Royal Botanic Gardens Kew

Queen Mary University of London

Novel insight into genome size and succulent trait evolution in Aloe (Asphodelaceae)

Kelda Elliott

2019

Dissertation submitted for the degree of Master of Science in Plant and Fungal Taxonomy, Diversity and Conservation awarded by Queen Mary, University of London.

https://doi.org/10.34885/66

Novel Insight into Genome Size and Succulent Trait Evolution in *Aloe* (Asphodelaceae)

KELDA F.V.A. ELLIOTT [1, 2]

[1]*School of Biological & Chemical Sciences, Queen Mary University of London, Mile End Rd., London E1 4NS, UK*
[2]*Jodrell Laboratory, Kew Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK*

Submitted 30 August 2019

Through the evaluation of genomic, anatomical, and physiological characters against the most taxa rich phylogenetic hypothesis of the genus **Aloe**, this research elucidated succulent **trait evolution** and relationships governing the **'succulence syndrome'** within this iconic group. Genome size estimates for 110 species produced 2.7-fold 1C variation and range of 13.49-36.42 (pg) and 1.7-fold 1Cx variation and range of 13.49-22.52 (pg), which demonstrates, **whole genome duplication** is responsible for the majority of genome size variation in the genus. Investigation into **stomatal density** shows that adaxial stomatal densities are higher across all species, that abaxial and adaxial stomatal densities are correlated almost 1:1, and that the stomatal density of both surfaces are negatively correlated to genome size. Water content was evaluated with the measure **saturated water content**, which produced a range of 7.74-41.91(g), raising methodological considerations that have long interrupted attempts to **quantify succulence**. Future analysis of climactic data and genome size could reveal relationships to traits pertaining to succulence, and could be used to inform conservation efforts.

KEYWORDS: Aloe - genome size - quantify succulence saturated water content - stomatal density - trait evolution - 'succulence syndrome' - whole genome duplication

INTRODUCTION

The evolution of traits designed to combat water stress has been necessary for plants throughout geological time. The term 'succulent' applies to a highly diverse paraphyletic group of angiosperms (approx. 3-5%), including upper estimates of 32 orders, 83 families, and 12,500 species acknowledged by the Angiosperm Phylogeny Group (APGIV, 2016) (Grace, 2019; Males, 2017; Arakaki *et al*., 2011; Nyffeler & Eggli, 2010). The major evolutionary innovation uniting these plants is 'succulence'; a condition characterized by the presence of tissue specialized for storing water, and that is often accompanied by a multitude of specific adaptations that together, are referred to as the 'succulence syndrome' (Ogburn & Edwards, 2010). Despite the taxonomic diversity of succulence, lack of phylogenetic signal in its distribution within the angiosperms shows it to be "evolutionarily accessible" (Griffiths & Males, 2017), suggesting environmental regulation. Many succulents occupy xeric habitats and the 'succulence syndrome' is understood to be an ecophysiological phenomenon, driven by the suite of conditions that characterize xeric environments, such as water and salt stress (Ogburn & Edwards, 2010). Arid environments command for development of traits that maximize water use efficiency, and these traits range from whole plant to anatomical and even nuclear scales. Knight & Ackerly (2002) established a lower frequency of plants with large genome sizes in environments with higher temperatures and lower precipitation, and studies examining the relationship between ploidy and genome size have found that on an aridity gradient, the number of polyploid species are positively correlated to aridity (Poggio *et al*., 1989). Both studies indicate that xeric environments negatively select against large genomes. Genome size influences biological aspects spanning traits at nuclear, cellular, and whole plant levels as well as environmental niche suitability and evolutionary persistence (Pellicer *et al*., 2018). The mean

genome size of a genus is negatively correlated to the number of species within the genus, and paleontological evidence supports a negative correlation between a family's mean genome size and the upper limit of its first appearance in geological time, positing that large genomes have been selected against throughout angiosperm evolution (Vinogradov, 2003). Genome size in succulent plants has been examined, but according to Males' 2017 paper, this is the first time it will have been scrutinized with the purpose of relating to succulence.

The focus group of this study is *Aloe* L. (Asphodelaceae, Asparagales), a genus comprised of >500 species native to Africa, Madagascar and the Arabian Peninsula (Grace *et al*., 2015; Carter *et al*., 2011). The biogeographical origin of the *Aloe* lineage has been traced to South Africa in the Early Miocene (ca.19 Mya), with diversification commencing ca.16 Mya; marked by several species radiations allowing dissemination upwards across the African continent to the Arabian Peninsula and outwards to Western Africa and Madagascar, (Grace *et al*., 2015). Many species are highly adapted to their environment, a phenomenon illustrated by high levels of endemism. Endemism has been estimated at 70% in Southern Africa, 90% in Ethiopia, and 100% in Madagascar (Grace *et al*., 2015). Both the wide geographic distribution, a distribution including species occupying habitats ranging from highly xeric to tropical, and adaptive specificity of species enable *Aloes* to provide a novel opportunity to investigate succulent traits in the context of environment (Grace *et al*., 2015; Carter *et al*., 2011).

Despite identification of qualitative traits characterizing succulence in the *Aloe* lineage, less is known quantatively. The genus *Aloe* contributes to the succulence-type characterized by tissues differentiated by ability to photosynthesize: chlorenchyma, photosynthetic tissue capable of providing nutrition for the whole plant body; and hydrenchyma, non-photosynthetic tissue specified to store water for sustaining metabolism during times of drought (Ihlenfeldt, 1985;

Males, 2017; Grace, 2019). This has been referred to as ‘storage succulence’ (Ihlenfeldt, 1985) or ‘partial succulence’ (Ogburn & Edwards, 2010). It is thought, all *Aloes* are succulent, however succulent traits are inherently variable and their plasticity is best interpreted as a spectrum (Ogburn & Edwards, 2010; Males, 2017; Grace, 2019), hence applying a binary view limits our current understanding. The continuous and therefore quantitative nature of succulent traits, allows them to be mapped against phylogenetic hypotheses. By doing so, this research hopes to quantify succulence and characterize the succulent syndrome in *Aloe* within a phylogenetic context.

The first objective of this project was to evaluate genome size variation in *Aloe.* Genome size, a quantitative measurement of DNA, is a trait of interest in botanical research owing to its association to phenotypic characters and adaptive or functional traits (Bennett & Leitch, 2005). A 2C-value quantifies DNA found within the nucleus of a diploid cell whereas a 1C-value is the amount in the nucleus of an unreplicated gamete or haploid cell (Pellicer *et al*., 2018; Dodsworth *et al*., 2015). Cell size, more specifically epidermal and stomatal guard cells, have been positively correlated to genome size, which leads to a negative correlation to stomatal density (Beaulieu *et al*., 2008). Stomatal density is a particularly appropriate measurement as it is associated with water use efficiency and is therefore pertinent to the ‘succulent syndrome’. The second objective was to establish the relationship between genome size and stomatal density in the genus *Aloe*. Water content, the amount of water a plant is capable of storing, is another metric relevant to cell size and succulence, leading to the third objective of quantifying succulence. Including this measurement was made possible by the development of a cost-effective and time-sensitive method for analyzing saturated water content (Ogburn & Edwards, 2012). Through quantification of genome size, stomatal density, and saturated water content as

well as an assessment of their relationships, this study aims to indicate possible methods for quantifying succulence by proxy in the genus *Aloe*.

METHODS

SAMPLING

The framework for sampling was based on a majority rule Bayesian consensus tree constructed using 7 plastid and nuclear genes (ITS, matK, trnL-F, rbcL, psbA, trnQ-rps16, rps16x1), created in 2015 (Grace *et al.*, 2015). The 189 taxa in the published paper were cross-referenced with availability in the Living Collections at the Royal Botanic Gardens, Kew (RBGK). Sampling was prioritized to uniformly reflect geographical and phylogenetic ranges. 110 *Aloe* spp. were sampled for genome size, 54 for stomatal density and saturated water content, and 2 outgroup spp. were sampled for all three measurements [See Appendix A for accession numbers].

GENOME SIZE

Nuclear DNA contents were estimated following the one-step flow cytometry procedure described by Doležel *et al*. (2007). Three isolation buffers were tested to optimize the method obtaining the most accurate results. Most successful across all samples was 'general purpose buffer' (GPB) (Loureiro *et al*., 2007) adjusted to include 3% Triton and supplemented with 3% PVP-40 and beta-mercaptoethanol (Pellicer & Leitch, 2014; Powell, Pers. Comm.), followed by Galbraith buffer (Galbraith *et al*., 1983), and Sysmex OxProtect buffer (Sysmex, United Kingdom). Several internal calibration standards were used; *Secale cereal* L. "Daňkovské" (1C=8.09 (pg)) (Doležel *et al*., 1998), *Pisum sativum* L. "Minerva Maple" (1C=4.86 (pg)) (Bennett & Smith, 1991), and *Allium cepa* L. "Ailsa Craig" (1C=16.77 (pg)) (Bennett & Smith, 1991). Leaf material was prepared by removing the hydrenchyma, leaving only chlorenchyma,

and was blotted to remove any remaining cellular contents of the hydrenchyma tissue. 1ml of isolation buffer was added to a petri dish, approximately 0.5cm^2 of calibration standard leaf material was chopped rapidly with a fresh razor blade into the buffer with approximately 2.5 cm^2 of leaf sample material, which was chopped slowly to avoid rupturing the nuclei (Powell, Pers. Comm.), another 1ml of isolation buffer was added, and the homogenate was filtered through a 30µm nylon mesh (Celltrics 30µM mesh, Sysmex, Goritz, Germany). For particularly difficult samples, further filtration was necessary, using a cottonwool ball soaked in the isolation buffer, and wedged in a pipette tip (Lee & Lin, 2005). This prohibited large polysaccharide chains characteristic of *Aloe* leaves from contaminating the sample. The suspension was incubated for 10 min. on ice in the dark. Prior to analysis, 100µl of propidium iodide (1 mg/mL) was added to the suspension and. The relative fluorescence of a minimum of 5000 particles was recorded using either a Partec Cyflow® SL3 (Partec GmbH, Münster, Germany) or Sysmex Partec CyFlow®Space (Sysmex Partec GmbH, Görlitz , Germany) flow cytometer, both fitted with a 100 mW green solid-state laser (532 nm, Cobolt Samba, Solna, Sweden). Output histograms were analyzed with FloMax®software v.2.4 (Partec GmbH, Münster, Germany). Three technical replicates were run per sample, and 2C-values calculated by averaging the three results. 1Cx-values were calculated by dividing the 2C-value by the ploidy. Not all ploidy levels were published. For these species, ploidy was estimated by comparing their genome size to the genome sizes of their closest diploid relatives (Viruel *et al*., 2019).Values were submitted to the Plant DNA C-values database at RBGK (https://cvalues.science.kew.org/).

