	5.96 ± 0.03	Solomon Islands	Sta. MD_L21A13
Sri Lanka Tall	SLT	Tall	4	6.07 ± 0.08	Sri Lanka	Sta. MD_L36A24
Tagnanan Tall	TAGT	Tall	3	5.93 ± 0.00	Philippines	Sta. MD_L38A25
Tahiti Tall	TAT	Tall	3	5.75 ± 0.03	Tahiti	Sta. MD_L03A08
Tahiti Red Dwarf	TRD	Dwarf	3	6.04 ± 0.13	Tahiti	Sta. MD_L14A26
Vanuatu Tall	VTT	Tall	3	5.95 ± 0.03	Vanuatu	Sta. MD_L44A24
West Africa Tall	WAT3	Tall	6	5.89 ± 0.06	West Africa	Sta. MD_L09A14

612

613

Appendix 2. Absolute genome sizes (pg) 2C estimated for Attaleinae species

Species	Abs. genome size /pg (2x)	x	Collection locality	Accession number
Allagoptera caudescens (Mart.) Kunze	5.35	4	RBG, Sydney	20091679
Attalea cohune Mart.	4.34	2	RBG, Sydney	20091583
Attalea phalerata Mart. ex Spreng.	4.02	2	RBG, Sydney	20091585
Astrocaryum alatum H. F. Loomis	4.36	2	RBG, Sydney	20091582
Bactris bifida Mart.	4.10	2	RBG, Sydney	20091209
Bactris gasipaes Kunth	9.43	4	RBG, Sydney	20100250
Beccariophoenix alfredii Rakotoarin et al.	7.47	4	RBG, Sydney	20100251
Beccariophoenix madagascariensis Jum. & H.Perrier	3.6	2	RBG, Sydney	20040914
Butia capitata (Mart.) Becc.	3.42	2	RBG, Sydney	932392
Butia eriospatha (Mart. ex Drude) Becc.	3.06	2	RBG, Sydney	780035
Elaeis oleifera. (Kunth) Cortes	4.43	2		Angiosperm 1C-values db
Jubaea chilensis (Molina) Baill.	5.10	2	RBG, Sydney	20090098
Jubaeopsis caffra Becc.	8.40	5		801080
Lytocaryum weddellianum (H. Wendl.) Toledo	3.72	2	RBG, Sydney	14451
Syagrus botryophora (Mart.) Mart.	4.32	2	RBG, Sydney	20090788
Syagrus coronata (Mart.) Becc.	3.96	2	RBG, Sydney	20091730
Syagrus glaucescens Glaz. ex. Becc.	6.90	2		Angiosperm 1C-values db
Syagrus romanzoffiana (Cham.) Glassman	6.10	2		Angiosperm 1C-values db
Syagrus sancona (Kunth) H.Karst.	3.90	2	RBG, Sydney	20091729
Syagrus schizophylla (Mart.) Glassman	4.00	2	RBG, Sydney	20091652
Voanioala gerardii J. Dransf.	6.32	19		Angiosperm 1C-values db

Notes: Abbrev: Sta. MD = CNRA Marc Delorme Coconut Research Centre in Côte G¶YRLUH$ULFD$%VWUDOLDQ1DWLRQDO%RWDQLF*DUGHQV&QEHUUDDQG5% SYD = Royal Botanic Gardens Sydney, Australia, ws = wild-sown and N = number of repeats.

Figure Legends

Fig. 1. Examples of flow cytometry histograms. A: Peak A: *Petunia* standard alone; B: Peak A: *Petunia* standard, Peak B: *Cocos nucifera* L. G1 represents the non-replicating cell phase.

625 Fig. 2. Boxplot of estimated nucleus DNA content. The thick horizontal line
626 corresponds to the median, the limits of the boxes are the first and the third quartiles.
627 Individual observations are represented by dots.
628 Fig. 3. Ancestral genome size reconstruction: Maximum likelihood phylogenetic tree
629 of Attaleinae based on seven *WRKY* nuclear loci using PhyML (Phylogenetic
630 Analysis of Maximum Likelihood). ML bootstrap supports are in parenthesis below
631 the branches. Sequence alignment will be deposited in Dryad database
632 (http://datadryad.org/). The numbers at the nodes refer to the inferred ancestral
633 genome sizes using maximum likelihood reconstruction approach implemented in
634 Mesquite. Numbers adjacent to the OTUs are the holoploid genome size (2Cx)
635 estimated using flow cytometry with ploidy levels in parenthesis, where 2x denote
636 diploids and >2x denote polyploids. The blue ovals indicate the polyploidy events.
637 Outgroups included were *Elaeis oleifera*, *Bactris major* and *B. brongniartii*.
638
639

