

**Ecological and Evolutionary Significance of
Crassulacean Acid Metabolism
in the Montane Genus *Puya* (Bromeliaceae)**



Juan David Beltrán

Magdalen College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2017

Table of Contents

List of Figures.....	vii
List of Tables.....	xi
Abstract.....	xiv
Acknowledgements.....	xvi
Chapter 1. General Introduction.....	1
Abstract.....	1
Crassulacean acid metabolism	2
The four phases of CAM.....	2
The C ₃ -CAM Continuum	4
Constitutive CAM vs. constitutive C ₃	5
The diversity of non-constitutive CAM.....	6
Evolution of CAM.....	7
Ecology of CAM	9
Outstanding question in CAM biology	10
Bromeliaceae, a diverse family	11
The genus <i>Puya</i>	12
The concept of climatic niche	16
Hypotheses tested in the present study	17
Chapter 2. Ecology and evolution of crassulacean acid metabolism in the montane genus <i>Puya</i>	19
Abstract.....	19
Introduction.....	20
Materials and Methods	25

Carbon-isotope values	25
Distribution of CAM and C ₃ species.....	28
Climatic data.....	29
Statistical analysis.....	32
Fundamental niche.....	34
Ancestral state reconstruction and phylogenetic signal.....	35
Results.....	36
$\delta^{13}\text{C}$ values in <i>Puya</i>	36
Distribution of CAM and C ₃ <i>Puya</i> species	37
Principal component analysis	39
Climatic differences given the photosynthetic pathway	40
Random forest	42
C ₃ and CAM fundamental niche.....	44
Ancestral state reconstruction and phylogenetic signal.....	48
Discussion.....	50
$\delta^{13}\text{C}$ values in <i>Puya</i>	51
Climatic niche differences between CAM and C ₃ species of <i>Puya</i>	53
Ellenberg's model	56
Heyduk's model.....	57
Topographic differences between CAM and C ₃	58
CAM and C ₃ niche overlapping.....	59
Why CAM species do not occur in the Northern Andes	60
Evolution of CAM in the genus <i>Puya</i>	63
Weak CAM in <i>Puya floccosa</i>	68
Conclusions	70

Chapter 3. Physiological and biochemical adaptations of CAM species to high-elevation environments in the genus <i>Puya</i>	72
Abstract	72
Introduction.....	73
Materials and Methods	77
Plant material and growing conditions.....	77
Cold resistance and freezing damage	78
PEPC activity	80
Temperature optimum of PEPC	81
Thermal stability of PEPC at 50 °C.....	81
Reduction in PEPC activity at 5 °C	82
Statistical analyses	82
Results.....	82
Cold resistance	82
Temperature optimum of PEPC	87
Thermal stability of PEPC at 50 °C.....	88
Relative reduction in PEPC activity at 5 °C.....	93
Discussion.....	96
Cold resistance	97
Cold tolerance or supercooling?	99
Temperature optimum of PEPC	100
Thermal lability of PEPC at high temperature	102
Relative reduction in PEPC activity at low temperature	103
Evolution of <i>Puya</i> and colonisation of the Andes	104
Conclusions	104

Chapter 4. Evolution of phosphoenolpyruvate carboxylase in monocots with emphasis on the origin of CAM in the genus <i>Puya</i>	106
Abstract	106
Introduction.....	107
Materials and Methods	112
Plant material.....	112
RNA extraction.....	112
RNA-seq assemble and relative expression of genes involved in CAM in <i>Puya</i> .	112
<i>Ppc-1</i> sequences in other plants.....	116
Phylogenetic analyses	116
Selection test in genes involved in CAM photosynthesis from eight <i>Puya</i> species	117
Selection test of <i>ppc-1</i> in monocots.....	118
Results.....	119
Expression of genes involved in CAM in <i>Puya</i>	119
Evolution of <i>ppc-1</i> in monocots plants.....	125
Genes involved in CAM evolving under positive selection in <i>Puya</i>	138
Positive selection on amino acid residues of <i>ppc-1</i> in CAM species	140
Discussion.....	146
The significance of relative expression of genes involved in CAM in non-CAM species in <i>Puya</i>	146
The evolutionary implications of positive selection in genes involved in CAM.	148
The evolutionary implications of the polymorphism Arg679 in <i>Puya</i>	148
The significance of evolution of CAM in two independent isoforms of <i>ppc-1</i> in monocots.....	150

Conclusions	151
Chapter 5. General discussion	153
Climatic niche differences between constitutive CAM and non-constitutive CAM species in <i>Puya</i>	153
Limitations.....	154
Evolution of CAM in <i>Puya</i>	156
AFLP Phylogeny limitations	156
Dating the origin of <i>Puya</i>	157
Evolution of CAM in <i>Puya</i>	157
Hybrid origin of CAM in <i>Puya chilensis</i>	159
The colonisation of the Andes by the genus <i>Puya</i>	169
Varadarajan's hypotheses.....	176
The ecological advantage of weak CAM.....	177
Physiological and biochemical adaptations to high-elevation environments in other organisms	178
Summary of the key results	180
Hypothesis 1: Constitutive CAM species and non-constitutive CAM species of <i>Puya</i> have different climatic niches.....	180
Hypothesis 2: CAM evolved multiple times in <i>Puya</i>	181
Hypothesis 3: There are physiological and biochemical differences between low- elevation and high-elevation species of <i>Puya</i>	181
Hypothesis 4: There are differences in the amino acid residues of PEPC between CAM and non-constitutive CAM species of <i>Puya</i>	181
CAM polymorphism.....	181
CAM polymorphism in <i>Puya</i>	182

Conclusions	182
Supplementary Figures.....	184
Supplementary Tables.....	207
Supplementary Material.....	230
References.....	246

List of Figures

Chapter 2

Figure 1. Distribution of <i>Puya</i> species based on the georeferenced location of 433 herbarium records from 147 species.	38
Figure 2. Principal component analysis of 32 climatic and topographic variables listed in Table 2 from 147 species of <i>Puya</i>	40
Figure 3. (A) Aridity index (mm mm^{-1} and (B) cloud cover seasonality (SD %) between C_3 (116 spp.) and CAM (31 spp.) species of <i>Puya</i>	41
Figure 4. Mean permutation importance of 32 climatic variables in 50 independent Random Forest ordinations in classifying the photosynthetic pathway of 147 species of <i>Puya</i>	43
Figure 5. Predicted suitable areas according to a MaxEnt model (32 variables) for CAM species of <i>Puya</i> (A and B) and C_3 species of <i>Puya</i> (C and D).	45
Figure 6. Differences between the predicted fundamental niche of CAM and C_3 species of <i>Puya</i> based on MaxEnt model (32 variables).	47
Figure 7. Maximum likelihood ancestral state reconstruction of the photosynthetic pathway in <i>Puya</i> based on the relationships supported by posterior probabilities obtained from the Bayesian analysis of AFLP data.	49
Supplementary Figure 1. Frequency histogram of $\delta^{13}\text{C}$ values of 161 species of <i>Puya</i> plotted in class intervals of 1.0‰.	184
Supplementary Figure 2. Correlation between $\delta^{13}\text{C}$ values and elevation in 147 species of <i>Puya</i>	185
Supplementary Figure 3. Correlation between $\delta^{13}\text{C}$ value and nitrogen content in 52 species of <i>Puya</i>	186

Supplementary Figure 4. Mean permutation importance of 32 climatic variables in 50 independent Random Forest ordinations to classify the photosynthetic pathway of 110 species of <i>Puya</i> from central and southern Andes and Chile.....	187
Supplementary Figure 5. Mean regression importance of 32 climatic variables to model elevation in 50 independent Random Forest regressions of 31 CAM species of <i>Puya</i>	188
Supplementary Figure 6. Bayesian inference phylogeny of 69 species of <i>Puya</i> based on AFLP information from Jabaily and Sytsma (2013).....	189
Supplementary Figure 7. Distribution of <i>Puya alpestris</i> and <i>Puya chilensis</i> by their carbon isotope values.....	190
Supplementary Figure 8. Linear correlation between aridity index, precipitation seasonality latitude and $\delta^{13}\text{C}$, the carbon isotope values and climatic variables of 5 species (51 individuals) of <i>Puya</i> from Chile.....	191

Chapter 3

Figure 1. Temperature 50% maximum effect (T_{50}) for four species of <i>Puya</i> (including both high- and low-elevation species) and five species from the subfamily Bromelioideae (both equatorial and non-equatorial) in a common garden experiment.	84
Figure 2. Temperature curves of PEPC activity recorded at 3.0 mM PEP activity of four species of <i>Puya</i>	88
Figure 3A. Summary of the percentage of the residual PEPC activity remaining after incubation for various times at 50 °C of extracts from four species of <i>Puya</i> and three species of the subfamily Bromelioideae.....	91

Figure 3B. Summary of the percentage of the residual PEPC activity remaining after incubation for 5 minutes at 50 °C of four species of <i>Puya</i> and three species of subfamily Bromelioideae.	91
Figure 4. Low-temperature activity of PEPC, measured at 5°C, for four species of <i>Puya</i> and three species of the subfamily Bromelioideae, expressed as a percentage of the activity recorded in the same extract at 25 °C.	94
Supplementary Figure 1. Optimisation of the electrolyte leakage protocol followed in the present study.	192
Supplementary Figure 2. Quasi-binomial logistic regression of relative electrolyte leakage and temperature.	193

Chapter 4

Figure 1. Ratio between the expression (TPM) of <i>ppc-1a</i> and <i>ppc-1b</i> isoforms in eight species of <i>Puya</i>	121
Figure 2. Maximum parsimony (MP) trees showing the evolution of <i>ppc-1</i> genes in monocots.	128
Figure 3. Neighbour-joining (NJ) trees showing the evolution of <i>ppc-1</i> genes in monocots.	131
Figure 4. Maximum likelihood (ML) trees showing the evolution of <i>ppc-1</i> genes in monocots.	134
Figure 5. Bayesian inference (BI) trees showing the evolution of <i>ppc-1</i> genes in monocots.	137
Figure 6. Fragment of the alignment of the <i>ppc-1</i> sequences in the order Asparagales and Poales.	145

Chapter 5

Figure 1. Relationships between the four climatic variables that discriminate more between the 31 constitutive CAM and the 116 non-constitutive CAM species in <i>Puya</i> found in Chapter 2.....	154
Figure 2. Ancestral state reconstruction by maximum parsimony of the CAM polymorphism in the residue 679 in the highest expressed PEPC isoform of eight species of <i>Puya</i>	159
Figure 3. Schematic representation of the origin of CAM in <i>Puya chilensis</i> by hybridisation with <i>P. alpestris</i> ssp. <i>zoellneri</i> and its consequences for the genome evolution.	162
Figure 4. Bayesian inference (BI) trees showing the evolution of <i>ppc-1</i> genes in <i>Puya</i> using nucleotide alignment.	165
Figure 5. Synthesis of the thesis.	172

List of Tables

Chapter 1

Supplementary Table 1. List of the taxonomic names mentioned in this thesis and the accepted taxonomic names of <i>Puya</i> species.....	207
---	-----

Chapter 2

Table 1. $\delta^{13}\text{C}$ values used in the present study from 51 species (107 individuals) of <i>Puya</i> in addition to the carbon isotope values reported in Crayn <i>et al.</i> (2015).	26
--	----

Table 2. List of 32 climatic and topographic variables used in the present study and their source.....	31
---	----

Supplementary Table 1. List of 32 climatic and topographic variables used in the present study and their source.	219
--	-----

Supplementary Table 2. Differences between CAM and C ₃ species in their mean, median, minimum and maximum values based on bootstrap test of 100 000 independent iterations. <i>p</i> -values indicates significant differences.....	220
---	-----

Supplementary Table 3. PCA eigenvalues of the climatic and topographic variables for PC1, PC2 and PC3.	221
--	-----

Supplementary Table 4. Accuracy of five different ordinations to classify CAM and C ₃ species using climatic data and random forest ordinations.	222
---	-----

Chapter 3

Table 1. Summary of the 50% maximum effect temperature (T_{50}) for each species. Plants were maintained for a minimum of 2 months in a common garden experiment.	85
---	----

Table 2. Summary of the Tukey HSD pairwise differences in T_{50} values between species.	86
Table 3. Summary of the Tukey HSD pairwise differences in T_{50} values between equatorial bromelioid species, low-elevation <i>Puya</i> species and high-elevation <i>Puya</i> species.	87
Table 4. Summary of the <i>post hoc</i> analysis using Tukey-adjusted pairwise comparisons of the residual PEPC activity after incubation at 50 °C for 5 minutes.....	92
Table 5. Summary of the <i>post hoc</i> analysis using Tukey-adjusted pairwise comparisons of the residual PEPC maximum activity after incubation at 50 °C.	93
Table 6. Summary of the Tukey HSD pairwise differences in the relative PEPC activity at 5°C compared with 25°C between species.....	95
Table 7. Summary of the Tukey HSD pairwise differences in the relative PEPC activity at 5°C compared with 25°C between equatorial bromelioid species, low-elevation <i>Puya</i> species and high-elevation <i>Puya</i> species.	96
Supplementary Table 1. Summary of the PEPC temperature optimum at saturated conditions based on the smooth function fitting a polynomial surface for each species.	223
Supplementary Table 2. ANOVA results from the linear mixed-effect model based on residual PEPC activity after incubation at 50 °C, time and species.....	223
Supplementary Table 3. ANOVA results from the linear mixed-effect model based on residual PEPC activity after incubation at 50 °C, time and species groups.....	224
Supplementary Table 4. Summary of the residual PEPC $V_{\max}$ activity (%) at 5 °C.	224

Chapter 4

Table 1. List of the 25 genes involved in CAM photosynthesis included in the present study.	115
Table 3. Summary of the ANOVA test of the differences in the gene expression of CAM and C ₃ species of <i>Puya</i>	125
Table 4. Summary of genes involved in CAM evolving under positive selection in the genus <i>Puya</i>	139
Table 5. Comparison between amino acid substitution models to identify residues under positive selection in the most expressed <i>ppc-1</i> isoform of 11 monocots species (5 CAM species; 6 C ₃ species).	141
Table 6. Summary of amino acid replacements under positive selection in the CAM lineages of monocots.	143
Supplementary Table 1. Summary of all the <i>ppc-1</i> genes used in the present study. In total 35 species; representing 7 orders and 13 families.	225
Supplementary Table 2. Isoforms of 25 genes involved in CAM photosynthesis in the eight species of <i>Puya</i>	227

Abstract

Little is known about the evolution and ecology of crassulacean acid metabolism (CAM) in the genus *Puya* Molina. CAM is a photosynthetic pathway typified by nocturnal CO₂ fixation and is regarded as a water-saving mechanism. *Puya* is one of the largest genera in the pineapple family (Bromeliaceae), with 226 species distributed across the Andes to Costa Rica and the Guiana Shield, and from sea level to 5000 m. About 21% of *Puya* species are CAM and at least 10 of these CAM species occur above 3000 m. The main aim of this thesis was to uncover new evidence to understand the ecophysiology and evolution of CAM in the montane genus *Puya*. The prevalence of CAM and C₃ species in *Puya* was estimated from carbon isotope values of 161 species. The climatic niche of constitutive CAM species and C₃ species of *Puya* was modelled using georeferenced herbarium records and climatic variables to evaluate the differences between their niches. The evolution of CAM in *Puya* was investigated by reconstructing the ancestral photosynthetic pathway on an AFLP phylogeny and by studying positive selection in the genes encoding the key enzyme phosphoenolpyruvate carboxylase (PEPC). The cold-resistance and the thermal lability of PEPC was investigated for high- and low- elevation CAM species of *Puya* to explore the potential molecular adaptations of CAM plants in high-elevation environments. The present study concludes that the common ancestor of *Puya* was a cold-resistant plant. This is suggested to explain the prevalence of *Puya* at high-elevations. The evolution of CAM was correlated with changes in the climatic niche, and occurred multiple times in *Puya*. These multiple origins were not independent because the common ancestor of *Puya* was likely to be a weak CAM plant (based on a diagnostic Arg679 residue in the PEPC sequence). It is likely that populations of *P. chilensis* gained CAM by introgression with *P. alpestris* ssp. *zoellneri*. Weak CAM photosynthesis and cold-

resistance allowed *Puya* to colonise the Andes from the south to the north; and, in the process, constitutive CAM and C₃ evolved. The later-evolving species in the genus are suggested to have lost their capacity for CAM as they radiated into more mesic habitats during their colonisation of the northern Andes.

Acknowledgements

First of all, I would like to thank my principal supervisor Professor Andrew Smith who always has time to discuss about Bromeliaceae and the ecology, evolution and physiology of CAM. His enthusiasm of CAM is contagious and I am proud of have been his student. I am very grateful to him because he taught me to be a more rigorous and precise scientist. I would also like to thank my second supervisor Dr Stephen Harris who helped immensely to the construction of this thesis. During our regular discussions about ecology and evolution he always provided guidance and helped me to be more succinct and direct. I would like to thank to Professor Paul Jarvis, who allowed me to use his spectrophotometer in the cold room. I am very grateful to Professor Lee Sweetlove and Professor Nick Kruger for their valuable comments during the assessment stage.

I am very grateful to the Louis-Dreyfus foundation, the Weidenfeld-Hoffmann Trust and the Colombian government (Colciencias) for their generosity funding my thesis. Without them, this thesis would not have been possible. I am grateful to Magdalen College throughout my DPhil.

I thank the University of Oxford Botanic Garden and The Ventnor Botanic garden for providing the plant material to complete this thesis. Similarly, to the Missouri Botanical Gardens Herbarium, Royal Botanic Gardens Kew Herbarium and the *Herbario Nacional Colombiano* for allowing me to use their collections to complete this thesis.

I want to thank to Dr Úrsula Flores-Pérez, Dr Qihua Ling, Dr Bruno Nevado and Dr Judith Paulus who helped me at different stages during my DPhil. I am very grateful to Dr Úrsula Flores-Pérez, who taught me all the procedures related with bacterial transformation and protein purification. Thanks to Pedro Bota for his help in the laboratory and thanks to Tim Wroe for looking after the maintenance of the green house.

Thanks to the people in the Plant Sciences Department who made my DPhil a very enjoyable experience such as Sam, Lottie, Gem, Rita, Claudia, Jacob and Marce.

This thesis would never have been completed without the love and support of my parents Elsa and Pablo Emilio, and my siblings Catalina and Pablo. They are always in my heart, and I hope every day they will be proud of me. I am very thankful to Mark Tushingham, for being with me, especially during complicated moments, he always points out the bright side of any situation, brightening up my life. I am very thankful to my friends Gaby, Paulis and Fay who were always there during my time in Oxford. Finally, to anyone who contributed to making my time in Oxford enjoyable.