STOMATAL DENSITY

Epidermal peels were made by halving leaves longitudinally along the margin, removing hydrenchyma, scraping the remaining chlorenchyma away with a razor blade until only the clear,

single-celled layer of epidermis remained, and mounting it onto a microscope slide in glycerol. A Leica DMLB compound microscope (Leica Microsystems CMS GmbH, Wetzlar, Germany) fitted with a CMOS Leica DMC5400 camera (Leica Microsystems CMS GmbH, Wetzlar, Germany) and Leica Application Suite v.4.13.0 [Build : 167] software (Leica Microsystems (Switzerland) Ltd. & Leica Microsystems CMS GmbH, Wetzlar, Germany) was used to view and capture epidermal images at 20X objective (200X total) magnification. Densities were calculated by counting the number of stomata in the image. The stomata per image area was then converted to the stomatal density standard of stomata/mm^2.

A primary investigation into the uniformity of stomatal density along a leaf was done concurrently with an investigation into whether abaxial and adaxial leaf surfaces have similar stomatal densities. An exemplar *Aloe* sp. of unknown identity was sampled by taking 6 evenly spaced epidermal peels along the length of both surfaces of 5 leaves. For each leaf, t-tests were run, and in each, the differences were statistically significant, leading to the decision to record densities of both surfaces as well as the ratio of the two.

Per species, a biological replication target of 3 accessions and 3 leaves per accession was established, with a technical replication target of each surface having measurements for 3 peels per leaf, and 3 fields of view per peel. The number or replicates between species varied according to the availability of plant material. The arithmetic mean of the measurements per species was then used as a representative value for the species. Fields of view were chosen by taking evenly spaced measurements along each peel, and adjusting the image field of view to have stomata on at least one widthwise and one lengthwise edge, creating a mechanism for consistency between measurements while maximizing the number of stomata within an image.

SATURATED WATER CONTENT

Saturated water content was measured using Ogburn & Edwards' (2012) direct saturated measure (SWCmeas), a proxy method for traditional P-V curve methods, that uses the equation:

SWCmeas = [leaf mass (g) at full hydration −dried leaf mass (g)]/ dried leaf mass (g).

A target for biological replicates of 3 accessions per species and 3 leaves per accession was established. Leaves were weighed directly after excision from the plant, followed by artificial hydration via submergence in deionized H_2O. Weight gain was monitored in 24hr intervals until it subsided to ≤ 3% increase, indicating saturation. They were then dried in a food dehydrator at 60°C until complete dehydration. Values were generated by entering saturated and dry weights into the SWCmeas equation, and averaging them across the accessions measured.

STATISTICAL & PHYLOGENETIC COMPARISON ANALYSIS

Statistical analysis was performed in R Studio (RStudio Team, 2015). Single pairwise trait modeling was used to compare all combinations of genome size, stomatal density, and saturated water content measurements. Non-normally distributed traits were log transformed to account for skewed distributions.

Phytools (Revell, 2012), picante (Kembel, 2010), ape (Paradis & Schliep, 2018), geiger (Harmon *et al*., 2008), and phylotools (Zhang, 2017) packages were installed. The original phylogenetic tree was uploaded alongside the dataset. The two were compared, and pruned to reflect species representations, and new files with these changes were written. These were applied to make ancestral trait reconstruction under Maximum Likelihood, providing values for the nodes and the 95% confidence intervals. The ancestral trait reconstruction, pruned tree, and pruned data set were input to create a map of the continuous data onto the tree. This produced a heatmap, a visual representation of the most likely evolution and distribution of quantitative traits throughout the history of the *Aloe* lineage. Phylogenetic uncertainties in the tree discredit

the reliability of values produced in ancestral trait reconstruction and were ignored. The heat maps created are to be used, merely, as an indication as to trait evolution over the history of the genus. [See appendix B for heatmap R scripts]

RESULTS

GENOME SIZE VARIATION

Genome size estimates for 110 species of *Aloe* produced a 2.7-fold 1C variation and range of 13.49-36.42 (pg) and 1.7-fold 1Cx variation and range of 13.49-22.52 (pg) (Figure 1). [see appendix A for species values]. When comparing the values and heatmaps of 1C and 1Cx genome sizes (Figure 2), it is evident that the total variation of genome size in the genus *Aloe* is caused by whole genome duplication rather than repetitive DNA insertions via transposable elements. To validate use of a 1Cx heatmap with inferred values, analysis was performed with only non-inferred values. Trends remained the same, supporting the original analysis.

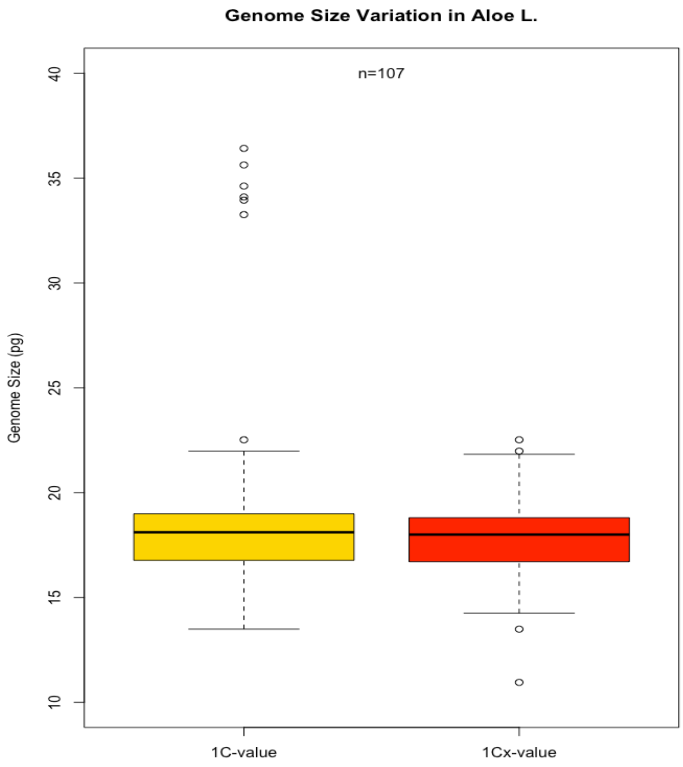

Figure 1. Distibutions of genome size variation for a sample of 107 Aloe ssp.; 1C-value distribution in yellow and 1Cx-value distribution in red

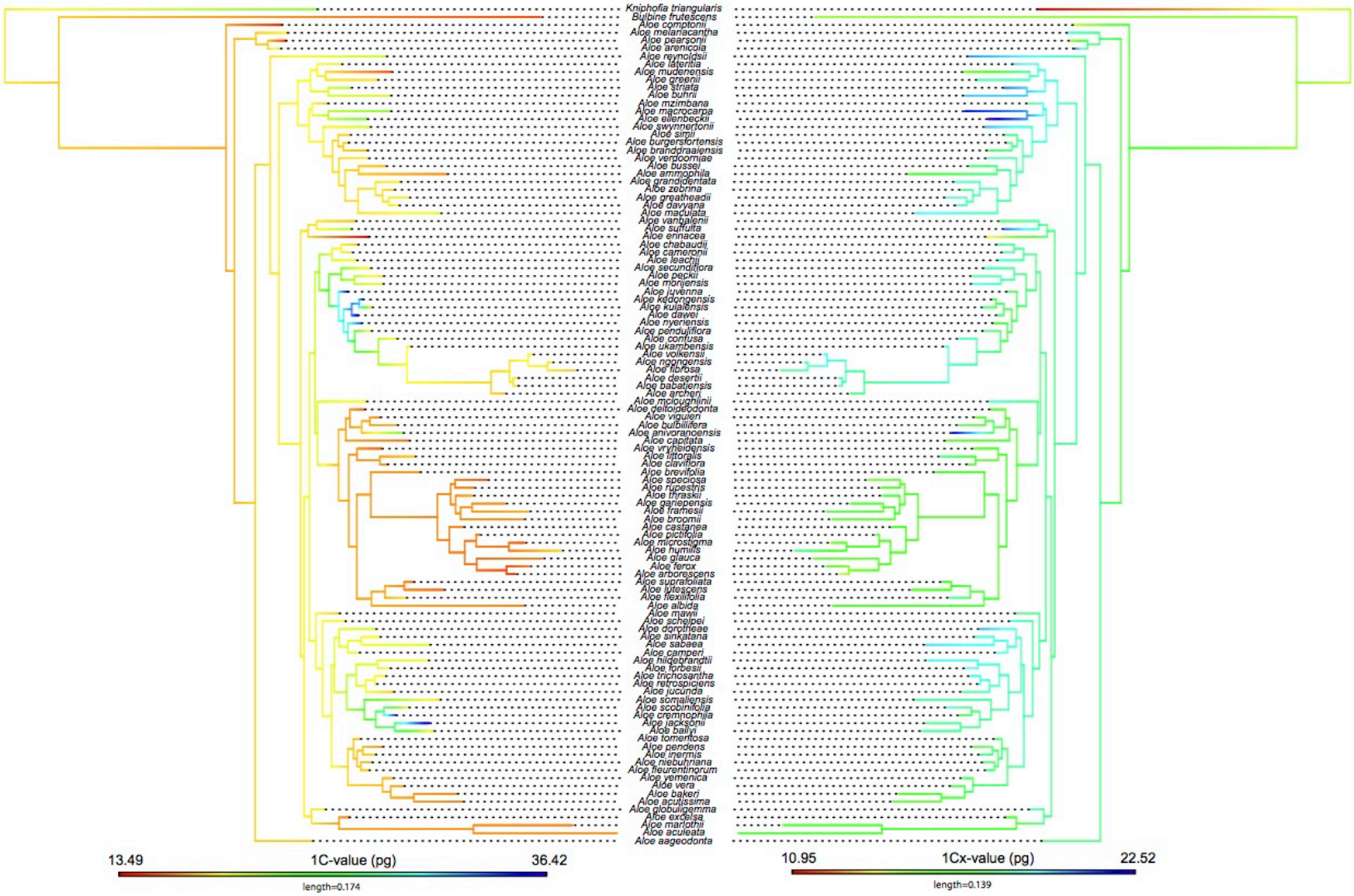

Figure 2. Heatmaps, created using Maximum Liklihood ancestral trait reconstruction values, showing the evolution of genome size throughout the history of the genus Aloe: 1C-values are mapped on the left and 1Cx-values on the right. This clearly illustrates that the majority of genome size variation in Aloes, is due to whole genome duplication

STOMATAL DENSITY

Stomatal density measurements were taken for 54 species, leading to an abaxial stomatal density range of 11.00-49.38 stomata/mm^2, an adaxial stomatal density range of 16.01-62.88 stomata/mm^2, and a range in the ratio of abaxial to adaxial stomatal densities of 1:1.1-1:2.535 (Figure 3). The correlation between abaxial and adaxial stomatal densities was almost 1:1 (Figure 4), but adaxial stomatal densities were consistently higher across all species. The ratio of stomatal densities was more stable across biological replicates, both between and within accessions, than any surface in isolation. The ratio of stomatal densities and abaxial stomatal density were positively correlated, but there was no significant correlation between the ratio of stomatal densities the adaxial stomatal density (Figure 4). Heatmaps for all three measurements show no obvious trends in trait evolution (Figure 5).