Figures

Click here to download Figure: Figs_Gunn_et al_Genome_size evolution.pdf

Fig. 1

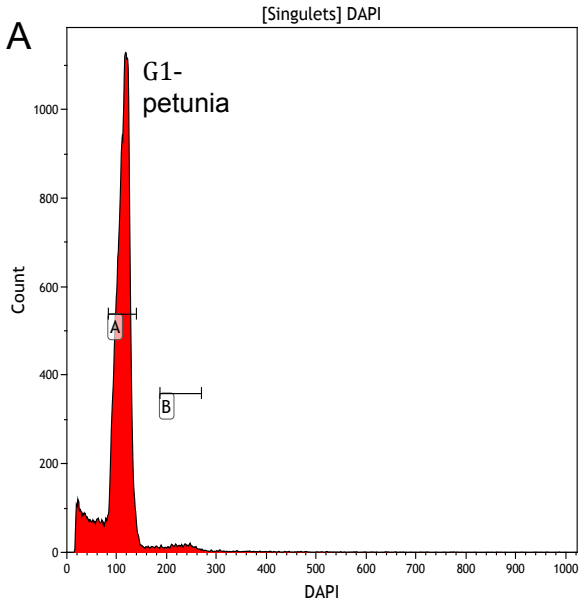


Fig. 2

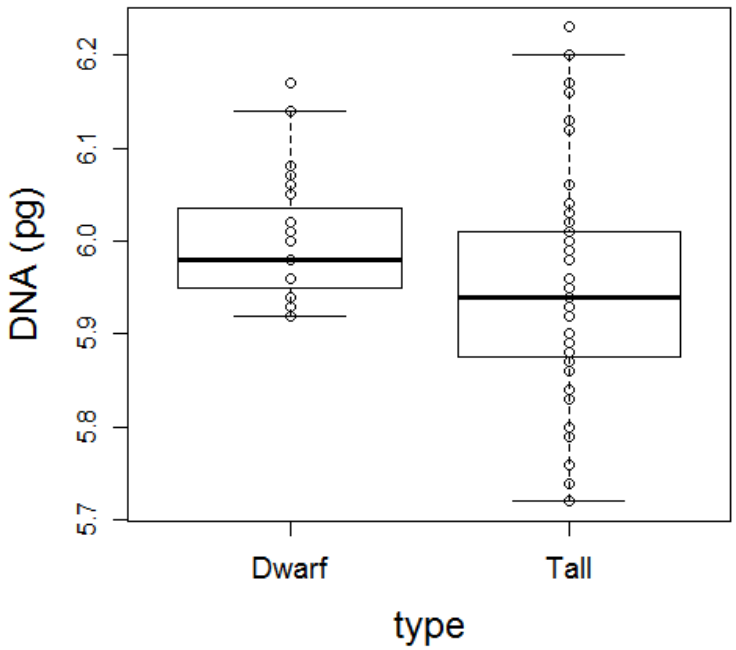
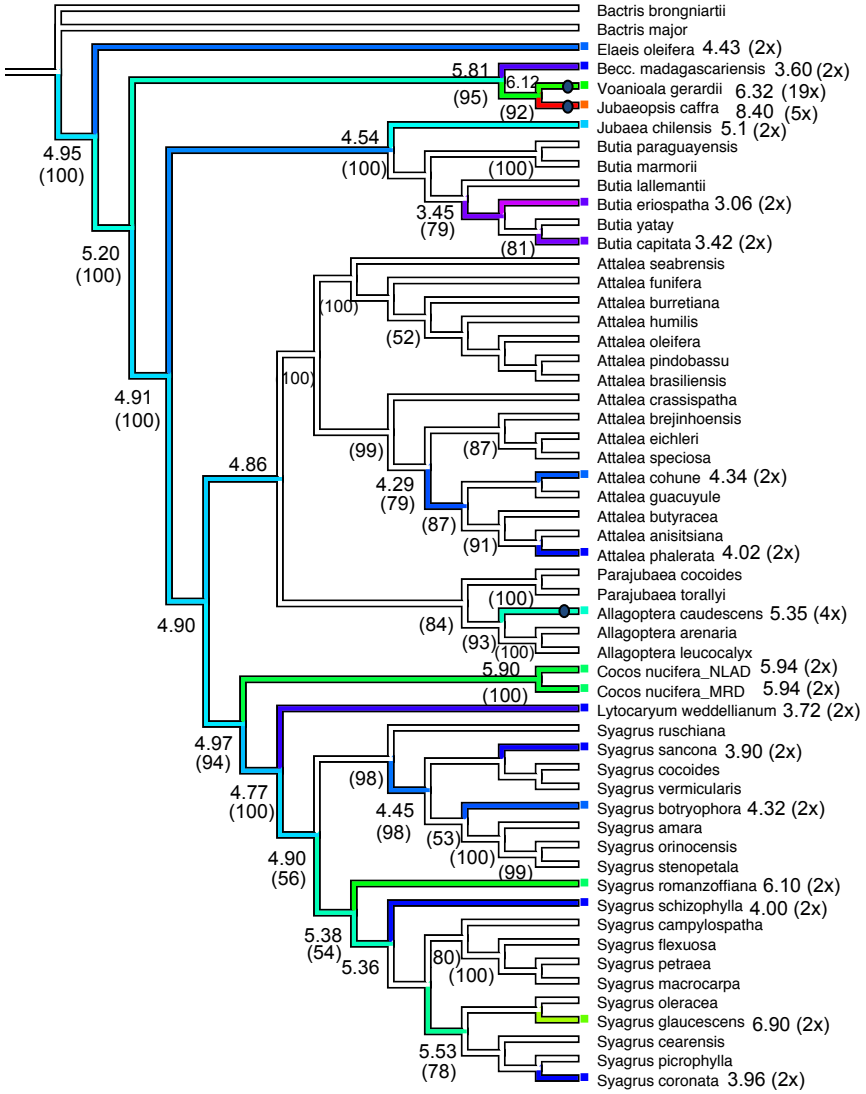