Chapter 1. General Introduction

Abstract

Understanding the origin of complex traits, and the ecological benefits of species exhibiting these traits, have been fundamental questions in biology. One remarkable complex physiological trait found in plants is crassulacean acid metabolism (CAM). CAM is a photosynthetic pathway whereby plants fix the majority of CO₂ into malic acid at night-time using the enzyme phosphoenolpyruvate carboxylase (PEPC). CAM is regarded as a water-saving mechanism. At least 6% of flowering plants are CAM species and this remarkable mechanism has evolved more than 60 times independently from C₃ ancestors.

The pineapple family (Bromeliaceae) contains both CAM and C₃ species and so represents an excellent model to understand the ecological and evolutionary significance of CAM. The family contains 3547 species in 76 genera. However, the majority of these genera contain either all CAM or all C₃ species. There are two relatively large genera that contain both CAM species and C₃ species: the polyphyletic and mainly epiphytic genus *Tillandsia*; and the monophyletic and montane genus *Puya*. *Puya* is one of the largest genera in the family with 226 species distributed across the Andes to Costa Rica and the Guiana Shield, and from sea level up to 5000 m. About 21% of *Puya* species are CAM and at least 10 of these CAM species occur above 3000 m. The main aim of this thesis is to uncover new evidence to understand the ecophysiology and evolution of CAM in the montane genus *Puya*.

The present chapter is divided in four sections. The first section presents a general introduction to CAM photosynthesis, with an emphasis on the concept of weak CAM. The second section reviews the literature about the family Bromeliaceae, with an emphasis on the genus *Puya*. The third section clarifies the niche concept that was adopted

in this thesis. Finally, the fourth section presents the general hypotheses of this thesis. Supplementary Table 1. lists all the taxonomic names mentioned in this thesis, including a list of accepted species names of *Puya*.

Crassulacean acid metabolism

The four phases of CAM

CAM is a modification of the basic C_3 photosynthetic pathway in which the majority of CO_2 assimilated from the atmosphere is fixed initially during the night-time by the enzyme phosphoenolpyruvate carboxylase (PEPC, EC 4.1.1.31; Osmond, 1978). Osmond (1978) demonstrated that the typical CAM cycle of *Kalanchoe daigremontiana* can be divided into four distinct physiological phases, based on gas exchange and malic acid accumulation. The four-phase model of CAM has been extensively reviewed (e.g. Osmond, 2007; Silvera *et al.*, 2010a; Borland *et al.*, 2011). This definition of the diel phases of CAM is convenient for establishing a general definition of CAM and for understanding the differential expression of CAM in different species (Dodd *et al.*, 2002).

Phase I of the CAM cycle is characterised by nocturnal CO_2 uptake. During this phase, the stomatal conductance is high, and the assimilation of atmospheric and respiratory CO_2 takes place. The atmospheric CO_2 is converted into HCO_3^- by carbonic anhydrase (CA). Then, HCO_3^- in the presence of phosphoenolpyruvate (a product of glycolysis) is assimilated into oxaloacetate (OAA) by PEPC. OAA is reduced to malate by NAD(P)-malate dehydrogenase (MDH). Subsequently, the malate is transported to the vacuole where it is accumulated as malic acid. At night-time, PEPC is phosphorylated by phosphoenolpyruvate carboxylase kinase (PPCK), which is regulated via the circadian clock (Hartwell *et al.*, 1996, 1999, 2005a,b).

Phase II occurs at the beginning of the day. It is characterised by a transiently increased rate of net CO₂ uptake. This is a direct consequence of the transition from CO₂ fixation by PEPC to ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) (Buchanan-Bollig and Smith, 1983).

During phase III, the stomata close and the majority of accumulated malate molecules are transported out of the vacuole to be decarboxylated in the cytoplasm. This process can occur by two different pathways (Smith and Bryce 1992; Christopher and Holtum 1996; Holtum *et al.* 2005; Silvera *et al.*, 2010a). One pathway involves the oxidation of malate into OAA by malate dehydrogenase (MDH) and its subsequent decarboxylation by PEP carboxykinase (PEPCK), yielding CO₂ and phosphoenolpyruvate (PEP). In the other pathway, malate is decarboxylated by NAD⁺- and/or NADP⁺-malic enzyme (ME) yielding CO₂ and pyruvate. The internal partial pressure of CO₂ increases as more molecules of malate are decarboxylated and the stomata remain closed until decarboxylation of the nocturnally accumulated malate is completed.

Finally, toward the end of the light period once malate decarboxylation is completed, the stomata may open again in phase IV if environmental conditions (such as water availability) permit. During this phase, direct fixation of atmospheric CO₂ occurs via RuBisCO starts and ceases with the onset of the following dark period.

The orchestration of several enzymes at different times shows why CAM is considered a complex physiological trait (Ceusters *et al.*, 2014). As a consequence of confining stomatal opening largely to the night-time, and of using PEPC for the initial capture of atmospheric CO₂, CAM plants show very high water-use efficiencies (Nobel and Hartsock, 1986, Winter *et al.*, 2015). These characteristics reveal the fundamental adaptive value of CAM photosynthesis in plants (Griffiths, 1988).

The phases described above are highly modulated by environmental conditions, in particular phases II and IV (Lüttge, 2004). For example, phase IV is reduced or even eliminated during drought-stress; while phases II and III are strongly influenced by light and temperature conditions (Smith and Lüttge, 1985; Griffiths 1988). Indeed, the contribution of nocturnal CO₂ fixation to the total daily carbon gain could vary between 100% and close 0% in CAM plants (Winter & Holtum, 2002, 2014; Winter *et al.*, 2015). The nocturnal CO₂ contribution to the total daily carbon gain also depends on the developmental stage and the particular species concerned (Silvera *et al.*, 2010a; Winter *et al.*, 2015). This huge variation makes CAM photosynthesis one of the best examples of phenotypic diversity and plasticity in plants (Dodd *et al.*, 2002; Silvera *et al.*, 2010a; Winter *et al.*, 2015).

The C₃–CAM Continuum

The diversity of CAM expression is better viewed as a continuum rather than a binary character of presence-absence. This has been recognised since Cockburn (1985) described several variations within CAM (Dodd *et al.*, 2002; Silvera *et al.*, 2010a, Winter and Holtum, 2014; Winter *et al.*, 2015). Winter *et al.* (2015) divided the continuum into three categories: constitutive C₃ plants; constitutive CAM plants; and any expression in between into the categories called C₃–CAM.

There is a correlation between the degree of succulence and CAM expression (Silvera *et al.*, 2010a). In the excellent review by Males (2017) about succulence in plants, he pointed out that it is very difficult to separate between CAM photosynthesis and succulence, due to their intimate relationship. However, not all succulent plants are CAM plants (Males, 2017). CAM plants need large vacuoles to accumulate the organic acids and this makes the tissues of CAM plants succulent (Barrera Zambrano *et al.*, 2014). It has been suggested that densely packed cells contribute to the efficiency of CAM by

reducing the possible leakage of CO₂, which is a product of the decarboxylation in phase III (Maxwell *et al.*, 1997; Nelson *et al.*, 2005; Nelson and Sage, 2008). Recent phylogenetic studies have demonstrated that at least in the sub family Agavoideae (Asparagaceae) the evolution of succulence pre-dated the origin of CAM (e.g. Heyduk *et al.*, 2016a).

Constitutive CAM vs. constitutive C₃

As mentioned above, PEPC is the key enzyme in CAM photosynthesis because it is responsible for fixing CO₂ at night-time. PEPC discriminates less against the heavier ¹³C isotope than does RuBisCO (Bender *et al.*, 1973; O'Leary and Osmond, 1980; Farquhar *et al.*, 1989; Chollet *et al.*, 1996). For this reason, it is possible to discriminate between CAM and C₃ species by using δ¹³C values (Osmond *et al.*, 1973). Moreover, the δ¹³C values are indicative of the relative contribution of nocturnal CO₂ to the total diel carbon gain. Less-negative values of δ¹³C indicate higher nocturnal CO₂ contribution (Winter and Holtum, 2002). Winter *et al.* (2015) suggested that, in view of the great variation of CAM, constitutive CAM plants should be defined by the amount of CO₂ fixed by PEPC, as reflected by δ¹³C values. Given the strong bimodal distribution of δ¹³C values found in C₃ and CAM plants (for example, but not only, in Bromeliaceae; Crayn *et al.*, 2015), Winter *et al.* (2015) suggested that plants with a δ¹³C value less negative than -20 ‰ should be classified as constitutive CAM plants.

Contrary to the views expressed in the majority of biology textbooks (for example Gurevitch *et al.*, 2006; Jones *et al.*, 2013) constitutive C₃ plants can open their stomata at night time and also fix nocturnal CO₂. Kunitake *et al.* (1959) demonstrated that excised tobacco leaves (*Nicotiana tabacum*) could fix nocturnal CO₂ by exposing them to ¹⁴CO₂. However, constitutive C₃ plants cannot achieve net fixation of CO₂ at night and accumulate the ensuing products such as malate in the vacuole. This biochemical capacity for nocturnal CO₂ fixation is not surprising given that all the enzymes necessary

for CAM (e.g. carbonic anhydrase, PEPC, malate dehydrogenase) are already present in C₃ plants. During the evolution of CAM, these enzymes and the regulatory genes should change to orchestrate the CAM cycle as described above (Borland *et al.*, 2016).

The diversity of non-constitutive CAM

The diversity of CAM expression is mirrored by the diversity of terminology to describe CAM expression. For example, Sipes and Ting (1985) illustrated the concept of CAM idling by studying the CO₂ exchange of *Peperomia camptotricha* in different water regimes and different developmental stages. CAM idling is characterised by total closure of the stomata during day and night, and by the nocturnal fixation of respiratory CO₂ by PEPC. Plants that show CAM idling typically exhibit small day–night differences in tissue titratable acidity

The other type of CAM expression is CAM cycling (Sipes and Ting, 1985), which is characterised by nocturnal re-fixation of respiratory CO₂ into organic acids while the stomata remain closed. The stomata remain opened during daytime and C₃ photosynthesis is exhibited during that period. CAM cycling is also described as weak CAM. Weak CAM has been considered to be an evolutionary intermediate between C₃ and CAM (Smith *et al.*, 1986; Silvera *et al.*, 2009, 2010a; Beltrán *et al.*, 2013; Winter and Holtum, 2014). Silvera *et al.* (2014) demonstrated that during the evolution of the tribe Oncidiinae (Orchidaceae) weak CAM species appeared to occupy an intermediate phylogenetic position between constitutive CAM and constitutive C₃ species. Until the present study, the only way to identify weak CAM plants was by measuring the day/night titratable acidity, $\delta^{13}\text{C}$ values and diel CO₂ exchange.

The expression of CAM described above can be modulated by the environment. The most remarkable example is facultative CAM plants (Winter and Holtum, 2014). For example, the ice plant *Mesembryanthemum crystallinum* can be induced to switch from C₃ to

CAM by salt stress, drought and age (Winter and von Willert, 1972; Winter and Holtum, 2014). The bromeliad *Guzmania monostachia* can show some CAM photosynthesis under drought and light stress (Medina *et al.*, 1977, Medina 1979, Griffiths and Smith, 1983; Maxwell *et al.*, 1997, 2002; Pierce *et al.*, 2002a, Beltrán *et al.*, 2013). The ecological relevance of facultative CAM is associated with environments which naturally show a temporal variation (e.g. seasonal) in the parameters that ultimately trigger the CAM expression in facultative CAM species (Borland *et al.*, 2011; Winter and Holtum, 2014). It is likely that facultative CAM is not an intermediate state between C₃ and CAM (Winter *et al.*, 2015).

The carbon isotope values of plants which exhibit CAM-cycling, CAM-idling, facultative CAM or weak CAM plants are usually more negative than −20 ‰ (Pierce *et al.*, 2002; Silvera *et al.*, 2005; Winter and Holtum, 2014). Given this level of CAM diversity, Winter *et al.* (2015) proposed a definition for all non-constitutive CAM types and grouped them all into a category called C₃–CAM species. This category is problematic because it joins facultative CAM with weak CAM species. These two mechanisms are physiologically very different. Some weak CAM plants exhibit nocturnal acid accumulation even in well-watered conditions (Silvera *et al.*, 2014) and they are therefore not facultative. More importantly, it is likely that weak CAM and facultative CAM evolved from different trajectories. For that reason, both terms weak CAM and facultative CAM should be preferred over the umbrella term C₃–CAM. It is clear that by measuring the $\delta^{13}\text{C}$ values, the bimodal distribution shows constitutive *vs.* non-constitutive CAM.

Evolution of CAM

CAM photosynthesis has been recorded in 33 families of vascular plants (Winter & Smith, 1996) and 6% of flowering plants are estimated to be CAM species (Silvera *et al.*, 2010a). CAM evolved in more than 60 independent lineages, which are comparable

to the number of origins of C₄ photosynthesis (Edwards & Ogburn, 2012). CAM evolved multiple times independently from C₃ ancestors (Edwards & Ogburn, 2012). There is currently much interest in understanding how such complex traits arise and the intermediate steps in their evolution. However, the studies of CAM evolution have been performed in a phylogenetic context only considering CAM *vs.* C₃ by measuring the $\delta^{13}\text{C}$ values. As mentioned above, this division only shows constitutive *vs.* non-constitutive CAM. For example, CAM evolution has been studied in the families Orchidaceae (Silvera *et al.* 2009, 2010a, 2014), Asparagaceae (Agavoideae; Heyduk *et al.*, 2016), Cactaceae (Edwards and Donoghue, 2006), Clusiaceae (Gehrig *et al.*, 2003), Crassulaceae (Mort *et al.*, 2001), Euphorbiaceae (Horn *et al.*, 2014) and Bromeliaceae (Crayn *et al.*, 2004, 2015; Givnish *et al.*, 2011, 2014).

Some lineages exhibit phylogenetic niche conservatism at the family level: almost all their members show CAM activity to some extent, such as in the cacti (Edwards *et al.*, 2005; Edwards and Diaz, 2006) and agaves (Nobel, 1988). However, in other families containing large numbers of CAM species, these may represent only a proportion of the total species, such as in Bromeliaceae (43% CAM species: Crayn *et al.*, 2015) and Orchidaceae (~15% CAM: Silvera *et al.*, 2009, 2010a,b). However, at finer taxonomic scales (species level and below) it is still unknown how ‘labile’ CAM has been as a trait, i.e. whether CAM has evolved multiple times within individual genera or populations, or whether CAM can be lost as well as gained. This question has been difficult to address principally because of a lack of sufficiently detailed information about both the photosynthetic pathway and phylogenetic relationships within clades containing large numbers of CAM and C₃ species. So far, the most complete information about the photosynthetic pathways at the family level has come from Crayn *et al.* (2004, 2015) for Bromeliaceae, in which just over half of the 3547 spp. have been sampled for their $\delta^{13}\text{C}$

values. However, the most intensively sampled family-wide phylogeny for Bromeliaceae contains only 90 spp. (Givnish *et al.*, 2011, 2014).

Ecology of CAM

CAM plants are found predominately in semi-deserts, Mediterranean climates, dry tropical forests and dry microclimates such as epiphytic habitats in tropical rain forests (Winter, 1985; Smith *et al.*, 1986; Silvera *et al.*, 2010a, b). This is congruent with CAM as a water-saving mechanism. Even though the habitat has been described in terms of the ecosystems where the CAM plants occur (Lüttge, 2004, 2010), there is still no formal definition of the climatic niche of CAM plants in comparison with their closely related C₃ species. A formal definition will increase our understanding of whether CAM (as a complex physiological trait) has conferred ecological advantages over their close relative C₃ sister-species. This topic will be explored in Chapter 2.

Another relevant ecological parameter is the correlation between CAM and elevation. The frequency of CAM species tends to decrease with increasing elevation. This phenomenon has been observed in epiphytic bromeliads from low elevations to cloud forests in the tropics (Griffiths and Smith, 1983; Pierce *et al.*, 2002b). Earnshaw *et al.* (1987) also showed a similar pattern for epiphytic orchids where CAM species were not found in the tropical alpine environment in Papua New Guinea. Two reasons have been given to explain this pattern. One is entirely evolutionary: CAM species colonised the cloud forests relatively recently and have therefore not yet radiated (Pierce *et al.*, 2002b). The second explanation is more ecological: cloud forests are more humid than lower elevations and this implies that CAM could be disadvantageous rather than advantageous (Earnshaw *et al.*, 1987).

However, terrestrial CAM species have been found in higher elevations in the Andes. There are a few remarkable examples of CAM plants occurring at altitudes above

4000 m (e.g. the cacti *Oroya peruviana* and *Tephrocactus floccosus* (Keeley and Keeley, 1989), and in Crassulaceae *Echeveria quitensis* (Medina & Delgado, 1976). These environments experience large day-night temperature changes of typically 20 °C and potentially as great as 40 °C, and during the night the air temperature usually drops to a minimum value below 0 °C (Rundel *et al.*, 1994). This is intriguing, because in typical CAM plants most of the CO₂ fixation (via PEPC) occurs at night, but no research has been performed on high-elevation CAM species to determine whether their enzymes possess special adaptations to maintain activity at low temperatures. This topic is explored in depth in Chapter 3.

Outstanding question in CAM biology

It is important to evaluate the climatic niche differences between constitutive CAM and non-constitutive CAM species in a single monophyletic clade. This will allow us to understand whether constitutive CAM could imply a change in the climatic niche as a consequence of its adaptive value. As mentioned above, another outstanding question in CAM biology is whether CAM has evolved multiple times within individual genera or even at population level. If there are differences in the climatic niche, using a phylogenetic reconstruction it is possible to understand whether the evolution of CAM was correlated with a change in the climatic niche. Another outstanding question is the potential physiological and biochemical differences between low-elevation and high-elevation CAM species. This topic had not been explored in the Andes and potentially could open a new research field for molecular adaptations of CAM to low temperatures. Finally, it is not clear whether there are amino acid residues which are unique for CAM species in the sequence of PEPC. The quest for these residues will add an additional marker that could discriminate between constitutive CAM, constitutive C₃ and weak CAM. This would

contribute to our understanding of the evolution of weak CAM and constitutive CAM on a large scale. These issues are all explored in Chapter 4.

To tackle these questions, it is necessary to find a model which contains constitutive and non-constitutive CAM species across a wide range of elevations and with sufficiently detailed information about both of the photosynthetic pathways and phylogenetic relationships. The present study considered that the best model to study these questions was the Neotropical family Bromeliaceae and in particular the montane genus *Puya*.