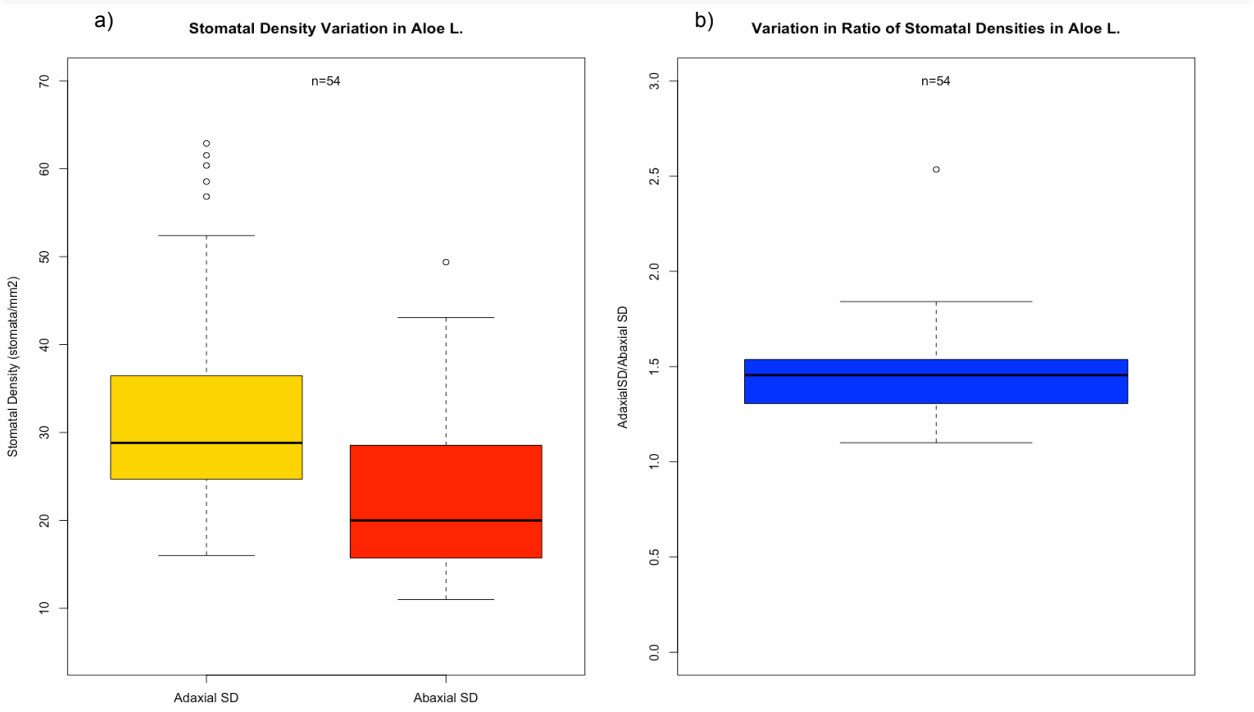

Figure 3. Distributions of stomatal density variation for a sample of 54 Aloe spp.: a) adaxial (yellow) and abaxial (red) stomatal denities b) the ratio between adaxial and abaxial stomatal densities

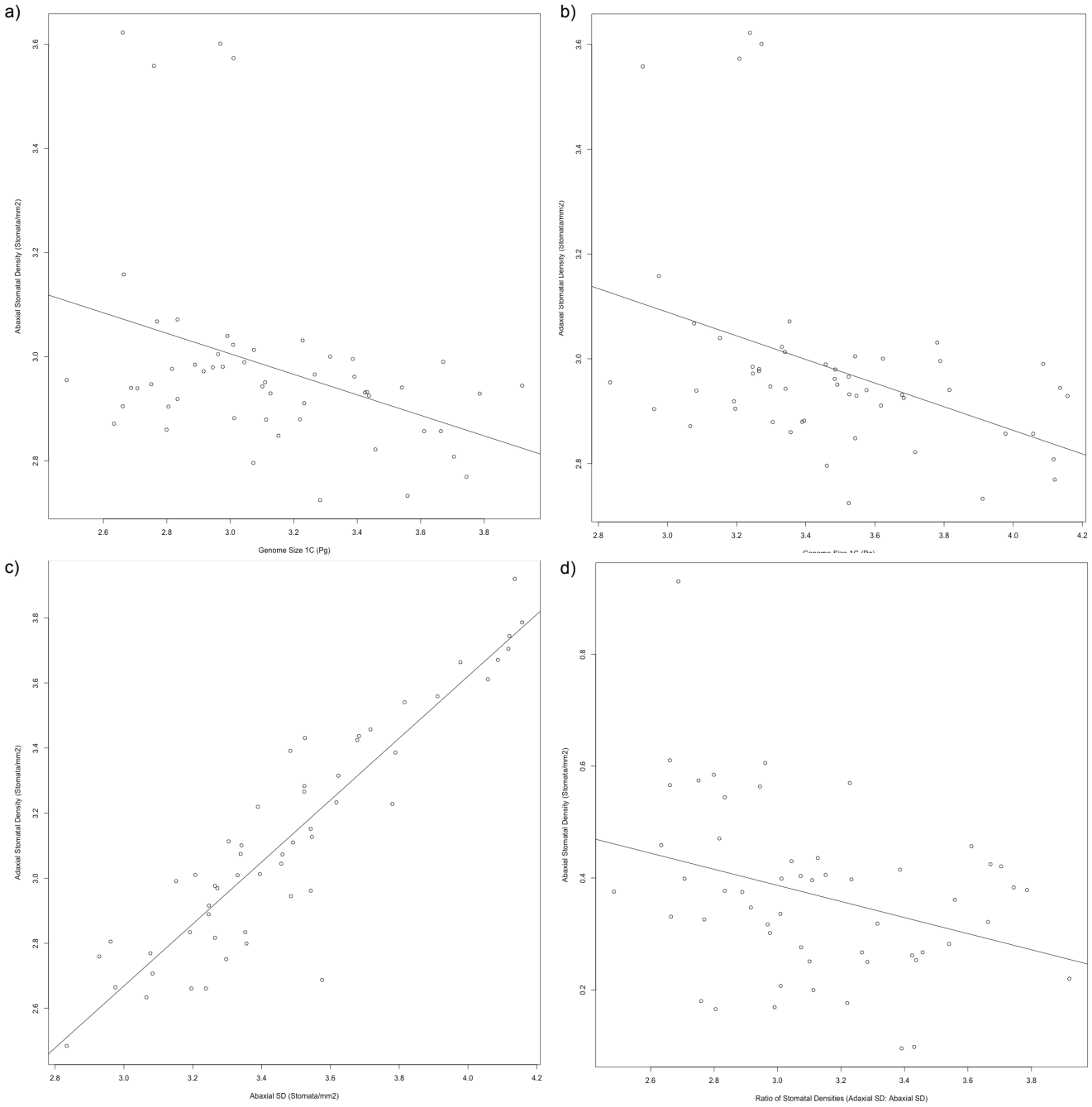

Figure 4. Statistically significant single pairwise linear regressions of 1C genome sizes and stomatal density measurements: a) 1C-value and abaxial stomatal density b) 1C-value and adaxial stomatal density c) Abaxial stomatal density and adaxial stomatal density d) The ratio of adaxial stomatal:abaxial stomatal densities and abaxial stomatal density

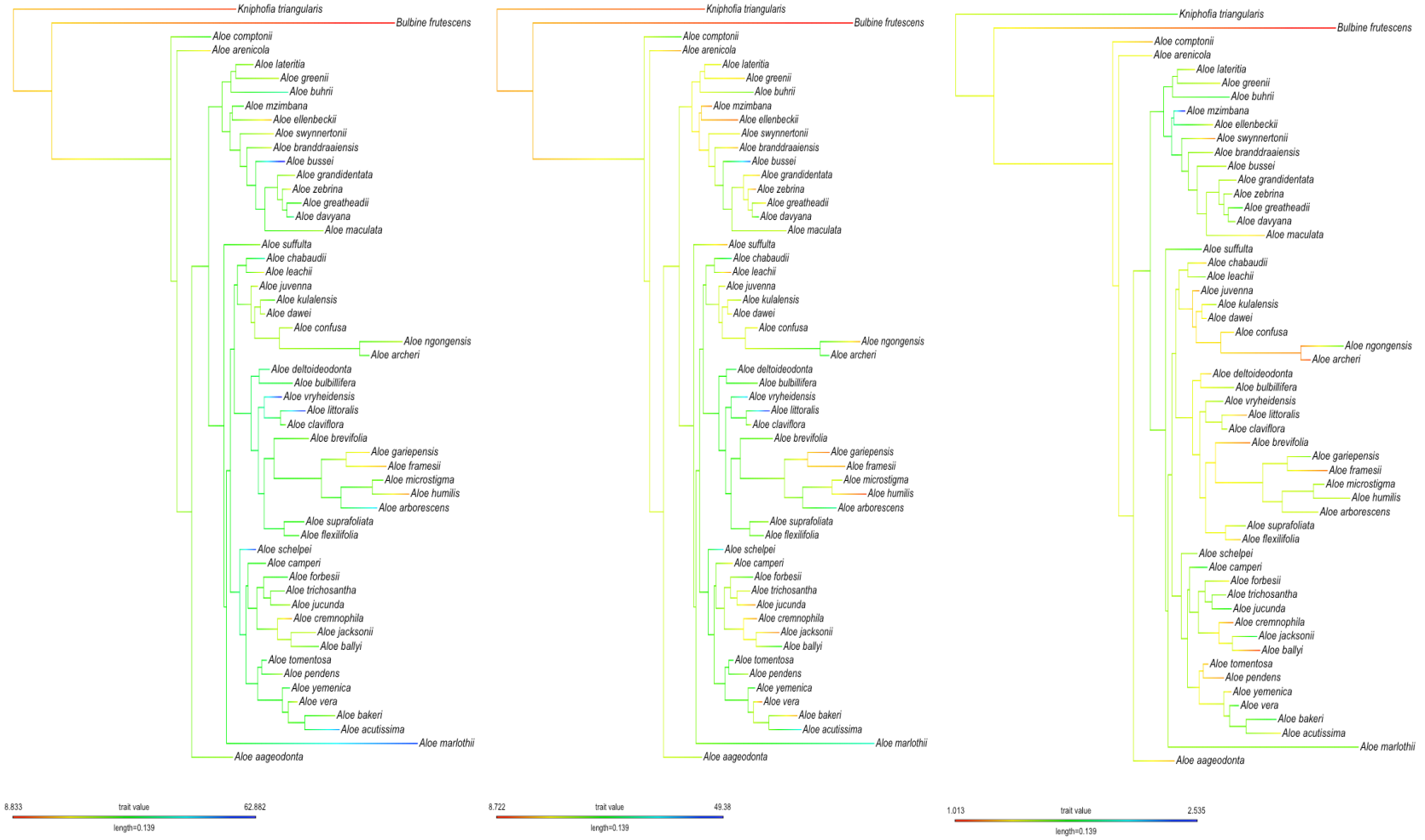

Figure 5. Heatmaps, created using Maximum Liklihood ancestral trait reconstruction values, of Adaxial stomatal density (left), abaxial stomatal density (center), and the ratio of adaxial:abaxial stomatal densities (right).