Fig. 3


Bactris brongniartii
Bactris major
Elaeis oleifera 4.43 (2x)
Becc. madagascariensis 3.60 (2x)
Voanioala gerardii 6.32 (19x)
Jubaeopsis caffra 8.40 (5x)
Jubaea chilensis 5.1 (2x)
Butia paraguayensis
Butia marmorii
Butia lallemantii
Butia eriospatha 3.06 (2x)
Butia yatay
Butia capitata 3.42 (2x)
Attalea seabrensis
Attalea funifera
Attalea burretiana
Attalea humilis
Attalea oleifera
Attalea pindobassu
Attalea brasiliensis
Attalea crassispatha
Attalea brejinhoensis
Attalea eichleri
Attalea speciosa
Attalea cohune 4.34 (2x)
Attalea guacuyule
Attalea butyracea
Attalea anisitsiana
Attalea phalerata 4.02 (2x)
Parajubaea cocoides
Parajubaea torallyi
Allagoptera caudescens 5.35 (4x)
Allagoptera arenaria
Allagoptera leucocalyx
Cocos nucifera_NLAD 5.94 (2x)
Cocos nucifera_MRD 5.94 (2x)
Lytocaryum weddellianum 3.72 (2x)
Syagrus ruschiana
Syagrus sancona 3.90 (2x)
Syagrus cocoides
Syagrus vermicularis
Syagrus botryophora 4.32 (2x)
Syagrus amara
Syagrus orinocensis
Syagrus stenopetala
Syagrus romanzoffiana 6.10 (2x)
Syagrus schizophylla 4.00 (2x)
Syagrus campylospatha
Syagrus flexuosa
Syagrus petraea
Syagrus macrocarpa
Syagrus oleracea
Syagrus glaucescens 6.90 (2x)
Syagrus cearensis
Syagrus picrophylla
Syagrus coronata 3.96 (2x)
5.81
(95)
6.13
(92)
4.54
(100)
(100)
4.95
(100)
3.45
(79)
(81)
5.20
(100)
(100)
(52)
100
4.91
(100)
(99)
(87)
4.86
4.29
(79)
(87)
(91)
4.90
(100)
(84)
(93)
(100)
5.90
(100)
4.97
(94)
(98)
4.77
(100)
4.45
(98)
(53)
(100)
(99)
4.90
(56)
5.38
(54)
5.36
80)
(100)
5.53
(78)