Bromeliaceae, a diverse family

Bromeliaceae is the fifth-largest family of monocots and is considered one of the most diverse Neotropical angiosperm families, with 3547 species and 76 genera (Benzing, 2000; Givnish *et al.*, 2011, 2014; Luther, 2012; Gouda *et al.*, 2016). Bromeliaceae species are found in a vast array of environments, from sea level to 5000 m, and across virtually all tropical and subtropical parts of the Americas (Smith and Downs, 1974, 1977, 1979). Bromeliads may be terrestrial, epiphytic, lithophytic, or even carnivorous plants (Benzing, 2000). They are perennial herbs, with spirally arranged rosulate leaves, often dilated at the base and forming a water-impounding ‘tank’ or phytotelma (Tomlinson, 1970). The shoots characteristically bear multicellular peltate trichomes, which in the epiphytic forms play an important role in water and nutrient absorption. In the extreme epiphytic forms in the genus *Tillandsia* L., the shoots are covered with a dense indumentum of absorptive trichomes and the root system is essentially reduced completely to the role of a holdfast (Smith, 1989; Benzing, 2000). Unfortunately, there is only one possible fossil record for the family, but there are uncertainties surrounding its age (Baresch *et al.* 2011). According to the most recent rate-calibrated phylogenies of the family, the crown group radiation of

modern representatives in the family is estimated to have begun about 20 Ma (Givnish *et al.*, 2011, 2014).

Within Bromeliaceae, 43% of the species are CAM plants, and this trait has evolved at least five times independently (Crayn *et al.*, 2004, 2015; Givnish *et al.*, 2014). The majority of the genera in the family are consistently CAM or C₃, reflecting a pattern of ecological niche conservatism at this taxonomic level (Givnish *et al.*, 2014). However, there are two major genera containing both C₃ and CAM species in the family: the polyphyletic genus *Tillandsia* and the monophyletic and montane genus *Puya*. In fact, 21% of *Puya* species may be CAM plants according to a $\delta^{13}\text{C}$ survey of 132 species in the genus (Crayn *et al.*, 2015). Surprisingly, this includes at least 10 species at high altitudes (> 3000 m) with $\delta^{13}\text{C}$ values indicative of constitutive CAM. *Puya* fulfil all the conditions to be considered as a model genus to test the outstanding questions in CAM biology mentioned above.

The genus *Puya*

Puya (Molina) is a monophyletic genus that provides a good model for studying Andean biogeography (Varadarajan, 1990; Jabaily and Sytsma, 2010, 2013). *Puya* is a relatively large, predominantly montane, genus with 226 species distributed across the Andes, central Chile, the Guiana shield and Talamancan paramo in Costa Rica (the distribution of *Puya* is reviewed in detail in Chapter 2). *Puya* has the widest elevation range of the family, from sea level up to 5000 m. (Luther, 2012; Jabaily and Sytsma, 2013). Although a few species extend to sea level, the genus is predominantly montane. Varadarajan (1990) proposed three main hypotheses about the biogeography of *Puya*. (1) *Puya* originated in the Guiana Shield and colonised the Andes from the north to the south. (2) Changes in the environment of *Puya* in the central Andes (from semi-xeric to xeric environments) triggered the evolution of CAM. (3) The colonisation of the Andes by *Puya*

was mainly “horizontal” because the common ancestor was a high-elevation species. Unfortunately, these hypotheses were based on two erroneous assumptions: (1) that *Puya* belongs to the subfamily Pitcairnioideae; and (2) that the origin of *Puya* was in the Guiana Shield (Varadarajan, 1990). The molecular phylogenies of the family and the genus show that these assumptions are wrong (Crayn *et al.*, 2004; Jabaily and Sytsma, 2010; 2013). However, the third hypothesis about the relationship between the climatic niche and the evolution of CAM has not yet been tested.

The tropical highlands are dynamic environments, where air temperature can vary from -10°C at night to 30°C or even higher during the day (Luteyn, 1999), where the wind speeds are higher than in the valleys, and where water availability may be limited, even in places with considerable precipitation, as the water is not necessarily available to the plant because it could be frozen (Leuschner, 2000). It has been claimed by Rundel *et al.* (1994) that *Puya* giant rosettes could have physical mechanisms to avoid sub-zero temperatures similar to *Espeletia* (Goldstein *et al.*, 1985; Rada *et al.*, 1985). For instance, *Espeletia* have their entire vegetative and reproductive body covered by a conspicuous indumentum (Luteyn, 1999) and also the dead leaves are still attached to provide insulation for the stem (Goldstein *et al.*, 1985). In contrast, the leaves of the majority of *Puya* species are usually glabrous on the adaxial surface, and with little or no indumenta on the abaxial surface. Some *Puya* species have conspicuous indumenta on their inflorescences, and this is only found in high-altitude species to protect the flowers and improve the seed set (Miller, 1986). To date, there are no studies comparing the ecophysiology of high- and low-elevation species of *Puya*.

Puya species are perennial, rosette-forming herbs with leaves that have coarsely spinose-serrate leaf margins. Their inflorescence are simple or compound, and the flowers are showy, with free sepals and petals, although the sepals are much shorter than the

petals. At anthesis, the sepals are spreading and then twisted, a characteristic that distinguishes *Puya* from all other bromeliads. The filaments are free; the ovary is always superior and glabrous; and the style is slender and elongated. The fruit is a loculicidal capsule; the seeds are broad and alate (Smith and Downs, 1974). Traditionally, the genus *Puya* has been split into two subgenera: *Puya* and *Puyopsis*. The former is characterized as having elongated sterile tips in the compound inflorescence (Smith and Downs, 1974), apparently as an adaptation to bird pollination (Benzing, 2000; Jabaily and Sytsma, 2010), where birds could perch. *Puyopsis* contains the majority of species in the genus which lack sterile bracts (Jabaily and Sytsma, 2010). However, this subdivision is not supported by the latest molecular phylogenies of *Puya* (Jabaily and Sytsma, 2010, 2013).

Traditionally, in the three-subfamily classification of Bromeliaceae, *Puya* was treated as part of the subfamily Pitcairnioideae (Smith and Downs, 1974). However, molecular data have shown Pitcairnioideae to be a paraphyletic group (Crayn *et al.*, 2004; Givnish *et al.*, 2011; 2014). *Puya* has been shown to be monophyletic and is accorded subfamily status as subfamily Puyoideae (Givnish *et al.*, 2011), which is sister to the huge subfamily Bromelioideae (933 species: Luther, 2012). Approximately 90% of subfamily Bromelioideae are CAM plants (Crayn *et al.*, 2004, 2015), in contrast to ~20% CAM species in Puyoideae. There is circumstantial evidence that suggest that the geographical origin of Bromelioideae and Puyoideae was in central Chile (Shulte *et al.*, 2009; Jabaily and Sytsma, 2013; Silvestro *et al.*, 2013). It is unknown whether CAM evolved independently in each of the two subfamilies or in the common ancestor of Bromelioideae and Puyoideae (Givnish *et al.*, 2014; Silvestro *et al.*, 2013).

The phylogeny of *Puya* has been tentatively reconstructed by means of several approaches, using morphological data (Hornung-Leoni and Sosa, 2008), chloroplast and nuclear DNA sequences (Jabaily and Sytsma, 2010) and AFLPs (Jabaily and Sytsma,

2013). Based on the molecular phylogenetic hypotheses the most likely origin of *Puya* was in central Chile. The chloroplast and nuclear phylogenies showed different topologies, consistent with a hybridisation event, which Jabaily and Sytsma (2010) explained by invoking a proposed introgression-hybrid model (a detailed explanation is provided in Chapter 5). The latest phylogenetic analysis of *Puya* (Jabaily and Sytsma, 2013) is the most complete until now in terms of sampling, encompassing 31% of the genus. Jabaily and Sytsma (2013) analysed altitudinal and latitudinal distribution and found that the peak of *Puya* diversity is in the central Andes at altitudes around 3000 m. They found that the Western Andean Portal (WAP) appeared to be an important biogeographic barrier for the northern Andean clade. Jabaily and Sytsma (2010) also stated that CAM had evolved multiple times in *Puya*, although no evidence was provided to support this statement.

In terms of reproductive ecology, it has been claimed that *Puya* is a predominately bird-pollinated genus (Hornung-Leoni *et al.*, 2007, 2013), particularly hummingbirds and passerine birds (Jabaily and Sytsma, 2010; information on some labels of herbarium sheets), with some exceptions for species being pollinated by bats such as *Puya aristeguietae*, *P. ferruginea*, *P. mirabilis* and *P. nitida* (Hornung-Leoni *et al.*, 2013). However, this information only concerns visitors, and until now there has been no serious study identifying the principal pollinator. High-elevation *Puya* species are semelparous plants (Smith and Downs, 1974), meaning that the plant only reproduces once in its life-time. However, mid- and low-elevation *Puya* species have the ability to produce pup rosettes (ramets) or clones at the base before the plant blooms and then dies. This strategy is called semi-semelparous. It has been suggested that the evolution of both systems in *Puya* is highly dependent on the environment and the availability of pollinators. Semelparous species are usually located in high-elevation environments which impose limitations for potential pollinators. Semi-semelparous are species usually located in lowlands where the

environmental conditions are better for poikilotherm pollinators such as hummingbirds (Jabaily and Sytsma, 2013). Dispersal of *Puya* seeds is usually poor (Smith and Downs, 1974; Garcia-Meneses and Ramsay, 2012, 2014).

The concept of climatic niche

The niche is a fundamental concept in ecology. Hutchinson (1958) defined the fundamental niche of a species or a population as the n-multidimensional space of environmental factors (conditions and resources) that interact with the requirements of the species (or the population) to “exist indefinitely”. Hutchinson (1958) distinguished between the fundamental and the realised niche by comparing the separation of the fundamental niches of two competitive species. The separation of niches occurs when the fundamental niches of the two species do not share points in common. Then, in the hypothetical case of two non-separated fundamental niches (of species A and B) the non-intersected part of the fundamental niche of species A with the fundamental niche of species B corresponds to the realised niche of species A. The definition of fundamental niche was maintained as the n-multidimensional space of environmental factors (biotic and abiotic). However, the realised niche was modified by referencing only one species. The realised niche is the intersection of the fundamental niche of species A and the biotope of species A (i.e. the physical space where the species occurs). As described by Begon *et al.* (2006), the difference between a fundamental niche and a realised niche is that the former refers to where an organism could live and the latter refers to where it actually lives.

Depending on the environmental factors to be considered, the interpretation and the scale (temporal and physical) of the niche could be classified as Grinnellian (Grinnell 1917; Whittaker *et al.*, 1973; James *et al.*, 1984; Jackson and Overpeck, 2000; Peterson,

2003; Soberón, 2007) or Eltonian (Elton, 1927; MacArthur, 1969; Vandermeer, 1972; Leibold, 1996, Soberón, 2007). The former emphasises abiotic conditions (e.g. temperature and precipitation) while the later emphasises biotic interactions (e.g. competition). During the last 20 years, due to computational advances and available georeferenced information, the distribution of species was modelled by studying their Grinnellian (climatic) niche. The species distribution models are empirical models based on field observations and environmental predictor variables to generate response surfaces based on probability of occurrence (Guisan and Zimmermann, 2000). They are powerful tools used to investigate conservation biology, and evolutionary and ecological questions (Guisan and Thuiller, 2005). Usually, the analysed output of the species distribution modelling is the translation of the niche into a physical map (the predictive distribution maps). The predicted distribution map is the intersection of the fundamental niche and the potential biotope (Guisan and Thuiller, 2005).

The present thesis used the multidimensional Hutchinson niche concept, with the clarification of the fundamental and realised niche concept as described in Begon *et al.* (2006).

Hypotheses tested in the present study

The aim of this thesis is to uncover new evidence to understand the ecophysiological and evolutionary significance of CAM in the montane genus *Puya*. This will provide evidence of the adaptive value of CAM in montane environments. There are four main hypotheses to be tested concerning the ecology (H₁), evolution (H₂ and H₄) and physiology (H₃) of CAM in *Puya*. The hypotheses are as follows:

H₁: Constitutive CAM species and non-constitutive CAM species of *Puya* have different climatic niches. (Chapter 2)

H₂: CAM evolved multiple times in *Puya*. (Chapter 2 and Chapter 4)

H₃: There are physiological and biochemical differences between low-elevation and high-elevation species of *Puya*. (Chapter 3)

H₄: There are differences in the amino-acid residues of PEPC between CAM and non-constitutive CAM species of *Puya*. (Chapter 4)

Chapter 2. Ecology and evolution of crassulacean acid metabolism in the montane genus *Puya*

Abstract

Despite the extensively studied physiological differences between CAM and C₃ species, the climatic niche differences between them has not been studied in depth, in particular in species that occur at high-elevations. The genus *Puya* (Bromeliaceae) contains 226 species distributed across a wide elevational and latitudinal range. About 21% of *Puya* species are CAM and at least 10 of these CAM species occur above 3000 m. This makes *Puya* a very remarkable genus and a good model to study the ecology and evolution of CAM photosynthesis in high-elevation environments. The present study: (1) tests whether the climatic niches of CAM and C₃ *Puya* species differ; (2) evaluates climatic niche differences within CAM species in *Puya* across an elevational range; and (3) investigates the evolution of CAM and the degree of ecological niche conservatism in the genus. The photosynthetic pathway of 160 species was determined from their $\delta^{13}\text{C}$ values. Climatic niche differences were tested using georeferenced locations for 147 *Puya* taxa with an expanded set of climatic and topographic variables by means of principal component analysis, bootstrap tests, random forest ordination and MaxEnt. A phylogenetic tree of 68 species of *Puya* was constructed using AFLP data, and ancestral character reconstruction of the photosynthetic pathway was performed using maximum likelihood. The present study concludes that the climatic niche of CAM *Puya* species (as opposed to C₃ species) is defined by a low aridity index (i.e. greater aridity), high cloud-

cover seasonality, low cloud cover and high precipitation seasonality. These variables are interconnected with light intensity and water availability. *Puya* species with constitutive CAM are restricted to the central Andes and central Chilean coast, but are not found in the northern Andes. This lack of CAM species in the northern Andes is explained by a selective biogeographic barrier for CAM species located at the *Nudo de Pasto* ($\sim 1^\circ\text{N}$). The 48% of the species in the southern and central Andes clade are constitutive CAM species, which contrasts with 21% of CAM species in the genus as a whole. CAM evolution in *Puya* was complex, with multiple origins and/or reversals of CAM linked to aridification processes in the late Miocene and mid-Pliocene.

Introduction

Crassulacean acid metabolism (CAM) is a complex physiological trait that is widely considered to be a photosynthetic adaptation to water-limited environments (Kluge and Ting, 1978; Osmond, 1978; Winter and Smith, 1996). CAM is characterized by: (1) nocturnal CO_2 uptake from the atmosphere and its fixation into malic acid by phosphoenolpyruvate carboxylase (PEPC, EC 4.1.1.31); and (2) inverted stomatal behaviour, with stomata being open at night but closing for much of the day (while the nocturnally accumulated malic acid is decarboxylated and the released CO_2 refixed by RuBisCO), thereby reducing water loss through transpiration (Osmond, 1978; Borland *et al.*, 2014). Across angiosperms, CAM is thought to have evolved independently from C_3 photosynthesis at least 60 times (Edwards and Ogburn, 2012), and in total approximately 6% of angiosperms exhibit CAM photosynthesis (Smith and Winter, 1996; Silvera *et al.*, 2010a, 2010b, 2014). Despite the well-established physiological dissimilarities between CAM and C_3 species, the climatic niche differences between them has not been evaluated in depth.

The extent of CAM activity can be assessed by determining the $^{13}\text{C}/^{12}\text{C}$ ratio in plant tissue, usually expressed relative to a standard ratio ($\delta^{13}\text{C}$ value: Osmond, 1978; Winter and Holtum, 2002). CAM plants are relatively enriched in ^{13}C compared with C_3 plants because PEPC discriminates less against the heavier ^{13}C isotope than does RuBisCO (Farquhar *et al.*, 1989; Cernusak *et al.*, 2013). However, not all CAM plants exhibit the same degree of nocturnal CO_2 fixation (Winter *et al.*, 2015). Constitutive CAM plants characteristically fix the majority of their atmospheric CO_2 at night using PEPC and therefore their tissues contain a relatively higher concentration of ^{13}C , showing $\delta^{13}\text{C}$ values less negative than -20‰ . In contrast, weak CAM plants (Winter and Holtum, 2015; part of C_3 -CAM plants *sensu* Winter *et al.*, 2015) only fix a small proportion of CO_2 at night, showing small changes in their titratable acidity and $\delta^{13}\text{C}$ values in the range -20 to -27‰ , which are not distinguishable from C_3 plants. Surveys of $\delta^{13}\text{C}$ values across families such as Bromeliaceae and Orchidaceae have tended to reveal strongly bimodal distributions, suggesting that the majority of species engage in constitutive CAM or C_3 photosynthesis, with relatively few weak CAM species (Pierce *et al.*, 2002a; Winter and Holtum, 2002; Crayn *et al.*, 2004; Silvera *et al.*, 2005, 2009; Winter *et al.*, 2015).

Terrestrial plants with CAM photosynthesis are typically succulent plants that show their greatest abundance in warm xeric habitats such as semi-deserts or seasonally dry environments with low precipitation and usually full exposure to solar radiation (Winter, 1985; Winter *et al.*, 2015; Holtum *et al.*, 2016). Ellenberg (1981) suggested that the occurrence of perennial stem-succulent plants, such as cacti and euphorbias, was determined by low annual precipitation ($<500\text{ mm year}^{-1}$) combined with relatively predictable inter-annual precipitation. Ellenberg (1981) defined the inter-annual variability as the ratio between precipitation of the wettest year and precipitation of the driest year over an extended sequence of years. Using this ratio, Ellenberg (1981) showed

that perennial stem-succulent plants were generally restricted to areas where the ratio was lower than 5. However, Ellenberg's hypothesis has not been tested for CAM plants with other growth forms, for example leaf-succulent giant rosettes. Moreover, Ellenberg's results did not explicitly characterize the contrasting niches occupied by CAM and C₃ plants.

The description of the climatic niche occupied by CAM plants may differ depending on the taxonomic group being evaluated. According to Alvarado-Cardenas *et al.* (2013), there is no environmental equivalence between the climatic niches of South African (Apocynaceae, Aizoaceae and Euphorbiaceae) and Mexican (Cactaceae) globular succulent plants with convergent morphology. The results of Alvarado-Cardenas *et al.* (2013) indicate that morphological similarity may not equate to niche equivalency. Therefore, in any attempt to define differences between CAM and C₃ climatic niches, it is important to control for both life form and evolutionary history in comparing CAM and C₃ species.

The current understanding of the climatic niche differences between CAM and C₃ species is incomplete. Edwards and Ogburn (2012) concluded there were no obvious differences in the climatic niche of CAM, C₄ and C₃ species in the "Portullugo clade" in Caryophyllales, a clade containing many CAM species in the suborder Portulacineae. This might be explained by the inclusion of species with different life-history strategies and growth habits (perennial *vs.* annual species and terrestrial *vs.* epiphytic species).