SATURATED WATER CONTENT

A range of 7.74-41.91(g) was found across the 54 species measured (Figure 6). The stability of the metric across all leaf measurements varied greatly; for some species, a remarkable range of values presented. This is best described by the extreme range in *Aloe greenii* Baker, which has a difference of 28.7g. The only significant correlation to this metric, is that saturated water content is negatively correlated to adaxial stomatal density (Figure 6). This relationship has not previously been reported in the literature and warrants further study.

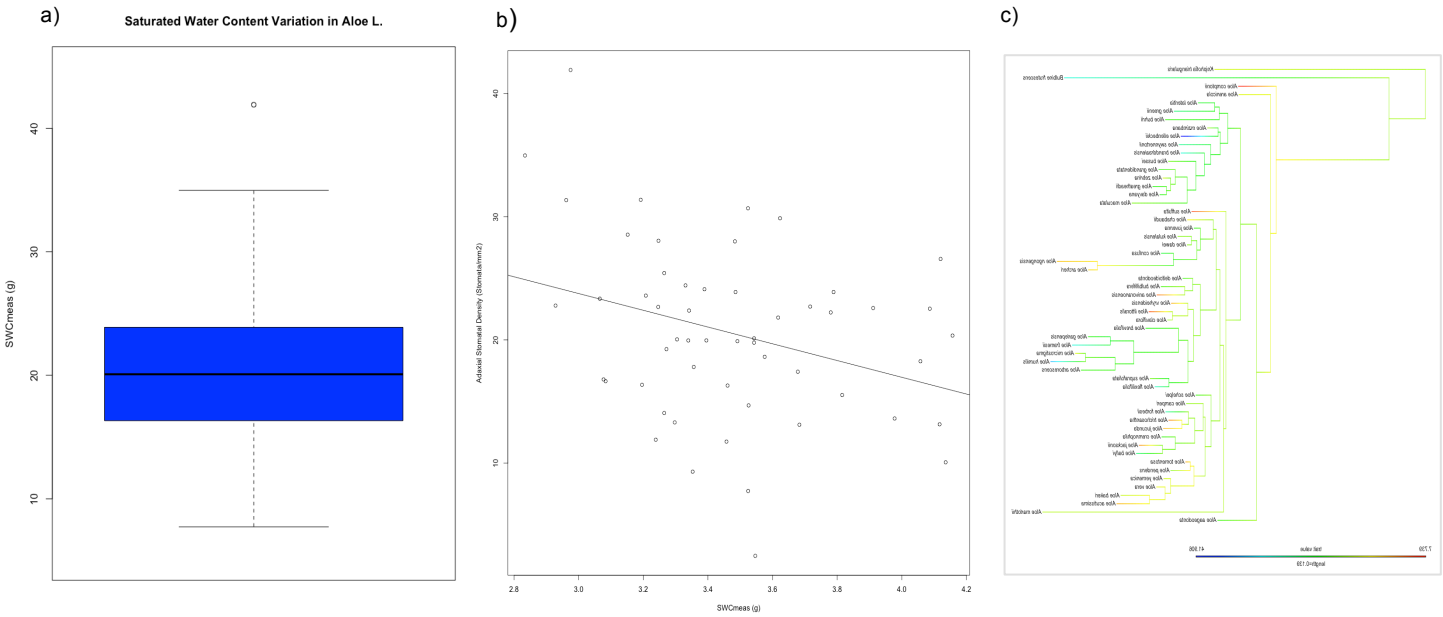

Figure 6. Analyses all used a sample of 54 Aloe spp.: a) Distribution of saturated water content variation b) The statistically significant single pairwise linear regression model of saturated water content and adaxial stomatal density c) Heatmap, created using Maximum Liklihood ancestral trait reconstruction values, of saturated water content

DISCUSSION

GENOME SIZE

This study provides genome estimates for 110 species, increasing total estimates to 178 (62%), extending the 1C-value range to 8.10-36.42 (pg) from 8.10-35.95 (pg) (Leitch *et al.*, 2019). When analyzing genome size variation, the mechanism is questioned: There are two mechanisms

accounting for the majority of genome size variation in plants, repetitive DNA insertions and whole genome duplication events, both of which are common in angiosperms (Knight *et al.*, 2005). Existing karyotypes of *Aloe* species show invariable base chromosome arrangement, with slight variations in morphology; that they all exhibit a conserved, highly orthoselected, bimodal karyotype of seven chromosomes, four large and three small (Brandham, 1971; Cutler *et al.*, 1980; Brandham & Carter, 1990; Brandham & Doherty 1998; Adams *et al.,* 2000; Kumari & Bijoy, 2010; Fentaw *et al.*, 2013; Nejatzadeh-Barandozi & Akbari, 2013; Sánchez-G. *et al.*, 2018). High levels of consistency in chromosome arrangement provides further rationale for the method of inferring ploidy levels for 1Cx-value calculations. Previously reported ploidies include 2n=2x=14, accounting for the vast majority of species, 4n=4x=28, with distribution concentrated in Eastern African and Horn of Africa centers of diversity (Brandham 1971; Brandham & Carter 1979, 1982; Cutler *et al* , 1980; Brandham & Johnson 1982; Brandham & Carter, 1990; Brandham & Doherty, 1998; Kumari & Bijoy, 2010; Fentaw *et al.*, 2013; Sánchez *et al.*, 2018), 2n=6x=42 in *A. cilliaris* Haworth (Muller, 1945), and individuals of 2n=3x=21 in populations of *A. humilis* (L.) Miller (Sharma & Mallick, 1965) and *A. commixta* A.Berger (Brandham, unpublished). This study reveals whole genome duplication or polyploidization to be responsible for the majority of genome size variation, and confirms that species with the largest genomes tend to exist in the Eastern African and Horn of Africa regions.

Even though some habitats in Eastern and Horn of Africa regions are at the extreme end of the aridity gradient, climatic conditions in these regions are highly variable (Camberlin, 2018), and it is likely that some *Aloes* in this region live in less water and nutrient limited environments. When water is the limiting nutrient, availability of other nutrients decreases, as they are not in solution, the state necessary for absorption. Due to the costly nature of DNA synthesis, N and P

requirements are demanding, and their availability has been shown to drive selection towards smaller genomes. This is exemplified by the finding that plants with large genomes can only persist and dominate when levels of N and P are high (Pellicer *et al*., 2018; Guignard *et al.*, 2016; Smarda *et al.*, 2013). Verification of smaller genome sizes in species from water limited environments, in this study, can bee seen in *Aloe erinaceae* D.S. Hardy, from Namibia and *Aloe comptonii* Reynolds, from South Africa, which generated the two lowest genome sizes. This finding aligns with current theoretical expectations. Perhaps it is the climate in the Eastern and Horn of Africa regions that accounts for the ability of polyploid species to persist. The aridity gradient present in *Aloe* species distributions, requires further investigation. Examining the concentration of polyploid *Aloe ssp.* along with accompanying climactic data is recommended.

Several limitations to exploring the role of genome size in contexts of succulence and *Aloes* exist: deficiencies in species representation in phylogenetic hypotheses, genome size estimates, and karyotypes; absence of geographical distributions data for *Aloe* species; and lack of corresponding climactic data. To provide a truly robust analysis, a comprehensive dataset is necessary. Previous phylogenetic analyses of *Aloes* have focused on subfamilial and infrageneric levels (Treutlein *et al.*, 2003; Daru *et al.*, 2013; Grace *et al.*, 2013; Manning *et al.*, 2014; Khodaei *et al.*, 2018), or specific geographic regions (Dee *et al.*, 2018), creating species poor delineations of the genus. Although a remarkable improvement on species inclusivity, the phylogenetic hypothesis made by Grace *et al*. in 2015 only covers a maximum of 20% of the genus, a limitation that underlies this study (Grace *et al,* 2015).

STOMATAL DENSITY

Two existing qualities acknowledged in previous research on stomatal density, were of curiosity. The first, that the stomatal density of the middle section of both length and width of a leaf is

representative of the mean stomatal density of the whole leaf (Smith *et al*., 1989; Willmer & Fricker, 1996; Poole *et al*., 2000). This stems from work on dicot groups, so to confirm the same in *Aloe*, a monocot, a preliminary investigation was performed in an unknown *Aloe* sp. Combined with evidence presenting during the stomatal density measurements, this was substantiated. The second, that because the ancestral trait of angiosperms is the presence of stomata on the abaxial surface, abaxial stomatal densities are usually higher (Mott *et al*., 1982; Willmer & Fricker, 1996). The opposite was found to be true in *Aloes*. Both this and the tight correlation of the two surfaces to each other, can perhaps be attributed to the rosette leaf form characterizing many *Aloe* species. Rosette phyllotaxy is associated with succulence, as it channels precipitation towards the root system and minimizes drying effects of wind and sunlight (Grace, Pers. Comm.). The rosette structure calls into question which surface behaves as the abaxial, as leaves are often upright or incurved (Carter *et al*., 2011), reducing concentrations of light to both surfaces, unlike plants exhibiting definitive upper and lower leaf surfaces and whose differences in surface stomatal densities reflects light exposure. Future investigation would be prudent to consider differences in stomatal densities between species with incurved versus decurved leaves and species with or without rosette phyllotaxy.