In contrast, Heyduk *et al.* (2016a) concluded that differences in the climatic niche between CAM and C₃ species in the Agavoideae (Asparagaceae) were driven primarily by temperature-related variables. Both studies focused on monophyletic clades and used principal component analysis (PCA) to investigate the differences between CAM and C₃ climatic niches. PCA is a statistical procedure that converts correlated *n*-variables into

uncorrelated n -components, where the first components retain the greatest possible variance (Rao, 1964). As such, it is a tool to summarise variances rather than test differences among groups (Legendre and Legendre, 2012). Moreover, after using PCA, neither Edwards and Ogburn (2012) nor Heyduk *et al.* (2016a) rigorously tested differences between CAM and C₃ groups, or their possible overlap, and both studies were limited to the 19 standard WorldClim bioclimatic variables (Hijmans *et al.*, 2005) and so did not take into account climatic factors important for CAM species, including inter-annual precipitation variability (Ellenberg, 1981), light intensity (Nobel, 1980a; Adams and Demmig-Adams, 1996) and measures of environmental water availability besides precipitation. Such factors may be important since CAM is a water-saving mechanism.

The Bromeliaceae, a diverse family of Neotropical angiosperms, comprises over 3500 species (Gouda *et al.*, 2016) and is an excellent model to study the ecology and evolution of CAM (Crayn *et al.*, 2015). In a survey covering nearly two-thirds of all Bromeliaceae species, Crayn *et al.* (2015) found that 43% of species were CAM plants. CAM appears to have evolved independently at least five times in the family (Crayn *et al.*, 2004, 2015; Givnish *et al.*, 2014). Bromeliaceae species occupy a wide range of ecological niches, from mesophytic to xerophytic environments and from sea level up to 5000 m in the tropical and subtropical Americas (Smith and Downs, 1974, 1977, 1979). The majority of the genera are consistently either CAM or C₃, reflecting a pattern of ecological niche conservatism at this taxonomic level (Crayn *et al.*, 2004, 2015; Givnish *et al.*, 2014). However, the large polyphyletic and mainly epiphytic genus *Tillandsia* and the monophyletic genus terrestrial *Puya* contain a mixture of C₃ and CAM species (Barfuss *et al.*, 2016; Crayn *et al.*, 2015) and therefore represent particularly interesting lineages for the comparative analysis of C₃ and CAM species.

Puya, a predominantly montane genus, comprises 226 currently described species (Gouda *et al.*, 2016) distributed through the Andes (from 38°S to 11°N), Guiana shield and Talamancan paramo (Costa Rica), and from sea level up to 5000 m (Smith and Downs, 1974; (Zizka *et al.*, 2009, 2013; Jabaily and Sytsma, 2010, 2013). Phylogenetic reconstruction suggests that colonization of the Andes by *Puya* occurred from south to north (Jabaily and Sytsma, 2010, 2013). Dispersal of *Puya* seeds is usually poor (Smith and Downs, 1974; Garcia-Meneses and Ramsay, 2012, 2014). Consequently, most *Puya* species have narrow geographical distributions (Varadarajan, 1990). Low dispersion may explain the strong association between geography and phylogeny in the genus (Varadarajan, 1990, Jabaily and Sytsma, 2010, 2013). Additionally, the $\delta^{13}\text{C}$ survey of Crayn *et al.* (2015) reported that 21% of the 132 species of *Puya* tested were CAM. *Puya* is sister to subfamily Bromelioideae (90% CAM), but it is unknown whether CAM evolved independently in each subfamily or in the common ancestor of Bromelioideae and Puyoideae (Givnish *et al.*, 2014; Silvestro *et al.*, 2013).

Consistent with the notion that CAM photosynthesis is an adaptation for water conservation in dry and hotter environments, the frequency of CAM species tends to decrease with increasing elevation (Griffiths and Smith, 1983; Earnshaw *et al.*, 1987; Pierce *et al.*, 2002b). However, *Puya* contains at least 10 constitutive CAM species that occur at more than 3000 m above sea level (Crayn *et al.*, 2015). In addition to *Puya*, only four Neotropical high-elevation taxa have been shown to exhibit direct evidence of night-time CO_2 fixation: *Echeveria quitensis* (syn. *E. columbiana*) (Crassulaceae; Medina and Delgado, 1976); *Oroya peruwiana* and *Tephrocactus floccosus* (syn. *Austrocylindropuntia floccosa*) (Cactaceae; Keeley and Keeley, 1989); and *Isoetes* (Isoetaceae; Keeley, 1998). The low abundance of high-elevation CAM species may be due to the low minimum night-time temperatures (usually below 0 °C) in the Neotropical highlands and an intrinsic cold

sensitivity of CAM plants (Goldstein *et al.*, 1985; Rada *et al.*, 1985; Keeley and Keely, 1989; Yamori *et al.*, 2014). Currently, no study evaluating the climatic niche or CAM evolution of high-elevation CAM species has been published. Studying the climatic niche of CAM across a broad range of elevations allows us to understand whether CAM plants occur only in warmer and drier environments (Winter *et al.*, 2015; Heyduk *et al.*, 2016a).

Puya is a monophyletic genus of bromeliads with a highly conserved growth habit, but which contains many CAM and C₃ species over a broad range of latitudes and elevations. Consequently, this genus can serve as a valuable model to: (1) test whether the realised and potential climatic niche of CAM species differ from the realised and potential climatic niche of C₃ species; (2) investigate the climatic niche differences between different CAM species; and (3) understand the evolution of CAM in the genus and the degree of ecological niche conservatism in CAM and C₃ species.

Materials and Methods

Carbon-isotope values

Plant tissue for ¹³C isotopic analysis and nitrogen content was collected from specimens in Missouri Botanical Garden (MO), Royal Botanic Gardens Kew (K), *Herbario Nacional Colombiano* (COL), material from Schulte *et al.* (2010) and from Jabaily and Sytsma (2013), Ventnor Botanic Garden (UK) and the University of Oxford Botanic Garden (UK). Typically, 1–3 mg of tissue dry biomass were ground using a pestle and mortar and samples analysed by mass spectrometry in the School of Archaeology at the University of Oxford. Reference CO₂ was calibrated against the original standard Pee Dee belemnite (PDB) via a secondary calibration against Vienna belemnite. The carbon-isotope ratio (δ¹³C value) was calculated as described in Crayn *et al.* (2015):

$$\delta^{13}\text{C} (\text{‰}) = [({}^{13}\text{C}_{\text{sample}} / {}^{12}\text{C}_{\text{sample}}) / ({}^{13}\text{C}_{\text{PDB}} / {}^{12}\text{C}_{\text{PDB}}) - 1] \times 1000$$

The photosynthetic pathway was determined according to the carbon isotope values of 160 species of *Puya* (132 species from Crayn *et al.* (2015) plus 28 species from this study: Table 1). Plants with $\delta^{13}\text{C}$ values less negative than -20‰ were classified as constitutive CAM; plants with $\delta^{13}\text{C}$ values equal to or more negative than -20‰ were classified as C_3 (Farquhar *et al.*, 1989; Crayn *et al.*, 2004, 2015; Winter *et al.*, 2015). Nitrogen content was determined in 52 individuals (40 *Puya* species) to test for any correlation between carbon-isotope value and the nitrogen content.

Table 1. $\delta^{13}\text{C}$ values used in the present study from 51 species (107 individuals) of *Puya* in addition to the carbon isotope values reported in Crayn *et al.* (2015). Taxonomic nomenclature based on based on Gouda *et al.* (2016). (*) No previous determination of $\delta^{13}\text{C}$ values in any other study.

Species	$\delta^{13}\text{C}$ values	Collector
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-25.7	C. Annable 2236 (NY)
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-23.9	Zizka <i>et al.</i> 8087-91
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-28.4	Zizka <i>et al.</i> 8093
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-26.3	Zizka <i>et al.</i> 8086.
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-25.2	Zizka <i>et al.</i> 8075-77
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-24.2	Zizka <i>et al.</i> 8066-69
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-22.5	Zizka <i>et al.</i> 8042-43
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-24.2	Zizka <i>et al.</i> 8144
<i>Puya alpestris</i> ssp. <i>alpestris</i>	-23.6	Zizka <i>et al.</i> 8143
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-22.6	Zizka <i>et al.</i> 8036-39
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-22.8	Zizka <i>et al.</i> 8003-4
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-19.1	Zizka <i>et al.</i> 8005
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-19.0	Zizka <i>et al.</i> 8009
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-16.1	Zizka <i>et al.</i> 8132-34
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-16.0	Zizka <i>et al.</i> 8125-27
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-15.4	Zizka <i>et al.</i> 8139
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-16.5	Zizka <i>et al.</i> 8149
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-20.3	Zizka <i>et al.</i> 8010-14
<i>Puya angelensis</i>	-25.4	J.M. Manzanares & W. Till 7970 (MO)
<i>Puya asplundii</i> *	-24.5	J.M. Manzanares & W. Till 7965 (MO)
<i>Puya assurgens</i>	-26.2	G.S. Varadarajan <i>et al.</i> 1257 (MO)
<i>Puya bermejiana</i>	-23.4	A. Grau <i>et al.</i> 1575 (MO)
<i>Puya brackeana</i>	-23.2	J.M. Manzanares & B. Patterson 7383 (MO)
<i>Puya bravoii</i>	-23.9	A. Grau <i>et al.</i> 1596 (MO)

<i>Puya cajasensis</i>	−24.3	J. Campos de la Cruz <i>et al.</i> 5309 (MO)
<i>Puya cardenasii</i>	−21.5	G.S. Varadarajan <i>et al.</i> 1463 (MO)
<i>Puya casmichensis</i> *	−19.8	R.S. Jabaily 223
<i>Puya chilensis</i>	−25.0	Quezada <i>et al.</i> (2014)
<i>Puya chilensis</i>	−25.3	Zizka <i>et al.</i> 8083
<i>Puya chilensis</i>	−28.1	Zizka <i>et al.</i> 8079–82
<i>Puya chilensis</i>	−25.8	Zizka <i>et al.</i> 8001
<i>Puya chilensis</i>	−22.8	Quezada <i>et al.</i> (2014)
<i>Puya chilensis</i>	−21.5	Zizka <i>et al.</i> 8006
<i>Puya chilensis</i>	−23.1	Zizka <i>et al.</i> 8109–12
<i>Puya chilensis</i>	−23.5	Zizka <i>et al.</i> 8114–18
<i>Puya chilensis</i>	−22.6	Quezada <i>et al.</i> (2014)
<i>Puya chilensis</i>	−21.7	Zizka <i>et al.</i> 8135–38
<i>Puya chilensis</i>	−21.7	Quezada <i>et al.</i> (2014)
<i>Puya chilensis</i>	−16.3	G.S. Varadarajan <i>et al.</i> 1487 (MO)
<i>Puya chilensis</i>	−19.7	Quezada <i>et al.</i> (2014)
<i>Puya chilensis</i>	−18.6	G.S. Varadarajan <i>et al.</i> 1484 (MO)
<i>Puya claudiae</i> *	−13.9	M.H. Nee 36230 (MO)
<i>Puya coerulea</i>	−24.5	Zizka <i>et al.</i> 8030–35
<i>Puya coerulea</i>	−23.8	Zizka <i>et al.</i> 8027
<i>Puya coerulea</i>	−23.2	Zizka <i>et al.</i> 8046–49
<i>Puya coerulea</i>	−23.7	Zizka <i>et al.</i> 8016–19
<i>Puya coerulea</i>	−22.3	Zizka <i>et al.</i> 8045–46
<i>Puya coerulea</i>	−21.8	Zizka <i>et al.</i> 8025
<i>Puya coerulea</i>	−22.1	Zizka <i>et al.</i> 8008
<i>Puya couerulea</i> var. <i>violaceae</i>	−21.9	O. Zöllner 8412 (MO)
<i>Puya cristata</i>	−13.3	M. Lewis 881103 (MO)
<i>Puya cueva</i>	−24.2	J. Campos de la Cruz <i>et al.</i> 5057 (MO)
<i>Puya dodsonii</i>	−25.0	C.H. Dodson & P.M. Dodson 15391 (MO)
<i>Puya entrierosensis</i> *	−29.9	M. Serrano 4729 (MO)
<i>Puya fulgens</i> *	−24.9	J.M. Manzanares <i>et al.</i> 7470 (MO)
<i>Puya gilmartinae</i> *	−21.8	G.S. Varadarajan <i>et al.</i> 1481 (MO)
<i>Puya glandulaosa</i>	−18.5	P.C. Hutchison <i>et al.</i> 6123 (MO)
<i>Puya herrerae</i>	−27.6	I. Huamantupa <i>et al.</i> 6902 (MO)
<i>Puya hirtzi</i> *	−27.6	J.M. Manzanares 5989 (MO)
<i>Puya humilis</i>	−25.4	I. Linneo <i>et al.</i> 773 (MO)
<i>Puya x joergensenii</i> *	−26.8	R.J. Jabaily 116
<i>Puya kuntzeana</i>	−25.6	J. Villalobos 472 (MO)
<i>Puya leptostachya</i> *	−24.2	M. Lewis 38513 (MO)
<i>Puya longispina</i> *	−23.0	J.M. Manzanares & W. Till 8010 (MO)
<i>Puya loxensis</i> *	−25.0	P.M. Jørgensen <i>et al.</i> 428 (MO)
<i>Puya macropoda</i>	−28.8	M.H. Nee 30283 (MO)

<i>Puya mimia</i> *	−16.9	A. Sagástegui 15392 (MO)
<i>Puya mucronata</i> *	−24.3	A. Sagástegui 16094 (MO)
<i>Puya navarroana</i> *	−27.3	J.M. Manzanares & B. Patterson 7386 (MO)
<i>Puya novarae</i> *	−22.5	R.J. Jabaily 158
<i>Puya olivacea</i>	−25.8	R. Lozano 1708 (MO)
<i>Puya pattersoniae</i> *	−27.1	J.M. Manzanares <i>et al.</i> 8087 (MO)
<i>Puya</i> × <i>pichinchae</i> *	−22.3	C.E. Cerón <i>et al.</i> 3042 (MO)
<i>Puya</i> aff. <i>pratensis</i>	−25.1	P.J. Barbour 3375 (MO)
<i>Puya pseudoeryngioides</i> *	−25.3	H. van der Werff <i>et al.</i> 14950 (MO)
<i>Puya pyramidata</i> *	−18.8	J. Campos de la Cruz & P. López 4952 (MO)
<i>Puya rauhii</i> *	−23.0	D.N. Smith <i>et al.</i> 12724A (MO)
<i>Puya roelzii</i>	−20.9	T.H. Goodspeed 33076 (MO)
<i>Puya roldanii</i> *	−22.2	R. Fonnegra G. & D. Tuberquia 4644 (MO)
<i>Puya roseana</i> *	−23.6	J.M. Manzanares <i>et al.</i> 8089 (MO)
<i>Puya tillii</i> *	−26.2	J.M. Manzanares 5738 (MO)
<i>Puya trollii</i> *	−24.2	M. Serrano <i>et al.</i> 6843A (MO)
<i>Puya venusta</i>	−23.7	Zizka <i>et al.</i> 8120–24
<i>Puya weberiana</i>	−21.6	G.S. Varadarajan <i>et al.</i> 1460 (MO)
<i>Puya weddelliana</i> *	−21.0	J. Villalobos <i>et al.</i> 964 (MO)

Distribution of CAM and C₃ species

Data on *Puya* species distribution were retrieved from the Global Biodiversity Information Facility (GBIF, 2014), and from specimens in Missouri Botanical Garden (MO), Royal Botanic Gardens Kew (K), *Herbario Universidad de Concepción*, Chile (CONC), *Herbario Nacional Colombiano* (COL), and *Botanischer Garten und Botanisches Museum Berlin-Dahlem* (B). In total, 3744 herbarium records of *Puya* were retrieved, of which 2027 (54.1%) records were georeferenced. The data were reviewed to identify common errors such as discrepancies between the municipality and the locality reported on the label, discrepancies between locality and the georeferenced coordinates, clear mistakes in the sign of longitude (e.g. specimens appeared in Asia rather than in South America, or specimens occurring in the middle of the ocean). This process was performed manually by reviewing each herbarium specimen and comparing label data with Google Earth

(Frank *et al.*, 2013) georeferenced coordinates. After data cleaning, 433 herbarium records, representing 147 *Puya* taxa, had consistent georeferenced information (accuracy of 0.5 km), locality and carbon isotope values (CAM: 31 spp., 56 records; C₃: 116 spp., 377 records). The distribution of these taxa was plotted using ArcGIS 10.2.2 (Law and Collins, 2013).

Climatic data

Thirty-two climatic and topographic variables were included in the analyses (Table 2). Nineteen bioclimatic variables, and elevation, were retrieved from WorldClim (Hijmans *et al.*, 2005). The ratio between mean annual precipitation and potential evapotranspiration (aridity index: Middelton and Thomas, 1997) and the ratio between actual and potential evapotranspiration (R-index: Yao, 1974) were retrieved from the Consultative Group for International Agricultural Research (CGIAR; Zomer *et al.*, 2007, 2008; Trabucco and Zomer, 2010). Annual cloud cover, cloud-cover intra-annual variability and inter-annual variability (seasonality) were retrieved from Wilson and Jetz (2016). Additionally, seasonality of the variables from CGIAR was calculated as the intra-annual standard deviation. Topological variables, slope, roughness and the north–south and east–west orientation of slopes were calculated using the function *terrain* in the R package raster (Wilson *et al.*, 2007; Hijmans and van Etten, 2014). Finally, precipitation inter-annual variability was calculated based on the dataset from New *et al.* (2002) from 1961–1990 annual precipitation means, using the ratio between precipitation of the wettest year divided by the precipitation of the driest year (see Ellenberg, 1981). Note that precipitation seasonality (bio 15; Hijmans *et al.*, 2005; O’Donnell and Ignizio, 2012) is the intra-annual coefficient of variation and it is different from the inter-annual variability of precipitation suggested by Ellenberg (1981). Resolution of the data was 30 arc seconds (*ca.* 1 km² at the Equator), except for the inter-annual precipitation variability at a

resolution of 10 min (*ca.* 340 km² at the Equator). Deblauwe *et al.* (2016) generated climatic variables based on remotely sensed data that improve species distribution model accuracy. However, this improvement is only significant at low elevations (e.g. the Amazon) where the density of climatic stations is very low compared with the Andes.

Spatial autocorrelation is a common problem in biogeographic analysis and species distribution models (Araujo *et al.*, 2005; Segurado *et al.*, 2006). Subsampling and bootstrapping are different approaches to avoid the non-independence problem of spatially autocorrelated data (Araujo *et al.*, 2005; Austin, 2007; Parolo *et al.*, 2008). All climatic variables were extracted for each georeferenced herbarium record to establish the climatic niche of CAM and C₃ *Puya* taxa, and for each variable the multiple individual record from one species were combined in one point per species to avoid heavily spatial autocorrelation and unbalance of common species over rare species.

Table 2. List of 32 climatic and topographic variables used in the present study and their source.

Climatic and topological variables	Units	Source
Annual mean temperature	°C	Hijmans <i>et al.</i> (2005)
Mean temperature of warmest quarter	°C	Hijmans <i>et al.</i> (2005)
Mean temperature of coldest quarter	°C	Hijmans <i>et al.</i> (2005)
Annual precipitation	mm	Hijmans <i>et al.</i> (2005)
Precipitation of the wettest month	mm	Hijmans <i>et al.</i> (2005)
Precipitation of the driest month	mm	Hijmans <i>et al.</i> (2005)
Precipitation seasonality (coefficient of variation of annual precipitation)	%	Hijmans <i>et al.</i> (2005)
Precipitation of the wettest quarter	mm	Hijmans <i>et al.</i> (2005)
Precipitation of the driest quarter	mm	Hijmans <i>et al.</i> (2005)
Precipitation of the warmest quarter	mm	Hijmans <i>et al.</i> (2005)
Precipitation of the coldest quarter	mm	Hijmans <i>et al.</i> (2005)
Mean temperature diurnal range	°C	Hijmans <i>et al.</i> (2005)
Isothermality (mean temperature diurnal range/temperature annual range)	mm mm ⁻¹	Hijmans <i>et al.</i> (2005)
Temperature seasonality	SD °C	Hijmans <i>et al.</i> (2005)
Maximum temperature of warmest month	°C	Hijmans <i>et al.</i> (2005)
Minimum temperature of coldest month	°C	Hijmans <i>et al.</i> (2005)
Temperature annual range	°C	Hijmans <i>et al.</i> (2005)
Mean temperature of wettest quarter	°C	Hijmans <i>et al.</i> (2005)
Mean temperature of driest quarter	°C	Hijmans <i>et al.</i> (2005)
Elevation	m	Hijmans <i>et al.</i> (2005)
Precipitation inter-annual variability	mm mm ⁻¹	This study
East slope orientation	m m ⁻¹	This study
North slope orientation	m m ⁻¹	This study
Roughness of the slope (Max. – Min. of the slope in a 9 km ²)	Radians	This study
SD of R-index	SD mm mm ⁻¹	This study
SD of aridity index	SD mm mm ⁻¹	This study
Slope (based on the elevation of nearest 4 grid km ²)	Radians	This study
Cloud cover inter-annual variability	SD %	Wilson and Jetz (2016)
Cloud cover seasonality (variability within a year)	SD %	Wilson and Jetz (2016)
Mean annual cloud frequency	%	Wilson and Jetz (2016)
R-index (actual evapotranspiration/potential evapotranspiration)	mm mm ⁻¹	Zomer <i>et al.</i> (2007)
Aridity index (annual precipitation/potential evapotranspiration)	mm mm ⁻¹	Zomer <i>et al.</i> (2007)

Statistical analysis

All statistical analyses were conducted using the software R (R Core Team, 2017) and the following packages: *sp* (Pebesma and Bivand, 2005); *raster* and *dismo* (Hijmans and van Etten, 2014; Hijmans *et al.*, 2015); *MASS* (Venables and Ripley, 2002); *lattice* (Sarkar, 2008); *e1071* (Meyer *et al.*, 2014); *ggplot2* (Wickham, 2009); and *ade4* (Dray *et al.*, 2007). The normality of the data was tested using the Shapiro–Wilk test (Shapiro and Wilk, 1965). Non-normal data were transformed using the formula $\ln(x^2 + 1)$ and the normality tested after transformation. Heteroscedasticity was tested using a Bartlett-test of homogeneity of variances. Data were standardised (mean = 0, SD = 1) and analysed by principal component analysis (PCA) and random forest ordination (RF). These two statistical approaches do not have strong requirements for normality or homoscedasticity.

Climatic variation was summarized by conducting a PCA. After PCA, the differences between the centroids of the CAM and C₃ clusters were tested using a stochastic sampling method (number of independent interactions = 100 000; sample size = 20; replacement was allowed, with the photosynthetic pathway included *a posteriori*). For each component (PC1, PC2, PC3), the probabilistic frequency of the difference between means of CAM and C₃ species were determined and tested to determine whether centroids were different between CAM and C₃.

Bootstrapping is a permutation test to determine whether or not an observed pattern is likely to be caused by chance (Legendre and Legendre, 2012). Bootstrapping is useful when the distribution of the data is unknown or the data do not conform to the assumptions required for parametric and non-parametric tests (Legendre and Legendre, 2012). The distribution curve is obtained by randomly permuting (bootstrapping) the data according to the null hypothesis to be tested. In the present study, the null hypothesis was that there were no significant differences between the niches of CAM and C₃ species. This

hypothesis was evaluated based on the mean, median, minimum and maximum values. Bootstrapping was particularly appropriate in this study as the data were not normally distributed and showed heteroscedasticity.

A bootstrap test was conducted to evaluate differences in climatic niche given the photosynthetic pathway. The differences between CAM and C₃ in each climatic variable was evaluated by conducting a bootstrapping test on the mean and median of each group (number of independent interactions = 100 000; sample size = 20; replacement was allowed). The overlap between groups was evaluated by conducting a bootstrapping test on the minimum and maximum of each climatic variable between CAM and C₃. The complete overlap was defined if there were no significant differences between CAM and C₃ in the median, mean, minimum and maximum values. The *p*-values generated by the bootstrapping test were adjusted using Bonferroni correction for multiple comparisons (Legendre and Legendre, 2012).

Random Forest (Breiman, 2001) is a non-parametric machine-learning method for classifying high-dimensional data, such as biogeographic problems or species distribution models (Garzón *et al.*, 2008; Lawler *et al.*, 2009; Hill *et al.*, 2017). It consists of a collection of independent decision trees (classification or regression trees), where each individual decision tree is built from a subset of the data using a subset of the variables. This process is repeated independently several times (bootstrapping), producing a collection of independent decision trees to classify the whole data, minimising noise and error. At the end of the analysis, a consensus tree accurately classifies high-dimensional data and ranks the contribution of each variable. Random Forest can be biased if different types of variables are included in the model (categorical *vs.* continuous) or if there is a disproportionate category within the data.

Fifty independent RF ordinations were conducted using the function *cforest* from *Party* R package (Hothorn *et al.*, 2006) to identify those variables capable of recovering the photosynthetic pathway given the climatic data. Ten thousand trees were used for each run; the maximum number of variables in each case was six; and replacement was not allowed, to prevent bias in the variable importance given the high correlation between the variables (Srobl *et al.*, 2007, 2008). To compensate for the natural imbalance in CAM and C₃ species in *Puya*, the data to build the model were equilibrated with the Synthetic Minority Over-sampling Technique, function *SMOTE* in *DMwR* (Torgo, 2010). The accuracy of the model was tested using the non-equilibrated data. The importance of each variable in the model was evaluated using a permutation test to optimize the area under the curve (AUC), as described in Janitza *et al.* (2013) for each independent RF ordination. The accuracy of nested ordinations was calculated using the most important 2, 3, 4 and 5 variables, based on a permutation test. These results were compared with the accuracy of the ordination using all variables. The same analysis described above was conducted with a partitioned dataset from the central Andes and Chilean region only, to test whether the results differed. A dataset with only CAM species was used to identify the variables that defined the high-elevation CAM niche. In this case, the variable of interest was elevation rather than photosynthetic pathway and the analysis was a RF regression.

Fundamental niche

The fundamental climatic niche of CAM and C₃ *Puya* species was modelled using MaxEnt (Phillips *et al.*, 2006). Fundamental CAM and C₃ niches were mapped to establish whether there were differences in the current distribution of CAM and C₃ *Puya* species and the potential niche. This approach is suitable for presence-only data to infer the fundamental niches (Elith *et al.*, 2011). In this analysis, each herbarium record (433

records) was taken as an independent point and was labelled according to its photosynthetic pathway. Bootstrap replicates with 100 000 background points from South America and the southern part of Central America were used as test parameters. Jackknife analysis was performed to determine the relative importance of individual variables in the whole analysis. Overall, 75% of the data were used to build up the model and the other 25% to test the model. All parameters of the model were tested using 100 bootstrap replicates and the output was mapped in ArcGIS 10.2.2 (ESRI, 2014). The difference between the fundamental niches of CAM and C₃ species was mapped in ArcGIS 10.2.2 to visualize suitable areas for CAM and C₃.

Ancestral state reconstruction and phylogenetic signal analysis

A Bayesian-inference phylogenetic tree of 68 species of *Puya* was constructed using an AFLP data set (Jabaily and Sytsma, 2013). Species with multiple accessions were pruned to single terminal before tree construction, except for *Puya chilensis* and *P. alpestris*, for which populations showing CAM and C₃ were included as independent taxa. An ultrametric tree was generated using the AFLP phylogeny and the function *chronos* (relaxed model, lambda 0, crown age: 9.4 Ma according to Givnish *et al.*, 2014) in package *ape*. The ancestral state reconstruction of photosynthetic pathway was performed using maximum likelihood under two different models: equal transition rates (ER) and all-different transition rates (ARD), using the function *ace* in the package *ape*. The best model was selected according to the highest log likelihood using a maximum likelihood statistical significance test, assuming a chi-square distribution and one degree of freedom (*G*-statistic). The most important variables suggested by the bootstrap analysis and the RF were analysed using Phylogenetic Generalized Least Squares (PGLS; Grafen, 1989; Martins and Hansen, 1997; Pagel, 1997, 1999; Rohlf, 2001).

Phylogenetic signal was calculated for the first three principal components (see above) using Blomberg's K (Blomberg *et al.*, 2003). Blomberg's K indicates no phylogenetic signal when $K = 0$. When $K = 1$, the trait is distributed across the phylogeny following Brownian Motion (BM). When $K > 1$, there is stronger phylogenetic signal than expected under BM. The significance of K (p -values) was estimated using 1000 randomisations using the function *multiPhylosignal* in the package *picante* (Kembel *et al.*, 2010). A phylogenetic ANOVA (function *phylANOVA* in *phytotools* package; Silsbe and Malkin, 2015; Rohlf and Nielsen, 2015) was conducted to establish the differences between the climatic variables suggested in the RF and bootstrap analyses and the first three principal components between CAM and C₃, after taking into account phylogenetic information.

Results

$\delta^{13}\text{C}$ values in *Puya*

Carbon isotope values were determined for 28 *Puya* species not previously included in the survey of Crayn *et al.* (2015). Out of the total of 161 species sampled (covering 75 % of the genus), 22% of the surveyed *Puya* species are constitutive CAM species (Table 1). The species with the least negative carbon isotope value was *P. meiziana* (−12.0 ‰; 4000 m elevation). The species with the most negative value was *P. grafii* (−31.9 ‰; 100 m elevation). Species with multiple accessions showed a range of carbon isotope values that were consistent with constitutive CAM $\delta^{13}\text{C}$ values or non-constitutive $\delta^{13}\text{C}$ values, except for *P. chilensis* (range −16.3 ‰ to −28.1 ‰) and *P. alpestris* (−16.1 ‰ to −28.4 ‰). The peak frequency in the distribution of $\delta^{13}\text{C}$ values in C₃ species was −23.9 ‰, and for CAM species was −14.8 ‰ (Supplementary Fig. 1). There was a positive significant correlation between elevation and $\delta^{13}\text{C}$ values in C₃ species (slope = 0.2270, p = 0.0147; Supplementary Fig. 2), but no significant correlation in CAM species (slope =

0.0370, $p = 0.8431$). There was no significant correlation between $\delta^{13}\text{C}$ value and nitrogen content in *Puya* ($p = 0.4916$, $n = 52$, Supplementary Fig. 3).

Distribution of CAM and C₃ *Puya* species

Puya is distributed across the Andes, from northern South America to southern Chile, with two species distributed in Costa Rica (Fig. 1). The northern distributed species of *Puya* is *P. alpicola* (C₃) found in Sierra Nevada de Santa Marta, Colombia (11° 7'N), the southernmost distributed species is *P. alpestris* (C₃ population) found in Chile (38° 19'S). CAM species are restricted to the central Andes and northern Argentina and Chile (Fig. 1). There are no constitutive CAM species in the northern Andes, the Guiana shield or the Talamancan paramo (Costa Rica). In contrast, C₃ species are distributed across the range of the genus (Fig. 1). The northernmost population of *P. lanata*, the most northerly distributed CAM species, is located at 3° 27' S, at the southern limit of the western Andean Portal.

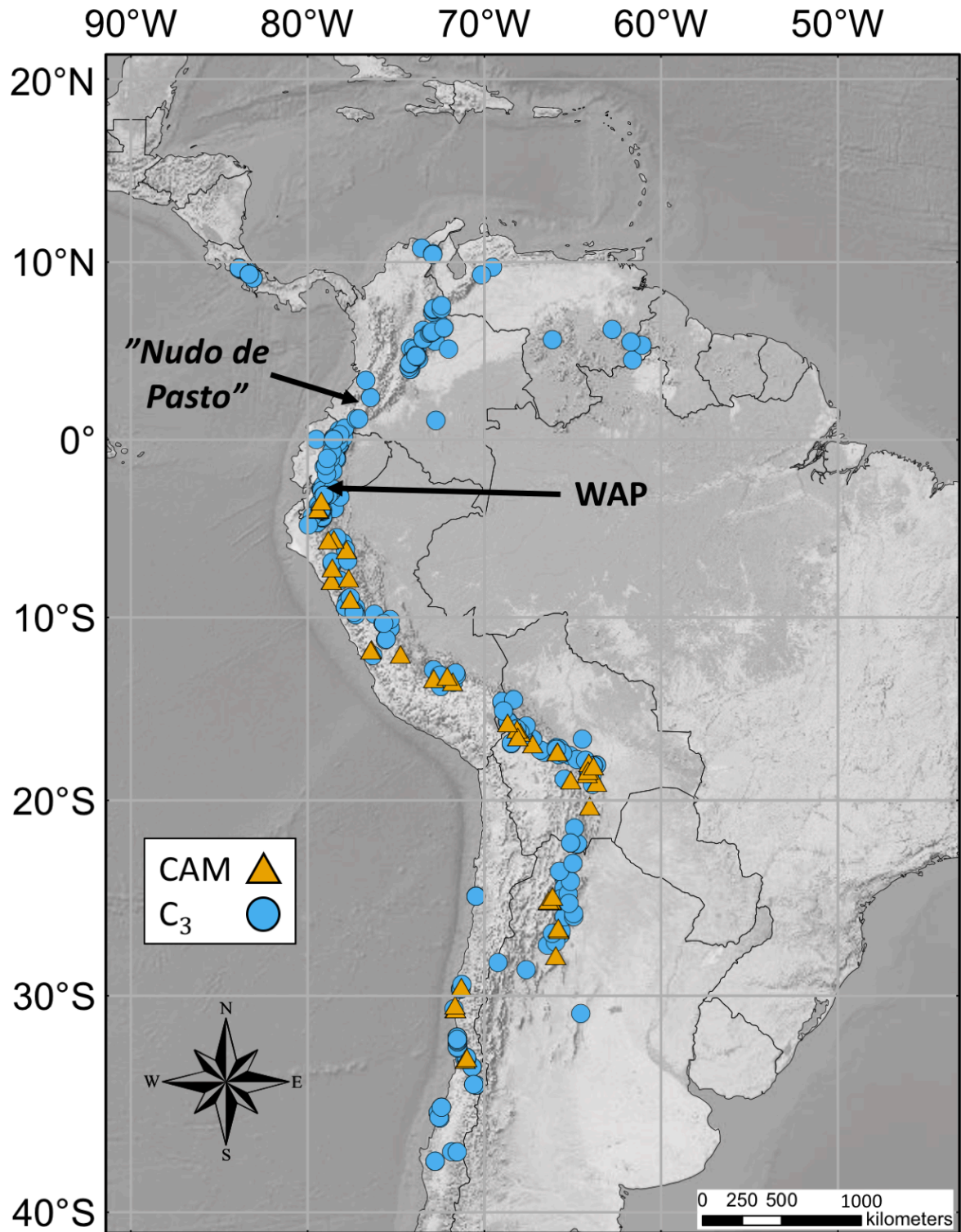


Figure 1. Distribution of *Puya* species based on the georeferenced location of 433 herbarium records from 147 species. Photosynthetic pathway was measured based on the carbon isotope values for each species. Orange triangles represent CAM species (56 records, 31 spp.) and blue circles represent C_3 species (377 records, 116 spp.). The arrows signal the Western Andean Portal (WAP) and the *Nudo de Pasto*. The distribution of CAM species is restricted to the central Andes and northern Chile.

Principal component analysis

None of the climatic and topographic variables was normally distributed before transformation. After the transformation, 25 of the 32 climatic variables were not normally distributed, whilst 17 variables showed significant heteroscedasticity (Supplementary Table 1). Principal component analysis was conducted to visualise whether or not CAM and C₃ species formed clusters (Fig. 2). All climatic variables were used in the PCA analysis. The first two principal components (PC1 and PC2) together accounted for 57.6% of the variance of the climatic data. CAM species were only located on the positive side of PC1, while C₃ species were dispersed across PC1 with a mean value on the negative side of PC1. PC2 did not separate CAM and C₃ species (Fig. 2; Supplementary Table 3). The three most important variables that correlated positively with PC1 were precipitation seasonality (eigenvector = 0.2383), intra-annual cloud cover variability (eigenvector = 0.2328), and temperature annual range (eigenvector = 0.2320). The most important variables that correlated negatively were R-index (eigenvector = -0.2715), precipitation of the driest quarter (eigenvector = -0.2502) and aridity index (eigenvector = -0.2477). The centroids of CAM and C₃ were tested by bootstrapping and only significant differences were found in PC1 ($p < 0.0001$) and not for PC2 ($p = 0.3788$).