Species with the lowest adaxial stomatal densitites are *A. framesii* L. Bolus, *A. humilis* (L.) Miller, *A. gariepensis* Pillans, which together form a species complex that is notoriously difficult to understand due to their sympatric distributions resulting in hybridization, as well as *A. ellenbeckii* A.Berger, and *A. cremnophila* Reynolds & P.R.O Bally (Grace, Pers. Comm.). The species complex is found in the dry succulent Karoo of Southern Africa, and the latter two species from Kenya and Somalia. Species with the highest adaxial stomatal densities include *A. bussei* A.Berger, *A. vryheidensis* Groenewald, *A. littoralis* Baker, *A. shelpei* Reynolds, and *A.*

marlothii A.Berger, a group exhibiting no taxonomic assemblage.

Intentions of investigating three stomatal density measurements, were to determine the most appropriate measurement and to use the ratio of stomatal densities to diminish the effects of variability between biological replicates. The addition of biological replicates influenced averages of stomatal density in abaxial and adaxial leaf surface measurements to a much greater extent than the ratio, corroborating its importance. Differences between heatmaps for stomatal density measurements prove the ratio to be less informative in explaining interspecific variation than stomatal densities of a surface in isolation. The hypothesis that the ratio would be significantly correlated to both adaxial and abaxial stomatal densities was not supported, and it is speculated that this is due to the increased variation in adaxial stomatal density. The complexity found in analysing the relationships between the three measurements, demonstrates that the combination is imperative to build the most accurate picture.

As the parameters of this study were set to reflect upper estimations of quantitative values, the method for choosing fields of view for all stomatal density measurements generated, was designed to reflect this goal, as well as create a method for consistency. The bias towards representing the highest stomatal density created in the methodology for selecting fields of view may have been stronger than intended. Stomatal clustering, a recognized phenomenon (Hoover, 1986; Serna & Fenoll, 1997; Xu & Zhou, 2008; Gan *et al.*, 2010; Franks & Casson, 2014; Lehmann & Or, 2015), has been noted as characteristic of plants in both arid (Hoover, 1986) and water stressed environments (Serna & Fenoll, 1997), conditions that most *Aloe* species experience. Upon reflection of stomatal distributions seen during imaging, this pattern was likely observed in multiple species. To confirm this, an investigation into stomatal densities on different scales of magnification followed by analysis of differences in stomatal densities at each

magnification, is required. This would explain whether measurements in this study are truly representative of mean stomatal densities or of stomatal clusters.

The correlation between genome size and cell size is strong, applies to different cell types, and is typified by the fact that cells of polyploid species are consistently larger than those of their diploid progenitors (Bennett & Leitch, 2005; Beaulieu *et al.,* 2008). Development of smaller guard cells, as they are more reactive to water stress, and higher stomatal densities that maximize diffusion of CO_2 when optimal photosynthetic conditions occur, are examples of traits designed to maximize water use efficiency and is a trait combination limited to species with smaller genomes (Beaulieu *et al.,* 2008). A negative correlation between genome size and stomatal density was previously established (Beaulieu *et al.,* 2008), and the present study finds this to be true for both abaxial and adaxial surfaces in the genus *Aloe*. Applying this concept to species with the lowest adaxial stomatal densities (*A. framesii* (1C=17.25 (pg)), *A. humilis* (1C=18.2 (pg)), *A. gariepensis* (1C=16.66 (pg)), *A. ellenbeckii* (1C=22.52 (pg)), and *A. cremnophila* (1C=34.1(pg))), no consistency to the expectation of seeing higher genome sizes presents, however species with the highest stomatal densitites (*A. bussei* (1C=17.71 (pg)), *A. vryheidensis* (1C=15.58 (pg)), *A. littoralis* (1C=18.00 (pg)), *A.shelpei* (1C=18.89 (pg)), and *A. marlothii* (1C=16.41 (pg))) have a consistency much closer to meeting the expectation. These genome size values are around or below the middle of the 1C genome size range, but no value is at the lowest end. Confirmation of the same negative correlation is a step in answering the question of the relationships existing in this lineage of plants. Future investigation into the environmental conditions of the same species in this study, has potential to reveal explanations for these results and make serious headway in disentangling the relationships.

SATURATED WATER CONTENT

Lack of visible phylogenetic patterns in the heatmap for this trait, suggests environmental conditions may be more responsible for saturated water content than phylogenetic relationships (Figure 6); however, there is no climactic evidence to support this.

Although proved robust in other succulent lineages, species of *Aloe* presented notable issues when using this method. Exudate from freshly excised leaves leaked into the water in which they were being submerged. This is responsible for immediate loss of physical mass, and may also be responsible for further loss from changes in the waters' ionic concentration, drawing mass from within the leaves. Another concern is the variability of weight gain and loss, from day to day, accession to accession, and leaf to leaf. Even though problems in the method were discovered at the beginning of data collection, measurements were continued in order to provide a thorough understanding of how and why these problems arose. If employing this method in the future, it is recommended that measurements be compared to PV curve analyses as well as increasing sample sizes across accessions and number of leaf replicates, to stabilize average values produced.

CONCLUSION & FUTURE RESEARCH CONSIDERATIONS

By applying the saturated water content method for quantifying succulence, evaluating genome size variation, and establishing relationships between genome size, stomatal density, and water content in the genus *Aloe,* this pioneering study lays the foundation for understanding how to approach the quantification of succulence within genomic and phylogenetic contexts. It highlights the necessity for optimization of methodologies applied to the investigation of succulent tissue, presents potential proxy methods for future research into the 'succulence

syndrome' specific to *Aloe*, and draws attention to the importance of relating these findings to an ecological perspective. Xeric habitats are inextricably linked to the development of adaptations that mitigate water stress (Ogburn & Edwards, 2010), and succulent traits range from whole plant to anatomical and even nuclear scales in the case of genome size (Knight & Ackerly, 2002) The 'succulence syndrome' is an ecophysiological theme that has arisen independently in a great diversity of plants, and given current persistent climate change, understanding the 'succulence syndrome' is important in revealing the ecological implications for succulent plants. Analysis of genome sizes of RedList species shows that genome size is positively correlated to a species' extinction threat (Vinogradov, 2003). At this time, developing strategic conservation efforts and priorities is of utmost importance, and although not a direct analysis of how to approach conservation, this research could provide a level of awareness necessary for the process.

Establishing causality when analysing complex relationships is nearly impossible, as in the case of genome size. Despite individual studies reaching decisive conclusions, generally, interactions between factors are not indicated (Turpeinen *et al.*, 1999; Suda *et al.*, 2003; Knight *et al.*, 2005; Basak *et al.*, 2019). It is, however, accepted that the influence of genome size on adaptive traits has significant influences on a plant's ecological niche and evolutionary persistence (Pellicer *et al*., 2018), although it is "often indirectly related to environmental factors" (Bennett, 1987). This is further complicated by cases where genome size does not even have a consistent relationship with an environmental trait across its quantitative range (Knight *et al.*, 2005). As illustrative examples, species with large genomes tend to be excluded from extreme environments, complemented by a consistent relationship between genome size and habitat variability, where species with small genomes tend to be distributed across highly variable habitats (Knight *et al.*, 2005). Other efforts to correlate environment to genome size

include altitude, latitude, and annual and monthly temperature and precipitation levels. Despite the difficulty in such analyses, attemps must be made, in order for the application of this research to be fruitful in aiding conservation strategy.

The first knowledge gap to fill, in reference to *Aloes,* is the absence of rigorously made species distributions and the gathering of climactic data. Characters to appraise are presence of rosette phyllotaxy, whether leaves are incurved or decurved, stomatal density on different scales to identify and account for stomatal clustering, stomatal guard cell length, epidermal cell area, stomatal index, stomatal conductance, ration of hydrenchyma:chlorenchyma, C isotope ratios to show the extent of CAM. Environmental conditions to regard include mean annual temperature, mean annual precipitation, altitude, growing season, and species habitat variability. Future statisical analysis of the data should include estimations of phylogenetic signal using Blomberg *et al's* test for K and Pagel's test for lambda (λ). Linear models for traits exhibiting phylogenetic signal should be done using least squares regressions so as to account for this, and provide a comparison to existing single pairwise regressions. This has potential to reveal some traits as being correlated to others, increasing the understanding of the 'succulence syndrome' in the genus *Aloe.*

ACKNOWLEDGEMENTS

First and foremost a huge thank you to my advisor Olwen Grace for her continued support, and for sharing her research interests and enthusiasm. Gratitude is also extended to Queen Mary University of London for funding, KewScience, Jodrell Laboratory, for supplies, and Richard Gianfrancesco and Andrew Leitch for directing the masters program. Many thanks to The South African National Biodiversity Institute (SANBI) for the donations to Kew's living collections that were sampled in this project as well as Paul Rees and Lara Jewitt for curating the living

collections and providing all sample material. This project would not have been possible without the academic support from Maria Conejero, Illia Leitch, Joseph Morton, Luca Pegoraro, Jaume Pellicer, Robyn Powell, Chrissie Pritchard, Juan Viruel, and Yannick Woudstra, as well as technical support from Ian Ondo, Charlotte Phillips, and Reinis Rozkalns. For this, my overwhelming gratitude is expressed.

REFERENCES

Adams, S. P., Leitch, I. J., Bennett, M. D., Chase, M. W., & Leitch, A. R. 2000. Ribosomal DNA evolution and phylogeny in *Aloe* (Asphodelaceae). *American Journal of Botany*, *87*(11), 1578–1583. https://doi.org/10.2307/2656733

Arakaki, M., Christin, P.A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R.M., Spriggs, E., Moore, M.J., Edwards, E.J. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences, 108* (20) 8379-8384. https://doi.org/10.1073/pnas.1100628108

Basak, S., Sun, X., Wang, G., Yang, Y. 2019. Genome size inafected by variation in morphological traits, temperature, and precipitation in turnip. *Applied Sciences. 9* (2): 253. https://doi.org/10.3390/app9020253

Beaulieu, J. M., Leitch, I. J., Patel, S., Pendharkar, A., Knight, C. A. 2008. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist*, *179*(4); 975–986. https://doi.org/10.1111/j.1469-8137.2008.02528.x

Bennett, M.D. 1987. Variation in genomic form in plants and its ecological implications. *New Phytologist, 106* (s1): 177-200. https://doi.org/10.1111/j.1469-8137.1987.tb04689.x

Bennett, M.D., Leitch, I.J. 2005. Genome size evolution in plants. *Genome Biology and Evolution*. https://doi.org/10.1016/B978-012301463-4/50004-8

Bennett MD, Smith JB. 1991. Nuclear DNA amounts in angiosperms. *Annals of Botany, 76* (2); 113-176. https://doi.org/10.1006/anbo.1995.1085

Brandham, P. E. 1971. The chromosomes of the Liliaceae: II. Polyploidy and karyotype variation in the Aloineae. *Kew Bulleitin, 25*(3): 381-399.