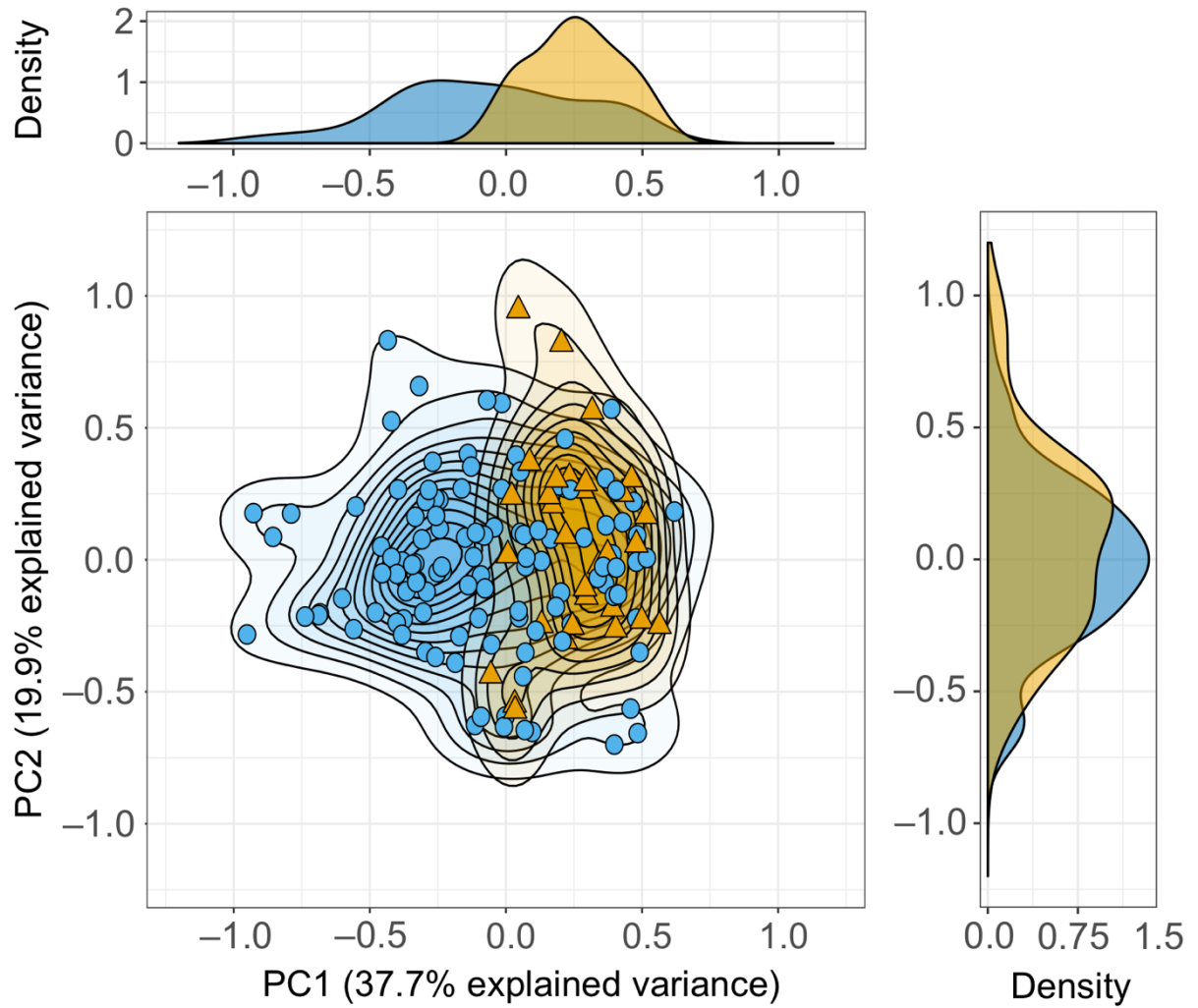


Figure 2. Principal component analysis of 32 climatic and topographic variables listed in Table 2 from 147 species of *Puya*. The first two principal components together explain 57.6% of the variance. CAM species are represented as orange triangles and C₃ species are represented as blue circles. The top and right panels show the density distribution of CAM (orange) and C₃ (blue) species across the PC1 and PC2, respectively. The shade of the isoclines represent density in the distribution of CAM and C₃ species. The centroids of CAM and C₃ distribution are significantly different only for the PC1 axes ($p < 0.0001$). There is a large overlap between CAM and C₃ distribution on the PC2 axes as shown on the right panel.

Climatic differences given the photosynthetic pathway

Only nine of 32 individual bioclimatic and topographic variables showed significant differences in their mean values between CAM and C₃ species according to

the bootstrap analysis (Supplementary Table. 2). Four climatic variables exhibited significant differences in both their mean and median values for CAM and C₃ species: aridity index ($p_{\text{mean}} = 0.0023$, $p_{\text{median}} = 0.0300$; Fig. 3a); cloud-cover seasonality ($p_{\text{mean}} = 0.0005$, $p_{\text{median}} = 0.0152$; Fig. 3b); cloud annual cover ($p_{\text{mean}} = 0.0250$, $p_{\text{median}} = 0.0374$); and cloud-cover inter-annual variability ($p_{\text{mean}} = 0.0409$, $p_{\text{median}} = 0.0370$). All variables showed a degree of overlap between CAM and C₃, even those for which there were significant differences in the mean and median values.

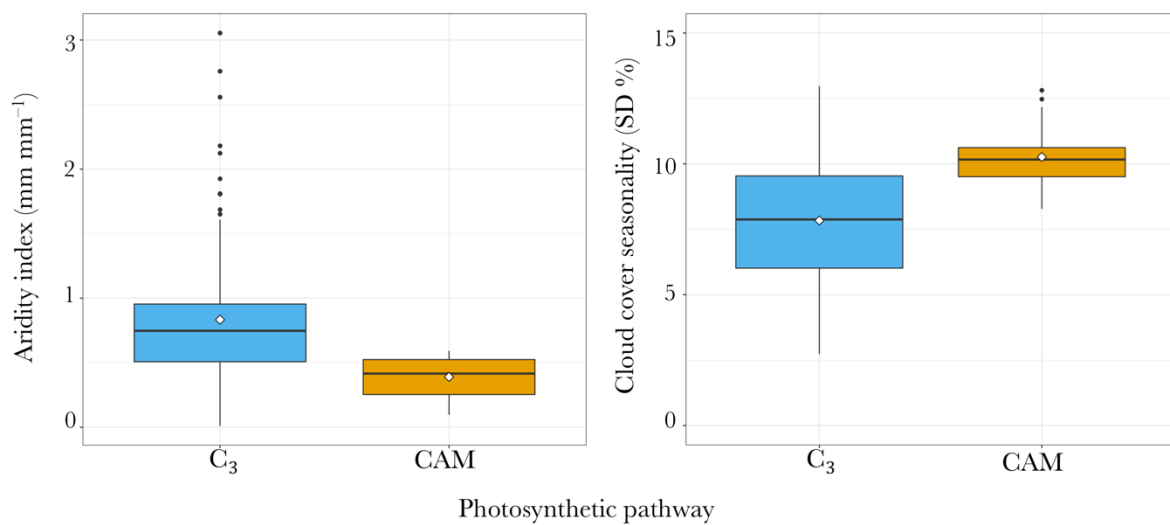


Figure 3. (A) Aridity index (mm mm⁻¹) and (B) cloud cover seasonality (SD %) between C₃ (116 spp.) and CAM (31 spp.) species of *Puya*. Boxes represent data from the 25th percentile to the 75th percentile. The white diamond and the black line in the boxes represent the mean and the median of the data, respectively. The whiskers represent data from the 25th or 75th percentile until the furthest data point that is within 1.5 times the inter-quartile range. The mean and median values difference between CAM and C₃ was calculated using bootstrap test (number of independent interactions = 100 000; sample size = 20; replacement was allowed). There are significant differences between CAM and C₃ species with respect to aridity index ($p_{\text{mean}} = 0.0023$, $p_{\text{median}} = 0.0300$) and cloud cover seasonality ($p_{\text{mean}} = 0.0005$, $p_{\text{median}} = 0.0152$).

When data were partitioned accounting for geographical region, only the central Andes and the Chilean region were considered since these were the only areas where

CAM and C₃ species occur together. Aridity index ($p = 0.0170$) and cloud-cover seasonality ($p = 0.003$) were the only variables that showed significant differences in the mean and median values between CAM and C₃ species.

Random forest

Random forest ordination was conducted to evaluate whether (given the climatic variation) the photosynthetic pathway was classified accurately. The model classified 91.1% of the data accurately (CAM 98.0%; C₃ 90.5%). *P. pyramidata* and *P. isabellina* were two constitutive CAM species that the model misclassified as C₃ species with a probability of 0.75 and 0.54, respectively. The model predicted inaccurately the photosynthetic pathway of 11 C₃ species, of which the majority of the species occur in the southern and central Andes, except for Chilean *P. gilmartiniae* and Colombian *P. bicolor*. Aridity index, cloud-cover seasonality, precipitation seasonality and temperature seasonality were the variables with highest importance in the model (Fig. 4 and Fig. 1 in Chapter 5). Precipitation inter-annual variability, annual mean temperature and annual precipitation were of marginal importance in the model.

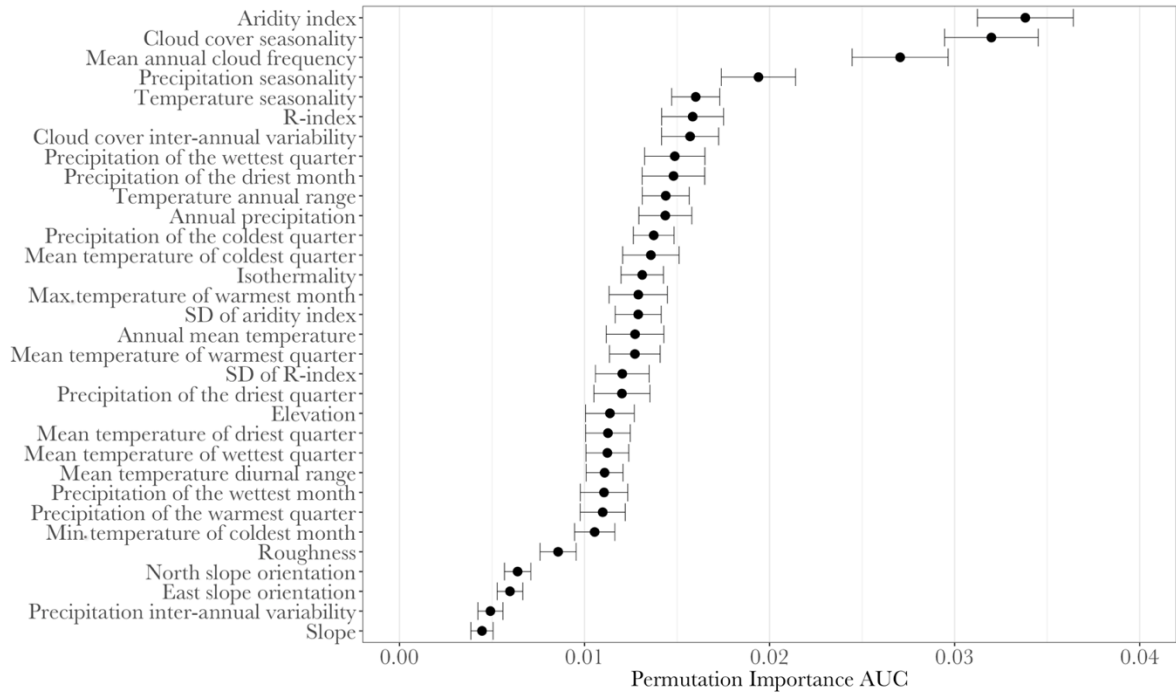


Figure 4. Mean permutation importance of 32 climatic variables in 50 independent Random Forest ordinations in classifying the photosynthetic pathway of 147 species of *Puya*. The permutation importance is a measurement of loss of accuracy before and after the variables in the y-axes were permuted. The higher the index the higher the importance of each variable in the model. The bars represent one standard deviation above and below the mean value. The permutation importance is calculated based on the change of accuracy of the area under the receiver operating characteristic curve (ROC) after permutation of each variable. To see the relationship between the most important variables see Fig. 1 in Chapter 5.

The model using the Central Andes and Chilean species reduces the error rate of classification and classifies 94% of the data accurately. The RF results of the partitioned data with only central Andes and Chilean species are similar to the results found using the whole dataset (Supplementary Fig. 4). The variables making the greatest contribution to classify CAM and C₃ species in the model were: cloud-cover seasonality; aridity index; R- index; temperature annual range; mean temperature of the coldest quarter; and temperature seasonality (Supplementary Fig. 4 and Fig. 1 in Chapter 5). The ordination

using these variables only perform better than the ordination using all the variables listed in Table 2 (Supplementary Table 4). The ordination with the first four variables (aridity index, cloud cover seasonality, mean annual cloud frequency and precipitation seasonality) perform much better classifying CAM species and C₃ species. Therefore, the 93.88 % of 147 species are successfully classified only using aridity index, cloud cover seasonality, mean annual cloud frequency and precipitation seasonality variables.

The RF regression using CAM species only explained 75.9% of the variance. Only temperature-related variables explained the differences within CAM species occurring at different elevations. The most important variables were: mean temperature of the driest quarter; mean diurnal range; mean temperature of the warmest quarter; minimum temperature of the coldest month; and annual mean temperature (Supplementary Fig. 5). The variables with an important role in classifying CAM and C₃ species based on the RF ordination were not relevant in the RF regression (Supplementary Fig. 5).

C₃ and CAM fundamental niche

The fundamental niche of CAM and C₃ species was evaluated by a MaxEnt model using the 32 bioclimatic and topographic variables in South and Central America. The CAM model (AUC = 0.9760, SD = 0.0030) showed potentially suitable areas for CAM species in the southern, central and northern Andes (Fig. 5). Potentially suitable areas of CAM in the northern Andes were separated from potential areas in the central Andes. The discontinuity was located north of the Western Andean Portal (1°–4°N, Fig. 5B). The distribution model for C₃ plants (AUC = 0.9890, SD = 0.0010) showed a continuous distribution across the Andes with no evident discontinuities in high-probability areas, except for the Panama isthmus and the Orinoco region between the Andes and the Guiana shield (Fig. 5C, D).

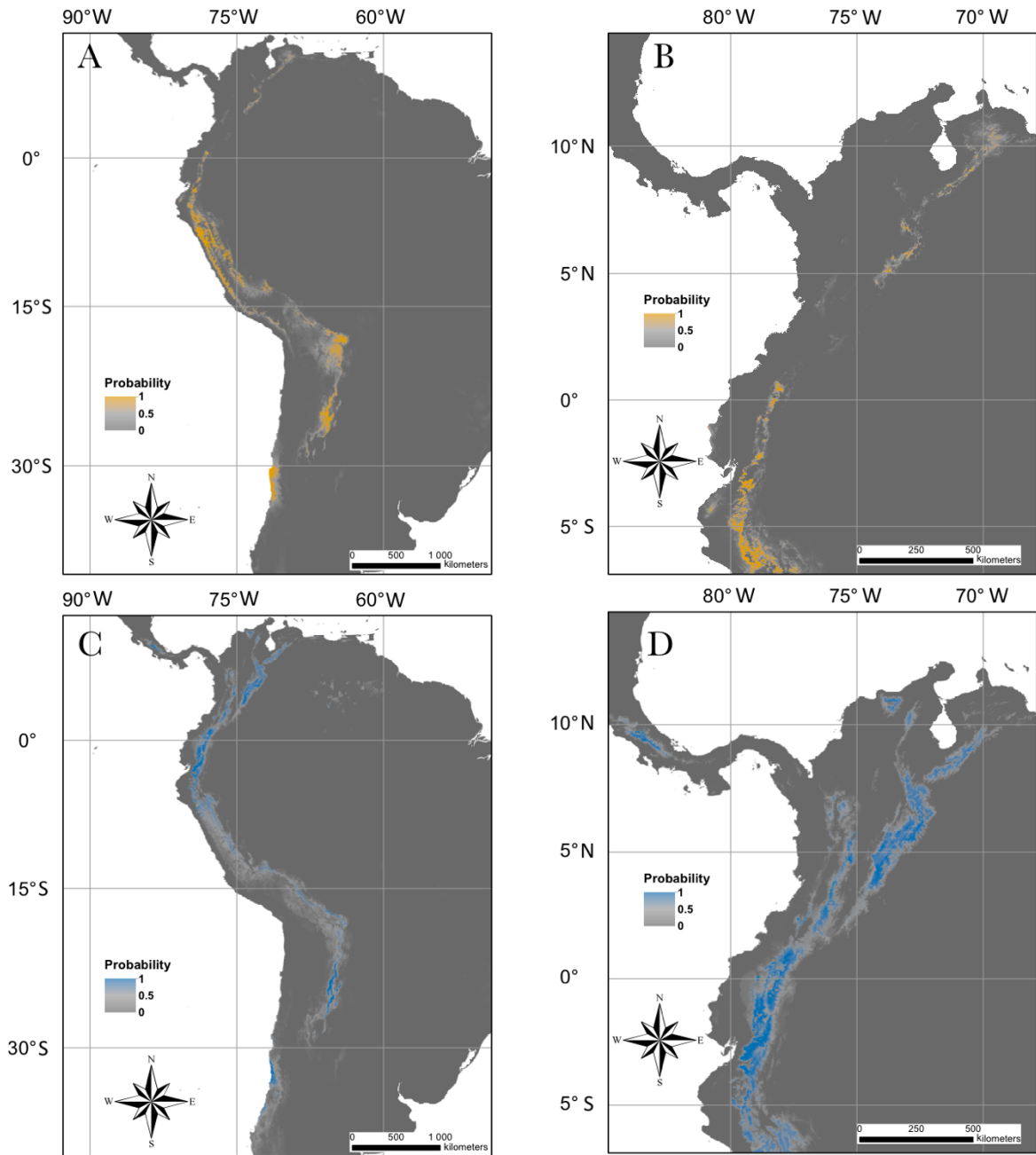


Figure 5. Predicted suitable areas according to a MaxEnt model (32 variables) for CAM species of *Puya* (A and B) and C₃ species of *Puya* (C and D). Panel A and C represent the predicted suitable areas in the whole of the distribution range of *Puya*. Panel B and D represent the predicted suitable areas in the northern Andes. In the panel A and B, the more intense orange the higher the probability of an area to be occupied by CAM species. In panel C and D, the more intense blue the higher the probability of an area to be occupied by C₃ species. There is a discontinuity in the distribution of suitable areas for CAM species in the northern Andes, north of the Equator.

The model also predicted areas with high probability of C₃ occurrence in the Guiana shield and in the Talamancan paramo. While predicted CAM areas in the central Andes were usually on the western side of the Andes, C₃ areas were on the eastern side (Fig. 6). The differences between the fundamental niche of CAM and C₃ species is summarised in Fig. 6. Positive values (> 0.5) indicate areas where CAM is highly predicted, negative values (< -0.5) indicate areas where C₃ is highly predicted, and values close to zero represent overlap or absence of CAM and C₃. The majority of the northern Andes region had negative values. However, there are at least three areas which had areas with positive values (Fig. 6).

In Chile, there is a predicted transition zone between CAM and C₃ suitability ($\sim 30^{\circ}\text{S}$ to $\sim 35^{\circ}\text{S}$). At the northern part of the transition zone, there is only CAM-suitable area. In the south, there is a C₃-suitable area, and in between there is a transition zone with values close to zero, showing both CAM and C₃ likely to be suitable (Fig. 6). The following species occur in the Chilean transition zone: *P. alpestris*, *P. chilensis*, *P. venusta* and *P. coerulea*. There were high-probability areas for CAM species in the Brazilian *caatinga* formation (Fig. 6), which is dominated by terrestrial and lithophytic CAM bromeliads.

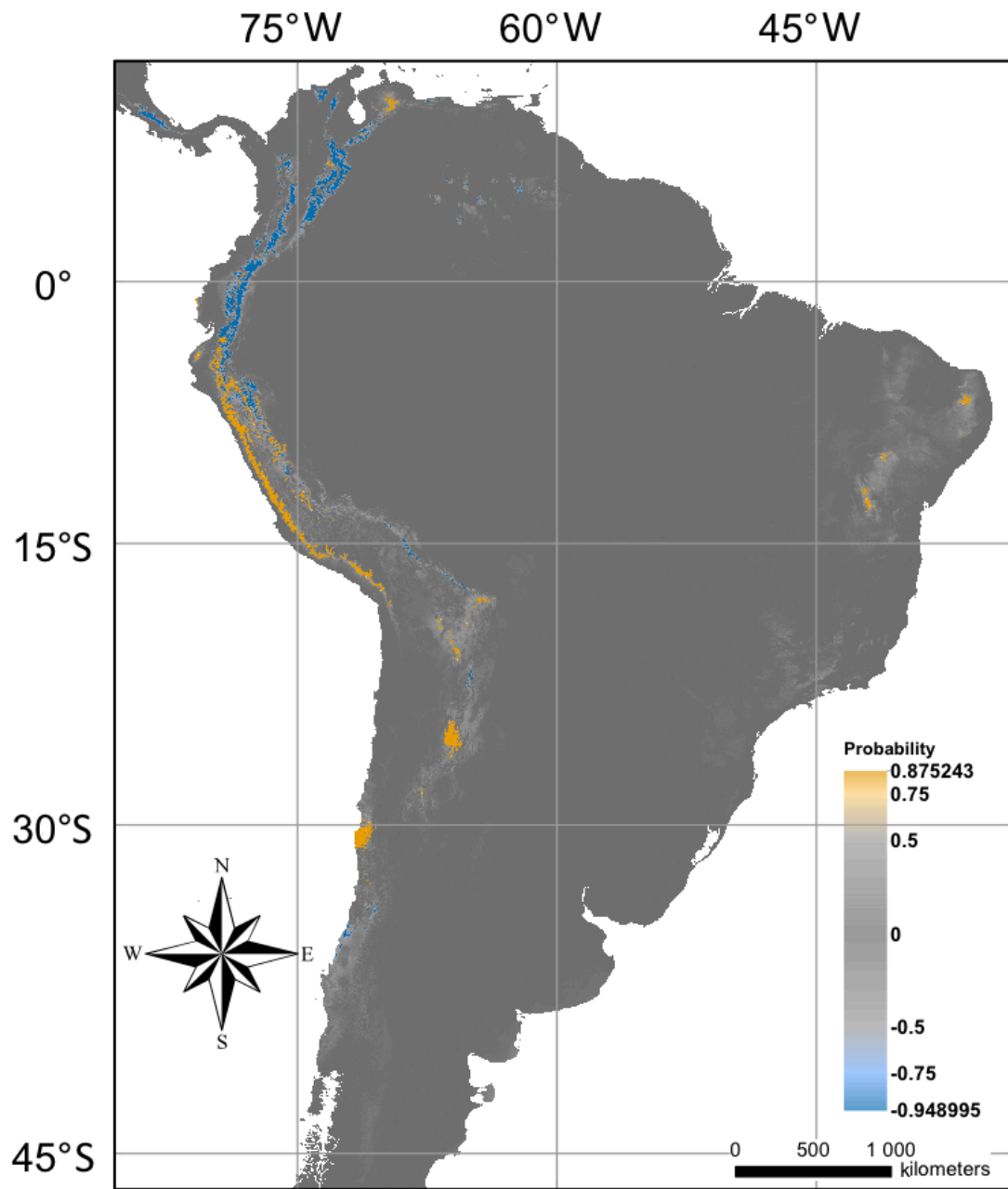


Figure 6. Differences between the predicted fundamental niche of CAM and C₃ species of *Puya* based on MaxEnt model (32 variables). Positive values indicate high suitability for CAM species over C₃ species (orange) and negative values indicate high suitability for C₃ species over CAM species (blue). North of 15°S and east of 45°W, there is a suitable area for CAM in the floristic region of the Caatinga (eastern Brazil).

Ancestral state reconstruction and phylogenetic signal

The *Puya* phylogeny reconstructed in the present study (Supplementary Fig. 6) did not differ from the phylogenetic tree obtained in Jabaily and Sytsma (2013); consequently, the names of the clades are conserved. The early-divergent clade with species in Chile possessing blue flowers is termed the blue clade. The sister clade of the blue clade is referred to as core *Puya* and is identified in the phylogeny with the letter C. Within core *Puya* there are the southern and central Andes clade (SC) and the northern and central Andes clade (NC). Within SC there are two subclades: clade A (mainly Argentinian species) and clade B (Bolivian and Peruvian species). Within NC there is the northern Andes clade (species exclusively occurring in the northern Andes) and the yellow clade (species from Chile that have yellow flowers).

The photosynthetic ancestral character state was reconstructed on the ultrametric tree using maximum likelihood. The log likelihood values of the equal transition rates (ER) and all-different transition rates (ARD) models were -32.8547 and -32.5066 , respectively. The ARD model had the less negative log likelihood value. However, after the maximum likelihood statistical significance test, the null hypothesis cannot be rejected ($p = 0.4041$) and the model with fewer parameters (ER) was chosen. According to the ancestral character state reconstruction, it was not possible to determine whether the photosynthetic pathway of the common ancestor of *Puya* was a constitutive C_3 or a constitutive CAM species (Fig. 7). Indeed, in the SC, yellow and blue clades CAM and C_3 showed multiple transitions between CAM and C_3 . On the contrary, the majority of species belonging to the NC clade were constitutive C_3 . There were no constitutive CAM species in the northern Andes clade. The photosynthetic pathway of the common ancestor of the SC clade is CAM, and therefore C_3 is likely to be a reversal from a CAM

(Fig. 7). The percentage of CAM species in the SC clade is 48%, while in NC clade it is 5%, *P. chilensis* and *P. casmichensis*.

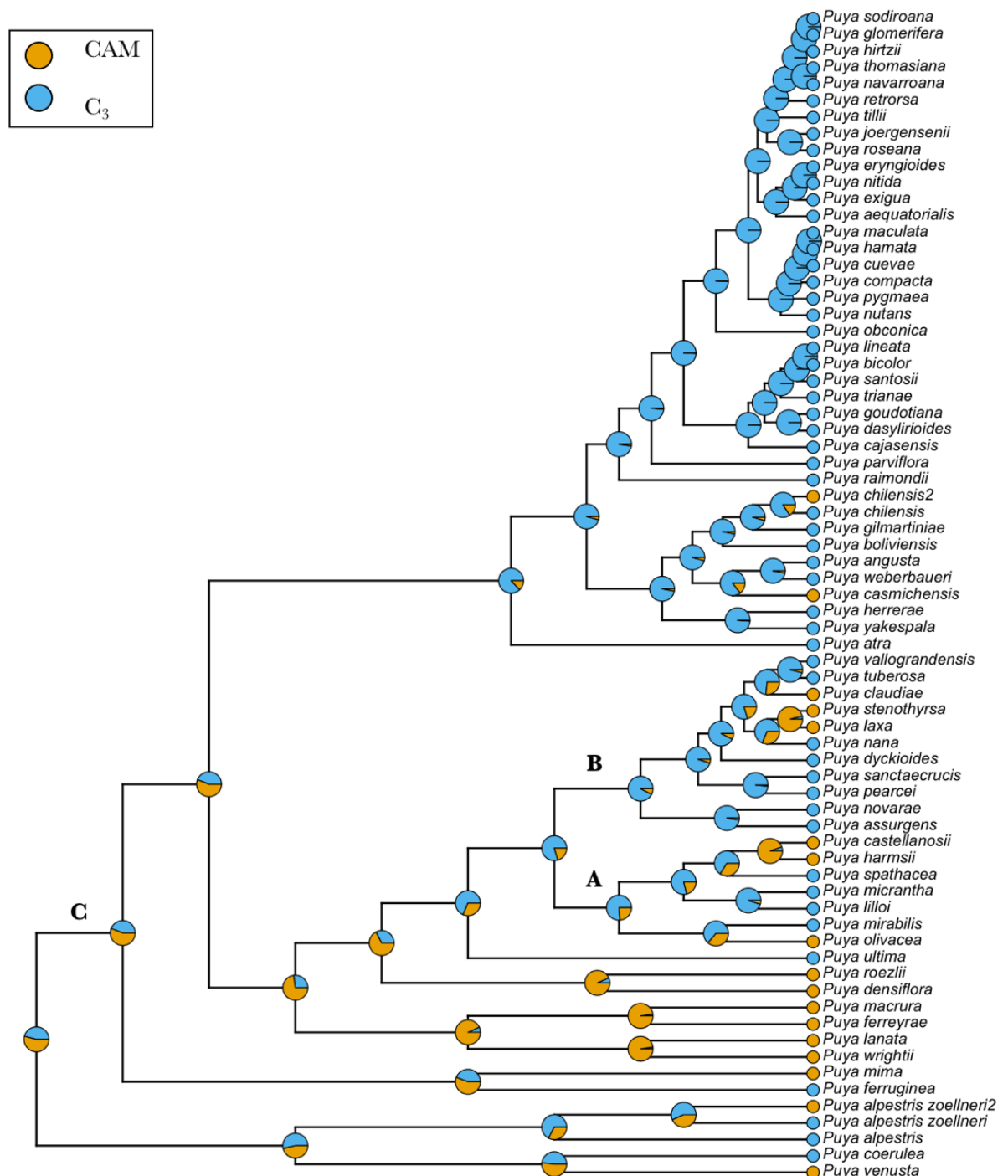


Figure 7. Maximum likelihood ancestral state reconstruction of the photosynthetic pathway in *Puya* based on the relationships supported by posterior probabilities obtained from the Bayesian analysis of AFLP data. The circles on the tips show the photosynthetic pathway according to $\delta^{13}\text{C}$ value (CAM photosynthesis represented in orange and C_3 photosynthesis in blue). Pie charts represent the probability of the ancestral

photosynthetic pathway at each node. Clade A: mainly species from Argentina and Bolivia. Clade B: mainly species from Bolivia and Peru. Clade C: core *Puya* after Jabaily and Sytsma (2013). The photosynthetic pathway of the common ancestor of *Puya*, the common ancestor of the Blue clade and the common ancestor of the Northern and Southern clades are undetermined. However, the common ancestor of the Northern clade was more likely to be C₃ plant and the common ancestor of the Southern clade was more likely to be CAM plant.

There was a strong and significant phylogenetic signal of PC1 ($K = 2.3416$, $p = 0.001$), and a weak but significant phylogenetic signal of PC2 ($K = 0.5794$, $p = 0.020$) and PC3 ($K = 0.7083$, $p = 0.0120$). This mean that two closely related species are more likely to have similar values of PC1 than two randomly species of *Puya*. Therefore, this result suggests that *Puya* exhibited a high degree of niche conservatism, given that these variables summarised the climatic and topographic variation of all *Puya*. Phylogenetic ANOVA showed significant differences between CAM and C₃ for aridity index ($p = 0.0012$), cloud-cover seasonality ($p = 0.0014$), precipitation seasonality ($p = 0.0023$), mean annual cloud frequency ($p = 0.0011$) and PC1 ($p = 0.0012$), but not for temperature seasonality ($p = 0.3532$), PC2 ($p = 0.0963$) or PC3 ($p = 0.5731$).

Discussion

Using the data on geographic distribution, carbon isotope values, climatic differences between CAM and C₃ *Puya* species, plus information on the ancestral state of the photosynthetic pathway and climatic niche phylogenetic signal, the following points will be discussed: (1) The differences in the bimodal distribution of $\delta^{13}\text{C}$ values in *Puya* and in the family Bromeliaceae are influenced by the montane environments where *Puya* mainly occurs. (2) The climatic niche differences between CAM and C₃ species of *Puya* do not correspond to the same differences found in other studies. (3) The overlap between

the CAM and C₃ bioclimatic niche in *Puya* is linked with the high degree of niche conservatism and perhaps with the highly heterogeneous microenvironments characteristic of this genus. (4) The lack of CAM species of *Puya* in the northern Andes is due to lack of suitable habitats or to a biogeographic barrier. (5) Finally, how different CAM evolved in each clade of *Puya*.

$\delta^{13}\text{C}$ values in *Puya*

Carbon isotope distributions showed 78% of *Puya* species are constitutive C₃ plants, or C₃ plants in which there is at most a very small amount of dark CO₂ fixation. There is a shift in the peak of the $\delta^{13}\text{C}$ values distribution between the C₃ species of *Puya* (−23.9 ‰) and C₃ bromeliads as a whole (−26.7 ‰: Crayn *et al.*, 2015). This shift is correlated with elevation, since the high-elevation C₃ species tend to have less negative $\delta^{13}\text{C}$ values (Supplementary Fig. 2). Körner *et al.* (1988, 1991) showed that high-elevation C₃ species have higher ¹³C concentrations in their tissues compared to low-elevation species. A similar pattern is found in the Bromeliaceae (Crayn *et al.*, 2015). The phenomenon is explained by the reduction of CO₂ partial pressure and temperature. At higher elevations, the ratio of internal to external partial pressure of CO₂ (p_i/p_a) is lower than in low-elevation environments (Körner *et al.*, 1988). Under low p_i/p_a , RuBisCO discriminates less against ¹³C than under high p_i/p_a (Figure 1 in Farquhar *et al.*, 1989). Accordingly, there is greater ¹³C abundance in the tissue of a C₃ plant at higher elevations. Moreover, Cernusak *et al.* (2013) proposed that lower temperatures increase water viscosity, thus reducing the ratio of intercellular to ambient CO₂ mole fractions (c_i/c_a), and ultimately increasing the relative ¹³C concentration in the plant tissue. Indeed, while the majority of *Puya* species occur at elevations ranging from 2500 to 3500 m (Jabaily and Sytsma, 2010), most C₃ bromeliads occur from 0 to 2000 m (Crayn *et al.*, 2015).

The shift in the peak of the $\delta^{13}\text{C}$ values of the C_3 *Puya* species suggests that the photosynthetic pathway of at least some species with values around -20‰ could be weak CAM (Supplementary Fig. 1). Of these species, it is impossible to determine whether they are constitutive CAM or weak CAM purely by reference to their $\delta^{13}\text{C}$ values. To resolve this matter, it is necessary to measure titratable acidity and nocturnal CO_2 uptake.

The focus of the current investigation is on constitutive CAM ecology and evolution. Within Bromeliaceae, few species are so far known to show weak CAM, for example in *Tillandsia cretacea* (Pierce *et al.*, 2002a); *Werauhia sanguinolenta* (Pierce *et al.*, 2002a, Beltrán *et al.*, 2013); *Puya floccosa* (Herrera *et al.*, 2010). However, it is not established whether this is explained by the rarity of weak CAM as an intermediate state, or because the taxon sampling is usually reduced in detailed physiological studies.

Not all C_3 *Puya* species occur at high elevations. In particular, C_3 species from the blue and yellow clades occur at low elevations (<1000 m, Gómez-Romero *et al.*, 2009; Zizka *et al.* 2009, Grau *et al.*, 2010) and have carbon isotope values close to -20‰ . These species diverge from the general tendency of the genus, in that these species show less negative $\delta^{13}\text{C}$ values at lower elevations. Moreover, *P. chilensis* (yellow clade) and *P. alpestris* (blue clade) show a considerable range of $\delta^{13}\text{C}$ values (Table 1). This result suggests that some populations are constitutive CAM (values less negative than -20‰) and some are non-constitutive CAM (values more negative than -20‰). The CAM populations of both species occur at the southern limit of the Atacama Desert (*P. alpestris* ssp. *zoellneri* $29^\circ 59' \text{S}$; *P. chilensis* $30^\circ 8' \text{S}$) whilst C_3 populations occur in the more temperate parts of Chile (Fig. 6, Supplementary Fig. 7). Variability in the $\delta^{13}\text{C}$ values within the populations is correlated with aridity index, precipitation seasonality and latitude (Supplementary Fig. 8). It is possible that CAM evolved as a consequence of colonisation of more arid environments in northern Chile. This has been reported in *P.*

chilensis (Quezada *et al.*, 2014; 2017). However, Quezada *et al.* (2014) used a threshold of -22‰ to distinguish between CAM and C₃ populations, only one population exhibited a $\delta^{13}\text{C}$ value around -19‰ . It was not until the survey of Crayn *et al.* (2015) where it was demonstrated that *P. chilensis* had also populations with $\delta^{13}\text{C}$ values typically found in other CAM plants (Supplementary Fig. 7–8). This is the first report of constitutive CAM populations in *P. alpestris*.

Climatic niche differences between CAM and C₃ species of *Puya*

The climatic niche of CAM *Puya* species (as opposed to C₃ species) is defined by low aridity index (indicating more arid areas) and high cloud-cover seasonality. CAM *Puya* species occupy areas with aridity index values in the range of 0.03 to 0.65 (Fig. 3A), i.e. arid to dry sub-humid environments (Middleton and Thomas, 1997). In contrast, the majority of C₃ species occupy humid environments. One extreme exception is *P. boliviensis* which occurs in the hyper-arid region on the southern limits of the Atacama Desert. *P. boliviensis* occupies the least xeric microenvironments in the *Lomas* formation, where dew-water is condensed on the rosette leaves and redirected to the soil where it can be absorbed (Rundel and Dillon, 1998; Janeba, 2017).

CAM species occur in areas with higher cloud cover seasonality than areas where C₃ species occur (Fig. 3). Mean annual cloud frequency could be interpreted as a variable correlated with precipitation and light intensity; and cloud cover seasonality with the seasonality of the precipitation and light intensity. Areas with high cloud cover seasonality experience at least one dry season, usually with clear skies, and another season of high precipitation and non-diffuse sunlight. In these areas, CAM may have an ecological advantage over C₃ plants. During the dry season, C₃ species would lose more water than CAM species as a consequence of the higher water use efficiency of CAM plants over C₃ plants in identical conditions (Winter *et al.*, 2005, 2015). Furthermore, high

photosynthetically-active radiation is correlated with high rates of nocturnal CO₂ fixation in the subsequent dark period and accumulation of malic acid in CAM plants (Nobel and Hartsock, 1983). In C₃ plants, high photosynthetically-active radiation could increase primary productivity in the short term (Graham *et al.*, 2003), but in the long-term their productivity decreases due to water stress (Malhi *et al.*, 2009; Meir *et al.*, 2009). Consequently, prolonged and predictable dry seasons could be ecological filters that benefit mainly CAM species. This is the case in the central Andes, where the rainy season extends from 4 to 9 months, depending on elevation and latitude, and is highly predictable (Garreaud, 2009). This is not the case of the northern Andes where there are two dry seasons in a year (Garreaud, 2009), and perhaps the extension of the drought stress is not long enough to be ecological filters that benefit mainly CAM species. This may explain the prevalence of CAM *Puya* species in the central Andes, and the lack of CAM species in the northern Andes.