Brandham, P.E. & Carter, S. (1979). The identity of Aloe juvenna. *Cactus & Succulent Journal Great Britain, 41*: 27-29.

Brandham, P.E. & Carter, S. (1982). Aloe juvenna rediscovered. *Cactus & Succulent Journal Great Britain, 44*: 70.

Brandham, P. E., & Carter, S. (1990). A Revision of the Aloe tidmarshii/A. ciliaris Complex in South Africa. *Kew Bulletin, 45*(4), 637-645. doi:10.2307/4113868

Brandham, P. E. & Doherty, M.-J. 1998. Genome Size Variation in the Aloaceae, an Angiosperm Family Displaying Karyotypic Orthoselection. *Annals of Botany*, *82*; 67–73. https://doi.org/10.1006/anbo.1998.0742

Brandham, P.E., & Johnson, M. A. T. 1982. Polyploidy and chromosome interchange in Aloe L. from Somalia. *Kew Bulletin,37*: 387-395.

Camberlin, P. 2018. Climate of Eastern Africa. *Oxford Research Encyclopedias: Climate Science*. doi: 10.1093/acrefore/9780190228620.013.512

Carter, S., Lavranos, J.J., Newton, L.E., Walker, C.C. 2011. Aloes: The definitive guide. *Royal Botanic Gardens, Kew, Richmond.*

Cutler, D.F., Brandham, P.E., Carter, S., Harris, S.J. 1980. Morphological, anatomical, cytological and biochemical aspects of evolution in East African shrubby species of Aloë L. (Liliaceae). *Botanical Journal of the Linnean Society, 80*(4):293-317. https://doi.org/10.1111/j.1095-8339.1980.tb01666.x

Daru, B. H., Manning, J. C., Boatwright, J. S., Maurin, O. , Maclean, N. , Schaefer, H. , Kuzmina, M., & van der Bank, M. 2013, Molecular and morphological analysis of subfamily Alooideae (Asphodelaceae) and the inclusion of *Chortolirion* in *Aloe*. *Taxon, 62*; 62-76. doi:10.1002/tax.621006

Dee, R., Malakasi, P., Rakotoarisoa, S.E., Grace, O.M. 2018. A phylogenetic analysis of the genus *Aloe* (Asphodelaceae) in Madagascar and the Mascarene Islands. *Botanical Journal of the Linnean Society, 187* (3):428-440. https://doi.org/10.1093/botlinnean/boy026

Dodsworth, S., Leitch, A.R., Leitch, I.J. 2015. Genome size diversity in angiosperms and its influence on gene space. *Current Opinion in Genetics & Development, 35*: 73-78. https://doi.org/10.1016/j.gde.2015.10.006.

Doležel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysák, M. A., Nardi, L., Obermayer, R. 1998. Plant genome size estimation by flow cytometry: inter-laboratory comparison. *Annals of Botany 82*(Suppl A):17–26. https://doi.org/10.1093/oxfordjournals.aob.a010312

Doležel J., Greilhuber, J. and Suda, J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature protocols, 2*(9); 2233.

Fentaw, E., Dagne, K., Rønsted, N., Demissew, S., Grace, O.M. 2013. Karyotypes in Ethiopian Aloe species (Xanthorrhoeaceae: Asphodelceae). *Kew Bulletin, 68*(4); 599-607. https://doi.org/10.1007/s12225-013-9475-8

Franks, P. J. & Casson, S. 2014, Connecting stomatal development and physiology. *New Phytol*logist, *201*(4): 1079-1082. https://doi.org/10.1111/nph.12673

Galbraith, D.W., Harkins, K.R., Maddox, J.M., Ayres, N.M., Sharms, D.P., Firoozabady, E. 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science, 220*(4601);1049–1051. https://doi.org/10.1007/BF02907241

Gan, Y., Zhou, L., Shen, Z.J., Shen, Z.X., Zhang, Y.Q., Wang, G.X. 2010. Stomatal clustering, a new marker of environmental perception and adaptation in terrestrial plants. *Botanical Studies, 51*(3); 325–336.

Grace, O. M. (2019). Succulent plant diversity as natural capital. *Plants, People, Planet*. https://doi.org/10.1002/ppp3.25

Grace, O. M., Buerki, S., Symonds, M., Forest, F., van Wyk, A. E., Smith, G. F., Klopper, R.R., Bjorå, C.S., Neale, S., Demissew, S., Simonds, M.S.J., Rønsted, N. 2015. Evolutionary history and leaf succulence as explanations for medicinal use in aloes and the global popularity of Aloe vera. *BMC Evolutionary Biology*, *15*(1); 29. https://doi.org/10.1186/s12862-015-0291-7

Grace O.M., Klopper, R.R., Smith, G.F., Crouch, N.R., Figueiredo, E., Rønsted, N., Van Wyk, A.E. 2013. A revised generic classification for *Aloe* (Xanthorrhoeaceae subfam. Asphodeloideae). *Phytotaxa, 76*(1); 7–14. http://dx.doi.org/10.11646/phytotaxa.76.1.2

Guignard, M. S., Nichols, R. A., Knell, R. J., Macdonald, A., Romila, C. A.,Trimmer, M., Leitch, I. J., Leitch, A.R. 2016. Genome size and ploidy influence angiosperm species' biomass under nitrogen and phosphorus limitation. *New Phytologist*, *210* (4); 1195– 1206. https://doi.org/10.1111/nph.13881

Harmon, L. J., Weir, J.T., Brock, C.D., Glor, R. E., & Challenger, W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics, 18*(1):129-131.

Hoover, W. S. 1986. Stomata and stomatal clusters in begonia: ecological response in two Mexican species. *Biotropica, 18*: 16–21.

Griffiths, H., Males, J. 2017. Succulent plants. *Current Biology, 27*(17); R890-R896, ISSN 0960-9822. https://doi.org/10.1016/j.cub.2017.03.021.

Ihlenfeldt, H-P. 1985. Lebensformen und Überlebensstrategien bei sukkulenten. *Plant Biology*, *98*(1):409-423

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics, 26*:1463-1464.

Khodaei, Z., Wyk, B.-E & Wink, M. 2018. Divergence Time Estimation of Aloes and Allies (Xanthorrhoeaceae) Based on Three Marker Genes. *Diversity, 10*(3); 60. https://doi.org/10.3390/d10030060

Knight, C. A., Ackerly, D. D. 2002. Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. *Ecology Letters, 5*(1); 66–76. https://doi.org/10.1046/j.1461-0248.2002.00283.x

Knight, C. A., Molinari, N. A., & Petrov, D. A. 2005. The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany*, *95*(1); 177–190. https://doi.org/10.1093/aob/mci011

Kumari, G. & Bijoy, K. R. 2010. Karyotype studies in dominant species of Aloe from eastern India, *Caryologia, 63*(1); 41-49. https://doi.org/10.1080/00087114.2010.10589707

Lee, H.-C., Lin, T.-Y. 2005. Isolation of plant nuclei suitable for flow cytometry from recalcitrant tissue by use of a filtration column. *Plant Molecular Biology Reporter, 23*(1); 53-58.

Lehmann, P., Or, D. 2015. Effects f stomata clustering on leaf gas exchange. *New Phytologist, 207*:1015-1025. https://doi.org/10.1111/nph.13442

Leitch, I. J., Johnston, E., Pellicer, J., Hidalgo, O., Bennett, M.D. 2019. Plant DNA C-values database (release 7.1, Apr 2019)

Loureiro, J., Rodriguez, E., Doležel, J. and Santos, C., 2007. Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Annals of Botany*, *100*(4); 875-888.

Males, J. 2017. Secrets of succulence. *Journal of Experimental Botany, 68*(9); 2121-2134. https://doi.org/10.1093/jxb/erx096

Manning, J., Boatwright, J., Daru, B., Maurin, O., Bank, M. 2014. A Molecular Phylogeny and Generic Classification of Asphodelaceae Subfamily Alooideae: A Final Resolution of the Prickly Issue of Polyphyly in the Alooids?. *Systematic Botany, 39*(1); 55–74. https://doi.org/10.1600/036364414X678044

Mott, K. A., Gibson, A. C., O'Leary, J. W. 2006. The adaptive significance of amphistomatous leaves. *Plant, Cell & Environment, 5*(6): 455-460. https://doi.org/10.1111/1365-3040.ep11611750

Muller, F. S. 1945. 'N chromosoomstudie van 'n aantal species van die genus *Aloe* Linn. Met spesiale verwysing na die morfologie en betenkenis van die somatiese chromosome. *Publ. Univ. Pretoria 2*; 1-157.

Nejatzadeh-Barandozi, F., Akbari, L. (2013). Karyotypic Variation of the Aloe vera L. and Aloe littoralis Baker in Iran. *Iranian Journal of Biotechnology*, 11(4), 233-237. doi: 10.5812/ijb.12985

Nyffeler, R. & Eggli, U. 2010. An up-to-date familial and suprafamilial classification of succulent plants. *Bradleya*, *28*; 125–144. https://doi.org/10.5167/uzh-35243

Ogburn, R. & Edwards, E. J. 2010. The ecological water-use strategies of succulent plants. *Advances in Botanical Research, 55*; 179-225. https://doi.org/10.1016/B978-0-12-380868-4.00004-1

Ogburn, R. M., & Edwards, E. J. 2012. Quantifying succulence: a rapid, physiologically meaningful metric of plant water storage. *Plant, Cell & Environment*, *35*(9); 1533–1542. https://doi.org/10.1111/j.1365-3040.2012.02503.x

Paradis E. & Schliep K. 2018. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics, 35*; 526-528.

Pellicer, J., Hidalgo, O., Dodsworth, S., Leitch, I. 2018. Genome size diversity and its impact on the evolution of land plants. *Genes*, *9*(2); 88. https://doi.org/10.3390/genes9020088

Pellicer, J., Leitch, I. 2014. The application of flow cytometry for estimating genome size and ploidy levels in plants. *Molecular Plant Taxonomy: Methods and Protocols*, Methods in

Molecular Biology, 1115; 279-307. doi: 10.1007/978-1-62703-767-9_14, © Springer Science+Business Media New York

Poggio, L., Rosato, M., Chiavarino, A. M., Naranjo, C. 1998. Genome size and environmental correlations in Maize (Zea mays ssp. mays, Poaceae). *Annals of Botany, 82*(Suppl. A); 107–115.

Poole, I., Lawson, T., Weyers, J. D. B., Raven, J. A. 2000. Effect of elevated CO_2 on the stomatal distribution and leaf physiology of *Alnus glutinosa*. *New Phytologist, 145*: 511–521. https://doi.org/10.1046/j.1469-8137.2000.00589.x

Revell, L. J. 2012. Phytools: An R package for phylogenetic comparative biology (and other things). *Methods, Ecology, and Evolution, 3;* 217-223. doi:10.1111/j.2041-210X.2011.00169.x

RStudio Team. 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. http://www.rstudio.com/

Sánchez-G., Y., Raymúndez, M.B., Imery, J., Acosta, M. C., & Moscone, E. 2018. Characterization of eight species of Aloe (Asphodelaceae) from the nucleolar organizing region. *Rodriguésia*, *69*(2); 363-372. https://dx.doi.org/10.1590/2175-7860201869208

Serna, L. and Fenoll, C. 1997, Tracing the ontogeny of stomatal clusters in *Arabidopsis* with molecular markers. *The Plant Journal, 12*(4); 747-755. doi:10.1046/j.1365-313X.1997.12040747.x

Sharma, A. K. & Mallick, R. 1965. Inter-relationships of the tribe Aloinea as reflected in its cytology. *Journal of Genetics 59*

Šmarda, P. M., Hejcman, A., Březinová, L., Horová, H., Steigerová, F., Zedek, P., Bureš, Hejcmanová, P., Schellberg, J. 2013. Effect of phosphorus availability on the selection of species with different ploidy levels and genome sizes in a long-term grassland fertilization experiment. *New Phytologist, 200*(3); 911– 921. doi: 10.1111/nph.12399

Smith S, Weyers JDB, Berry WG. 1989. Variation in stomatal characteristics over the lower surface of *Commelina communis* leaves. *Plant, Cell & Environment* 12: 653–659.

Suda, J., Kyncl, T., Freiova, R. 2003. Nuclear DNA amounts in Macronesian angiosperms. *Annals of Botany, 92*; 152–164.

Treutlein, J., Smith, G. F., Van Wyk, B.-E., Wink, M. 2003. Phylogenetic relationships in Asphodelaceae (subfamily Alooideae) inferred from chloroplast DNA sequences (*rbcL*, *matK*) and from genomic finger- printing (ISSR). *Taxon, 52*(2); 193-207. https://doi.org/10.2307/3647487

Turpeinen, T., Kulmala, J., Nevo, E. 1999. Genome size variation in Hordeum spontaneum populations. *Genome, 42*;1094–1099.

Vinogradov, A. E. 2003. Selfish DNA is maladaptive: evidence from the plant Red List. *Trends in Genetics*, *19*(11); 609–614. https://doi.org/10.1016/j.tig.2003.09.010

Viruel, J., Conejero, M., Hidalgo, O., Pokorny, L., Powell, R. F., Forest, F., Kantar, M. B., Soto-Gomez, M., Graham, S. W., Gravendeel, B., Wilkin, P. & Leitch, I. J. 2019. A target capture-based method to estimate ploidy from herbarium specimens. *Front. Plant Sci.,10*; 937. https://doi.org/10.3389/fpls.2019.00937

Willmer C, Fricker M. 1996. *Stomata*. London, UK: Chapman and Hall.

Xu, Z., Zhou, G. 2008. Responses of leaf stomatal density to water status and its relationship with photosynthesis in a grass, *Journal of Experimental Botany*, *59*(12); 3317–3325. https://doi.org/10.1093/jxb/ern185

Zhang, J. 2017. Phylogenetic tools for building PHYLIP supermatrix and more

SUPPLEMENTARY MATERIAL

APPENDIX A - Sampling & Raw Data

Species	RBGK Accession #s	Region	GS (2C)	GS (1C)	GS (1Cx)	Ploidy	SWCmeas	Abaxial SD(1mm)	Adaxial SD(1mm)	SD Ad:SD Ab
Aloe aageodonta	2011-1187	EA	35.94	17.97	17.97	2	22.383	21.218	27.262	1.285
Aloe aculeata	1995-4276	SA	33.50	16.75	16.75	2	-	-	-	-
Aloe acutissima	2007-919	MAD	32.81	16.41	16.41	2	13.625	37.999	52.402	1.379
Aloe acutissima var. antanimorensis	2010-1941; 1965-67512				0.00	2				
Aloe albidia ~	2019-1172	SA	32.71	16.35	16.35	2	-	-	-	-
Aloe ammophila	1973-2574	SA	33.51	16.75	16.75	2	-	-	-	-
Aloe anivoranoensis	2000-921	MAD	43.96	21.98	21.98	2	10.928	-	-	-
Aloe arborescens	1974-29906; 1973-3969; 1973-2525	SA	28.75	14.38	14.38	2	22.585	34.120	48.963	1.435
Aloe archeri	1977-2960	EA	35.54	17.77	17.77	2	14.686	29.898	32.984	1.103
Aloe arenicola	1995-3288; 1977-6915	SA	40.98	20.49	20.49	2	16.793	14.945	20.692	1.385
Aloe babatiensis	1974-4465	EA	37.02	18.51	18.51	2	-	-	-	-
Aloe bakeri	2007-2342; 1998-3512	MAD	32.93	16.46	16.46	2	17.807	15.431	27.680	1.794
Aloe ballyi	1996-4735	EA	36.65	18.33	18.33	2	27.997	28.702	31.575	1.100
Aloe bellatula ~	2019-1174	MAD	32.79	16.39	16.39	2	-	-	-	-
Aloe branddraaiensis	1957-14502	SA	35.04	17.52	17.52	2	31.370	16.010	23.340	1.458
Aloe brevifolia	2009-1310; 1953-23205	SA	33.62	16.81	16.81	2	24.116	24.015	28.644	1.193
Aloe broomii	2003-2526	SA	33.59	16.79	16.79	2	-	-	-	-
Aloe buetneri ~	2010-1747	WA	48.65	24.33	24.33	2	-	-	-	-
Aloe buhrii	1993-1708; 1985-695	SA	39.44	19.72	19.72	2	22.231	24.227	42.821	1.768
Aloe bulbillifera	1971-2681; 1965-67515	MAD	34.74	17.37	17.37	2	21.813	24.352	36.238	1.488
Aloe bulbillifera var. paulianiae	1993-787									
Aloe burgersfortensis	1965-72105	SA	36.14	18.07	18.07	2	-	-	-	-
Aloe bussei	1990-1816	EA	35.42	17.71	17.71	2	20.353	43.079	62.882	1.460
Aloe cameronii	1970-3558	ZAM	36.92	18.46	18.46	2	-	-	-	-
Aloe camperi	1973-13456	HOA	38.37	19.18	19.18	2	19.777	18.324	33.563	1.832
Aloe capitata	1999-3428	MAD	32.45	16.23	16.23	2	-	-	-	-
Aloe castanea	1984-2332	SA	32.86	16.43	16.43	2	-	-	-	-
Aloe chabaudii	1996-1526; 1984-3808	ZAM	35.86	17.93	17.93	2	15.530	33.490	44.399	1.326
Aloe chabaudii var. verekeri	1973-2865									
Aloe claviflora	2016-934; 2008-2157	SA	33.70	16.85	16.85	2	19.960	19.338	28.819	1.490
Aloe comptonii	2002-488; 1986-4265	SA	28.50	14.25	14.25	2	7.739	25.654	32.934	1.284
Aloe confusa	1977-5436; 1949-52501	EA	37.39	18.70	18.70	2	25.425	18.614	25.172	1.352
Aloe cremnophila	1981-3493; 1981-847; 1981-857	HOA	68.19	34.10	17.05	4	22.785	14.788	17.698	1.197
Aloe davyana	1973-2542	SA	38.00	19.00	19.00	2	23.887	28.548	43.207	1.514
Aloe dawei	1973-1966; 1973-1967; 1951-35701	EA	71.26	35.63	17.81	4	19.253	18.471	25.356	1.373
Aloe deltoideodonta	2008-2159; 1973-2855	MAD	31.61	15.81	15.81	2	22.705	30.727	40.116	1.306
Aloe deltoideodonta var. candicans	1973-2853					2				
Aloe desertii	1970-1749	EA	37.13	18.56	18.56	2	-	-	-	-
Aloe dorotheae	2003-2532	EA	39.95	19.97	19.97	2	-	-	-	-
Aloe ellenbeckii	1977-3962; 1973-2107	HOA	45.03	22.52	22.52	2	41.906	13.357	18.589	1.392
Aloe erinacea	1978-3757	NAM	26.99	13.49	13.49	2	-	-	-	-
Aloe excelsa	1986-1125	ZAM	33.12	16.56	16.56	2	-	-	-	-
Aloe fero	copenhagen/ not accessioned yet	SA	32.36	16.18	16.18	2	-	-	-	-
Aloe fibrosa	1977-3619	EA	35.98	17.99	17.99	2	-	-	-	-
Aloe fleurentinorum	1979-1476	AP	36.15	18.08	18.08	2	-	-	-	-
Aloe flexilifolia	1990-1811	EA	36.82	18.41	18.41	2	30.682	25.204	32.920	1.306
Aloe forbesii	1967-39304	AP	38.17	19.09	19.09	2	29.873	26.518	36.454	1.375
Aloe framesii	2018-62	SA	34.50	17.25	17.25	2	31.347	15.528	18.324	1.180
Aloe gariepensis	2015-1806; 2013-1204	NAM	33.32	16.66	16.66	2	23.345	12.924	20.446	1.582
Aloe glauca	1973-4058	SA	33.06	16.53	16.53	2	-	-	-	-
Aloe globuligemma	1951-35708	SA	37.48	18.74	18.74	2	-	-	-	-
Aloe grandidentata	1973-2520	SA	37.57	18.78	18.78	2	22.677	16.974	24.690	1.455
Aloe greatheadii	1996-1525	ZAM	37.37	18.68	18.68	2	23.887	18.003	31.634	1.757
Aloe greenii	2006-1113	SA	37.07	18.53	18.53	2	28.045	17.465	24.711	1.415
Aloe hildebrandtii	1981-887	NAM	38.51	19.26	19.26	2	-	-	-	-
Aloe humilis	2002-3308	SA	36.41	18.20	18.20	2	34.967	10.995	16.010	1.456
Aloe inermis	1975-4502	AP	36.51	18.26	18.26	2	-	-	-	-
Aloe jacksonii	1969-342	HOA	72.84	36.42	18.21	4	11.902	13.309	24.497	1.841
Aloe jucunda	1981-779; 1981-749; 1957-63007	HOA	36.11	18.05	18.05	2	13.306	14.660	26.040	1.776
Aloe juvenna	1982-3184	EA	69.25	34.62	17.31	4	23.597	19.289	23.725	1.230
Aloe kedongensis *	1980-2740	EA	67.89	33.95	16.97	4	-	-	-	-
Aloe kulalensis	1948-15901	EA	36.23	18.12	18.12	2	19.898	21.411	31.827	1.486
Aloe lateritia var. graminicola	1973-2058; 1970-1744; 1955-12801	EA	39.10	19.55	19.55	2	24.425	19.267	26.956	1.399
Aloe leachii	1990-1820	EA	38.02	19.01	19.01	2	-	13.888	21.989	1.583
Aloe littoralis	1967-60106	NAM	36.01	18.00	18.00	2	10.069	49.380	61.532	1.246
Aloe lutescens	1973-2869	ZAM	31.69	15.85	15.85	2	-	-	-	-
Aloe macrocarpa	1972-4103	HOA	43.67	21.83	21.83	2	-	-	-	-
Aloe maculata	2006-822	SA	38.71	19.35	19.35	2	19.955	20.639	27.197	1.318
Aloe marlothii	2010-804	ZAM	32.83	16.41	16.41	2	18.269	36.006	56.838	1.579
Aloe mawii	1985-3809	EA	36.10	18.05	18.05	2	-	-	-	-
Aloe mcloughlinii	1984-4966	HOA	37.62	18.81	18.81	2	-	-	-	-
Aloe melanacantha	2015-1807	SA	37.81	18.90	18.90	2	-	-	-	-
Aloe microstigma	1990-1899	SA	30.75	15.38	15.38	2	16.296	20.603	30.844	1.497
Aloe morijensis	1981-2423	EA	38.42	19.21	19.21	2	-	-	-	-
Aloe mubendiensis ~	1973-2845	EA	36.70	18.35	18.35	2	-	-	-	-
Aloe mudenensis	1947-52506	SA	30.54	15.27	15.27	2	-	-	-	-
Aloe musapana ~	1996-2446	ZAM	29.90	14.95	14.95	2	26.571	41.278	60.567	1.467
Aloe mzimbana	1985-3812	EA	35.85	17.92	17.92	2	18.628	13.695	34.720	2.535
Aloe ngongensis	1975-903; 1973-5447	EA	37.23	18.62	18.62	2	14.070	15.720	25.172	1.601
Aloe niebuhriana	1979-2869	AP	35.25	17.62	17.62	2	-	-	-	-
Aloe nyeriensis	1974-4112	EA	66.52	33.26	16.63	4	-	-	-	-
Aloe officinalis	1975-4505	AP	35.44	17.72	17.72	2	2.476	21.796	33.691	1.546
Aloe officinalis var. officinalis	1984-1589									
Aloe pearsonii	2015-2136	SA	28.96	14.48	14.48	2	-	-	-	-
Aloe peckii	1981-1140	HOA	36.62	18.31	18.31	2	-	-	-	-
Aloe pendens	1999-3276; 1977-3258; 1977-2624	AP	33.61	16.80	16.80	2	20.044	21.487	26.233	1.221
Aloe penduliflora	1973-1911	EA	36.01	18.01	18.01	2	-	-	-	-
Aloe pictifolia	1986-1122	SA	32.70	16.35	16.35	2	-	-	-	-
Aloe retrospiciens	1954-63008	HOA	37.97	18.98	18.98	2	-	-	-	-
Aloe reynoldsii	1995-3310	SA	40.42	20.21	20.21	2	-	-	-	-
Aloe rupestris	1958-63020	ZAM	32.45	16.23	16.23	2	-	-	-	-
Aloe sabaea	1973-2559	AP	39.00	19.50	19.50	2	-	-	-	-
Aloe schelpei	1964-42705; 1962-57801	HOA	37.78	18.89	18.89	2	22.533	38.289	58.542	1.529
Aloe scobinifolia	1981-1110	HOA	37.82	18.91	18.91	2	-	-	-	-
Aloe secundiflora	1973-2067	EA	37.07	18.54	18.54	2	-	-	-	-
Aloe simii	2008-2160	SA	36.22	18.11	18.11	2	-	-	-	-
Aloe sinkatana	1987-4087	SUD	37.73	18.87	18.87	2	-	-	-	-
Aloe somaliensis	1981-817	HOA	36.48	18.24	18.24	2	-	-	-	-
Aloe speciosa	1973-3208	SA	30.74	15.37	15.37	2	-	-	-	-
Aloe striata *	1993-1708	SA	41.03	20.51	20.51	2	-	-	-	-
Aloe suffulta	1961-56203	ZAM	41.14	20.57	20.57	2	9.309	16.010	27.583	1.723
Aloe suprafoliata	2003-1825	SA	32.52	16.26	16.26	2	20.128	22.375	33.563	1.500
Aloe swynnertonii	1970-2395	ZAM	39.80	19.90	19.90	2	28.541	18.903	22.375	1.184
Aloe thraskii	1947-29907	SA	34.68	17.34	17.34	2	-	-	-	-
Aloe tomentosa	2003-2489; 1986-1110; 1979-2870	AP	35.27	17.64	17.64	2	13.115	30.091	38.771	1.288
Aloe trichosantha	2015-2405; 2003-2483	HOA	37.74	18.87	18.87	2	11.747	19.996	30.734	1.537
Aloe trichosantha ssp. trichosantha	1972-2771									
Aloe ukambensis	1952-53104	SA	37.60	18.80	18.80	2	-	-	-	-
Aloe vanbalenii	2008-2076	AP	35.47	17.74	17.74	2	-	-	-	-
Aloe vera	1993-1717; 1975-3937; 1969-12338	SA	34.52	17.26	17.26	2	16.357	13.309	23.436	1.761
Aloe verdoorniae	1965-12201	MAD	37.38	18.69	18.69	2	21.664	-	-	-
Aloe viguieri	1980-2088	EA	32.81	16.41	16.41	2	-	-	-	-
Aloe volkensii	1970-1748	SA	38.33	19.17	19.17	2	-	-	-	-
Aloe vryheidensis	2013-1202	AP	31.16	15.58	15.58	2	13.152	39.639	60.374	1.523
Aloe yemenica	1978-3736	ZAM	35.50	17.75	17.75	2	17.419	29.705	38.578	1.299
Aloe zebrina	2007-2178	ZAM	35.80	17.90	17.90	2	16.649	13.984	20.832	1.490
Bulbine frutescens	1973-3211	SA	29.90	14.95	7.48	4	31.14	8.722	8.833	1.013
Kniphofia triangularis	2009-3302	SA	21.89	10.95	10.95	2	17.05	11.222	11.778	1.681

** = sampled as, but labeled differently*
~ = not included in analysis
^ = Specimen donated to RBGK by SANBI
ploidy assumed from GS & ploidy of close rel.

Key to Regions
AP = Arabian Penninsula
EA = Eastern Africa
HOA = Horn of Africa
MAD = Madagascar
NAM = Namibia
SA = South Africa
SUD = Sudan
WA = Western Africa
ZAM = Zambian Region

APPENDIX B

R script for creating a phylogenetic hypothesis with continuous data mapped onto it, using adaxial stomatal density as an example. Output = heatmap

---

```
###SETUP

#setup required packages
install.packages ("phytools", dependencies = TRUE)
install.packages ("picante", dependencies = TRUE)
install.packages ("ape", dependencies = TRUE)
install.packages ("geiger", dependencies = TRUE)
install.packages ("phylotools", dependencies = TRUE)
library (picante)
library (phytools)
library (ape)
library (geiger)
library(phylotools)

#set your working directory
setwd ("~/Desktop/Aloe Analysis")

#upload original tree file
tree <- read.tree(file = "kelda1.tre")
tree
plot(tree, cex = 0.3)

### SD ad analysis

# upload data
SDaddata <- read.csv("SD for R.csv", header = T,row.names = 1)
SDaddata

#format data names to match names in tree
species_SDad <- gsub(" ", "_", rownames(SDaddata))
species_SDad

#set columname according to your dataset
data_SDad <- setNames (SDaddata$Ad.SD, species_SDad)
data_SDad

#this compares both the dataset and the tree (original or pruned) and will
drop the data missing on the tree and also viceversa
test2_SDad <- match.phylo.data(tree, data_SDad)
plot(test2_SDad$phy, cex = 0.8)

#write files for the pruned data and tree
write.csv(data_SDad, file = "pruned_SDad.csv") #check file and add row.names
to the csv
```

```
write.tree(test2_SDad$phy, file = "pruned_SDad.tre") #this tree is the one
you should use from now onwards

data1_SDad <-read.csv("pruned_SDad.csv", header = T, row.names = 1)
data1_SDad

data2_SDad <-setNames(data1_SDad$x,rownames(data1_SDad))
data2_SDad

pruned_tree_SDad <-read.tree(file  = "pruned_SDad.tre")


#compare names in dataset and tree --> list of differences *cannot be any for
the ancestral state reconstruction step
obj_SDad<- name.check(pruned_tree_SDad, data2_SDad) #compares names in
dataset and the tree and creates lists of differences
obj_SDad # should be OK

# this is the ancestral reconstruction under Maximum Likelihood. It provides
value for the nodes and the 95% CI
fit_SDad<-fastAnc(pruned_tree_SDad,data2_SDad,vars=TRUE,CI=TRUE)
fit_SDad
write.csv(fit_SDad$ace, file = "Aloes-ancestral trait reconstuctuion_SDad")
#creates a file with the ancestral reconstruction

#now duplicate node labels and save a new tree. Open this tree in figtree to
see node numbers
pruned_tree_SDad$node.label<-((length(trees$tip)+1:((length(trees$tip)*2)-
1)))


# continuous data mapping on tree, output is a heatmap
obj2_SDad<-contMap(pruned_tree_SDad,data2_SDad,res=1000,plot=FALSE)

#specifiations for the image
plot(obj2_SDad, show.tip.label=FALSE,
res=1000,lwd=1,fsize=c(0.9,0.7),outline=FALSE,legend=0.2*max(nodeHeights(tree
)))

#fsize c(x,y) x=tiplabels and y=legend
#outline=FALSE No black lines framing the colour ones
#type= "phylogram" (the default), "cladogram", "fan", "unrooted", "radial"
#show.tip.label=TRUE/FALSE

#if desired reverse color scale and plot again
obj2$cols[]<-rev(rainbow(1001,start=0, end=0.75))

plot(obj2, show.tip.label=FALSE,res=1000,lwd=1,fsize=c(0.7,0.7),
outline=FALSE,legend=0.2*max(nodeHeights(tree))
```