Aridity index and cloud-cover seasonality differ from variables suggested by Ellenberg (1981; Ellenberg's model) and Heyduk *et al.* (2016a; Heyduk's model). Aridity index is defined as the ratio between annual precipitation and potential evapotranspiration. This compound variable, which reflects the amount of water potentially available in the soil for the plant, is very important in highly heterogeneous environments such as the Andes, instead of using only precipitation or only potential evapotranspiration. For example, in circumstances where precipitation is high but potential evapotranspiration is even higher, aridity index would reflect an environment with potential water deficit for the plants (arid). In contrast, areas with relatively low precipitation and low potential evapotranspiration would be considered relatively humid areas using the aridity index (humid).

Males and Griffiths (2017) showed that the relationships between drought-resistant traits, aridity index and precipitation seasonality were correlated by the functional type of bromeliads. The aridity index used in Males and Griffiths (2017) is defined as the ratio between mean annual precipitation to mean annual temperature. The precipitation-temperature aridity indexes are useful to study paleo climatic differences, mainly because potential evapotranspiration is not easy to calculate in paleo climatic data (Quan *et al.*, 2013). Furthermore, different indexes involving precipitation and evapotranspiration have been used to describe the aridity. These indexes are more informative and preferable because they include temperature and radiation information (Quan *et al.*, 2013). Maire *et al.* (2015) used the ratio between potential evapotranspiration and precipitation. This index has limitations, especially in places where the mean annual precipitation is zero or close to zero. Moncrieff *et al.* (2015) used the difference between potential evapotranspiration and mean annual precipitation as a proxy for aridity. However, the linearity correlation of this potential variable with the majority of precipitation variables could reduce the potential variables to use into a model, and this scenario is not ideal. The Males and Griffiths (2017) aridity index could be problematic for terrestrial plants because the radiation, the run-off of water and the evaporation factors were ignored. However, this proxy might be useful to compare different types of bromeliads (epiphytes and terrestrials), mainly because epiphytes are independent of the soil. The aridity index used in the present study incorporates precipitation, temperature and radiation (CGIAR; Zomer *et al.*, 2007, 2008; Trabucco and Zomer, 2010). The importance of variables such as annual cloud frequency and cloud cover seasonality need to be explored in other taxa containing CAM and C₃ species.

Ellenberg's model

Ellenberg (1981) suggested that annual precipitation < 500 mm partly defines the niche of stem-succulent plants. However, the current study shows annual precipitation does not define the niche of CAM and C₃ *Puya* species. Indeed, the CAM species *P. pyramidata*, *P. pitcairnioides* and *P. densiflora* occur in places where the annual precipitation is > 1000 mm. The contrast between Ellenberg's predictions and the observed distribution of *Puya* species may be because Ellenberg (1981) did not include saxicolous plants while all the *Puya* species mention above are saxicolous. Possibly the majority of precipitation that is received by the *Puya* species mention above will run off the cliffs. Therefore, unlike the aridity index, annual precipitation does not reflect the amount of water available to plants in the soil.

Ellenberg (1981) identified low inter-annual variability of precipitation as important for defining the stem-succulent climatic niche. However, in the current study no significant differences in inter-annual variability of precipitation were found between CAM and C₃ *Puya* species, this being the second-least-important variable in the RF ordination. This pattern is explained because many C₃ species occur in areas where the precipitation inter-annual variability is low as well as the precipitation (e.g. *P. assurgens*, *P. dyckiioides*, *P. minima*, *P. volcanensis*, *P. bermejiana*); Ellenberg's model predicts CAM stem-succulent plants in areas with low precipitation (< 500 mm year⁻¹) and low inter-annual variability of precipitation (< 5). Furthermore, Chilean CAM species (*P. alpesstris* ssp. *zoellneri* and *P. chilensis*) and Peruvian CAM species (*P. roezlii*) occur in areas where the precipitation was less than 500 mm year⁻¹ at the southern and northern limits of the Atacama Desert, respectively. The inter-annual precipitation variability of these species is very high, reaching values typically found for non-stem-succulent CAM species according to Ellenberg's model. Dew (rather than precipitation) is the main water source

for bromeliads in this region, and in the northern parts of the Atacama Desert (Rundel and Dillon, 1998). The present investigation suggests that a better model for understanding differences between the climatic niche of CAM and C₃ plants should include aridity index and cloud cover-related variables.

Males and Griffiths (2017), who investigated whether inter-annual precipitation variability was correlated with drought resistance-related traits in Bromeliaceae, suggested that precipitation seasonality (bio15) was equivalent to inter-annual precipitation variability. However, precipitation seasonality is a measure of intra-annual precipitation (Hijmans *et al.*, 2005; O'Donnell and Ignizio, 2012). Precipitation seasonality and inter-annual precipitation variability are different variables, which are not necessarily correlated. For example, the southern limits of the Atacama Desert have low precipitation seasonality and high inter-annual precipitation variability (Rundel and Dillon, 1998), whilst the equatorial Andes (~1°S to ~4°N) has low precipitation seasonality and low inter-annual precipitation variability (Garreaud, 2009). In the present study, precipitation seasonality is an important variable in the RF, showing significant differences in the mean value between CAM and C₃.

Heyduk's model

Heyduk *et al.* (2016a) showed that temperature-related variables were the most important explanatory variables distinguishing the CAM and C₃ bioclimatic niches in Agavoideae. In contrast, the present study of CAM and C₃ species of *Puya* found the most important variables were not directly linked to temperature. One possible explanation for the difference is that Heyduk *et al.* (2016a) did not consider cloud cover-related variables, aridity index, or topographic variables. Heyduk's model attributed the lower contribution of precipitation-related variables to the fact that C₃ and CAM species of Agavoideae coexisted in arid areas. However, precipitation alone is not a proxy for aridity.

Another reason for the discrepancy might be the broader range of elevations where *Puya* occurs, irrespective of the photosynthetic pathway (Supplementary Figure 2). This makes any temperature variable (which is usually correlated with elevation) less likely to define the CAM and C₃ niches in *Puya*. Agavoideae mainly occur in the southwestern deserts of North America (typically dry areas) and the semi-arid regions of Central America, where the climatic conditions are more spatially homogeneous than in the Andes. Furthermore, the Andes form a mosaic of dry and humid areas, changing abruptly (Engel *et al.*, 2014).

More fundamentally, Heyduk's model has two problems that potentially contribute to biased conclusions. Firstly, the significance of the variables identified by a PCA were not tested by incorporation into a climatic model. The present study has shown that important variables in the PCA are not necessarily variables that define the CAM and C₃ niches. For example, temperature annual range is an important PCA variable but has a very low importance in the RF ordination, which classifies the niche of CAM and C₃ species given the whole climatic variance. And secondly, there was no reference to spatial autocorrelation, which can lead to bias in the analysis of a species distribution model (Araujo *et al.*, 2005). The present study accounted for spatial autocorrelation by merging multiple records of each species into single points.

Topographic differences between CAM and C₃

Topographic variables did not contribute significantly in the RF model. One reason is the scale of these variables. Finer-scale analyses may produce better results, but general patterns may be more easily revealed at larger scales. For example, the variable east slope orientation would be more useful in terms of whether the species was located on the east or west cordillera than in the east or west face of a 1 km² mountain. Unfortunately, all topographic variables have a scale of 30 arc-sec (i.e. 1 km²);

consequently, there is a discrepancy between the relative importance of topographic variables in the RF and the fundamental niche of CAM and C₃.

The Andes is a highly heterogeneous and diverse environmental landscape (Tovar *et al.*, 2013; Engel *et al.*, 2014). The different landscapes can change from humid to dry within a few kilometres due to changes in wind direction, topography and reduction in cloud cover (Killien *et al.*, 2007; Garreaud, 2009; Engel *et al.*, 2014). Given the classification of the Andean biomes proposed by Tovar *et al.* (2013) together with the results from the present study, CAM species of *Puya* occur in the xeric pre-puna, xeric puna, montane scrubland and seasonally dry tropical montane forest. In contrast, C₃ species occur in the humid puna, montane scrubland, and the paramo. The present study identified a ‘characteristic floristic element’ of the biomes of paramo, humid puna and xeric puna through the photosynthetic pathway of the *Puya* species that occur in each area.

CAM and C₃ niche overlapping

Despite the climatic niche differences between the CAM and C₃ species of *Puya*, there is some overlap between niches. Complete overlap occurred in climatic variables that are unimportant in distinguishing between CAM and C₃ species. Partial overlap occurred in variables that are important in the RF and according to the bootstrap analysis. All variables showed a degree of overlap between CAM and C₃ climatic niches. One explanation for this overlap is the high degree of niche conservatism in *Puya* and the putatively repeated transitions between C₃ and CAM photosynthesis in *Puya*. A high degree of niche conservatism means that closely related species tend to occupy more similar niches than non-related species. This increases the potential overlap between C₃ and constitutive CAM sister species, especially given that CAM evolved (or there was reversal to C₃) multiple times, as indicated by the ancestral state reconstruction.

Another explanation is that CAM and C₃ *Puya* species could coexist in the same environment but be distributed in different microhabitats. This pattern is common in CAM and C₃ epiphytic bromeliads, which occupy different areas of the canopy according to their photosynthetic pathway (Griffiths and Smith, 1983; Smith, 1989). CAM epiphytes occupy the most exposed areas of the canopy which, even in tropical forests, have low water availability (Zotz and Ziegler, 1997; Silvera and Lasso, 2016). In contrast, C₃ epiphytes occur in the less exposed part of the canopy and the lower canopy. This pattern could be reflected in the montane scrubland where both CAM and C₃ *Puya* species coexist, with CAM species occupying more exposed habitats (i.e. saxicolous), whilst C₃ species occupy less exposed environments or colonize substrates close to a water source.

Weak CAM *Puya* species potentially increase the overlap between CAM and C₃ climatic niches. These species could occupy different environments than C₃ species and possibly the niche would be more similar to the niche of constitutive CAM plants. Nevertheless, there are only a few species of Bromeliaceae that show weak CAM (Pierce *et al.*, 2002a, Winter *et al.*, 2015). Only two *Puya* species have been reported as weak CAM, *P. chilensis* and *P. floccosa* (Herrera *et al.*, 2010; Quezada *et al.*, 2014, 2017). It is likely that weak CAM species only represent a few points in the biogeographical analysis and are therefore unlikely to bias the analysis. Despite these possible reservations, there is a strong signal that differentiates the CAM and C₃ niches.

Why CAM species do not occur in the Northern Andes

Three reasons could explain the absence of CAM species of *Puya* in the northern Andes: (i) survey incompleteness; (ii) biogeographical selectivity; and/or (iii) absence of suitable CAM habitats in the northern Andes. Survey incompleteness can be dismissed. The present study determined the $\delta^{13}\text{C}$ values of 70% of *Puya* species from Colombia,

Venezuela and Costa Rica; all showed values typical of C₃ plants rather than constitutive CAM.

Initially, the Western Andean Portal (WAP, located between $\sim 3^{\circ}$ and 5°S) was considered a potential selective biogeographic barrier for CAM species of *Puya*. This is mainly because the northernmost CAM *Puya* species (*P. lanata*) occurs in the northern limits of the WAP, and CAM species are absent further north. The WAP is an important biogeographic barrier for Andean plants, birds and mammals (Antonelli *et al.*, 2009). From the Eocene to middle Miocene ($\sim 40\text{--}12$ Ma), intermittent Pacific marine incursions limited the dispersal of plants and animals north and south of the WAP (Antonelli and Sanmartin, 2011). However, from the middle Miocene onwards, the Pacific marine incursions stopped due to the uplift of the northern Andes. This uplift allowed bidirectional migration. The split between Puyoideae and Bromelioideae is estimated to have occurred around 10–8 Ma ago, and *Puya* evolved in the southern Andes around ~ 9 Ma ago (Jabaily and Sytsma, 2010, 2013; Givnish *et al.*, 2014), when the WAP allowed bidirectional migration. Furthermore, the age of the northern Andes *Puya* clade is ~ 1.58 Ma (Madriñan *et al.* 2013; 2015). Thus, when *Puya* colonised the northern Andes, the WAP was not a barrier, and suitable areas for CAM species existed across the WAP.

The potential niche of CAM and C₃ species of *Puya* was evaluated to understand the lack of CAM species in the northern Andes. There is a considerably large gap (~ 444 km, 4° in latitude) between suitable areas for CAM between the WAP and northern Andes. The gap is located in the northern part of Ecuador and in the south of Colombia, where the northern Andes is divided into three cordilleras (Duellman, 1999). This discontinuity is not found in the C₃ potential areas (Fig. 5D), but could act as a significant biogeographical barrier for CAM species, taking into account the size of the gap and the low dispersal abilities of *Puya* species. Moreover, the region around this gap is

characterised by almost constant high precipitation over the year due to orographic precipitation and orographic cloud drizzle (Garreaud, 2009). This pattern is the product of the moist air from the Amazon and from the Gulf of Panama, converging in the *Nudo de Pasto* region (Duellman, 1999; Vuille *et al.*, 2000; Bendix *et al.*, 2005). These conditions have been maintained since the closure of the Panama Isthmus and the consequent modification of ocean currents, even though the final closure date is still debated (Ibaraki, 1997; Montes *et al.*, 2015; O'Dea *et al.*, 2016; Jaramillo *et al.*, 2017). Recent geological evidence suggests the Central American seaway vanished around the middle of the Miocene (13–15 Ma; Montes *et al.*, 2015), contradicting biological and geological evidence that places the closure at ~3 Ma (O'Dea *et al.*, 2016). The *Nudo de Pasto* has therefore been an area of constant and high precipitation since at least ~3 Ma. Colonisation of the Northern Andes by *Puya* happened around ~1.58 Ma ago (Madriñan *et al.*, 2013) and at that time the *Nudo de Pasto* was a region of constant rainfall (Vuille *et al.*, 2000; Londoño *et al.*, 2014) and constant cloud cover (Vuille *et al.*, 2000; Bendix *et al.*, 2005). Given these characteristics, the *Nudo de Pasto* region is highly suitable for C₃ species of *Puya*. Moreover, CAM species may have been at a positive disadvantage compared with C₃ species, given the extra energetic cost of CAM and the unfavourable environmental conditions at the *Nudo de Pasto*, where there is no direct sunlight and no obvious water stress. Therefore, the *Nudo de Pasto* could act as a biogeographic barrier for CAM species but not C₃ species of *Puya*.

The northern Andes experiences two dry seasons each year (Garreaud, 2009), whilst the central Andes only has one dry season that can extend for 9 months. The two relatively short dry seasons experienced in the northern Andes could limit the ecological advantage of constitutive CAM plants, because short dry seasons produce less water stress

for the plants. Therefore, bimodal precipitation seasonality in the northern Andes imposed an extra limitation for CAM species to compete against C₃ *Puya* species.

Evolution of CAM in the genus *Puya*

Uplift of the central and northern Andes and aridification processes may have shaped the evolution of constitutive CAM in *Puya*. The central Andes uplift (10–6 Ma; Garzzone *et al.*, 2008) is congruent with the origin of *Puya* and its colonisation of the Andes. Eight to 5 million years ago a global aridification, associated with the increasing abundance of C₄ grasslands, was possibly linked with origin of constitutive CAM in several groups (Arakaki *et al.*, 2011; Edwards and Ogburn, 2012). In South America, periods of strong aridification took place 6.1 to 5.2 Ma and 3.6 to 3.3 Ma (Amidon *et al.*, 2017). During this time, the evolution of constitutive CAM may have been triggered in the two subclades of southern and central Andes clade (SC) clade in *Puya*.

Pleistocene glaciation-interglaciation cycles expanded and reduced vertically the distribution of puna and the paramo, acting as corridors when expanded or refuges when contracted (Sklenář *et al.*, 2011; 2014). The Milankovitch oscillations during the last 100 Ma, in particular during the Pleistocene, have changed speciation and extinction rates (Dynesius and Jansson, 2000) and expanded and contracted humid and dry puna in the central Andes during Plio-Pleistocene and the mid-Holocene (Simpson, 1983; Ledru *et al.*, 2013, Engel *et al.*, 2014). Finally, uplift of the northern Andes around 5 to 2 Ma (Gregory-Wodzicki, 2000) allowed the high-elevation C₃ species to colonise the north.

The present study established that there was a high degree of niche conservatism in the genus *Puya*. This is evidenced by the positive and significant *K*-values for the first three components in the PCA. Niche conservatism is in great part responsible of the differentiation of niches at local and global scale in plants (Donoghue, 2008). The high-degree of niche conservatism has been proved in high-elevation plant genera such as

Valeriana and *Calceolaria* (Sklénář *et al.*, 2011). On the other hand, the phylogenetic ANOVA showed that the evolution of CAM or the reversal of C₃ was linked in a changed in the PC1, aridity index, cloud-cover seasonality, mean annual cloud frequency and precipitation seasonality. These results suggest that the evolution of constitutive CAM was correlated with change in the climatic niche. Therefore, constitutive CAM had an adaptive value in terms of changing the climatic niche, despite the high degree of niche conservatism. Similar results were found in the Montiaceae related with the growth form (Ogburn and Edwards, 2015).

The evolutionary history of the CAM in *Puya* is complex with multiple gains and losses. For example, the common ancestor of the SC clade is likely to have been a constitutive CAM plant, with reversals from CAM to C₃ occurring within the clade, and with possible secondary origins of CAM (Fig. 7). In the blue clade (low-elevation species), it is uncertain whether constitutive CAM evolved from a C₃ ancestor or whether C₃ arose by reversion from a CAM ancestor. In contrast, in the NC clade it is more likely that constitutive CAM evolved from a C₃ ancestor. The percentage of constitutive CAM species in the NC clade is 5%, with CAM appearing only in *P. chilensis* and *P. casmichensis*. In contrast, 48% of species in the SC clade are constitutive CAM. It is possible that the clades likely to evolve CAM are clades in which the most recent ancestor was not necessarily a constitutive C₃ plant. Consequently, all the biochemical and morphological changes from C₃ to constitutive CAM are more likely and more readily evolved.

It remains uncertain whether the common ancestor of *Puya* was a constitutive CAM or a C₃ species. However, circumstantial evidence suggests that the photosynthetic pathway of the common ancestor of *Puya* and Bromelioideae could be weak CAM. In Bromelioideae, the early-divergent genus *Ochagavia* contains weak CAM species, and the early-branching monotypic genus *Deinacanthon* is constitutive CAM; the former is

distributed in the south and central Chile while the latter is distributed in Paraguay. This circumstantial evidence suggest that it is likely that the geographical origin of the subfamily Bromelioideae and Puyoideae could be in central Chile. Given that CAM is a complex physiological trait, and that at least some of the earlier-divergent species of *Puya* are also CAM, it is possible that the last common ancestor of Bromelioideae and Puyoideae could have had some capacity to fix nocturnal CO₂ and accumulate organic acids overnight. An intermediate photosynthetic pathway in between C₃ and CAM could explain the evolution of constitutive CAM in both subfamilies once species colonised environments where water stress was high.

Phylogenetic data indicate that the Andes was colonised by *Puya* twice from south to the north (Jabaily and Sytsma, 2013): once by the common ancestor of the SC clade, and secondly by the common ancestor of the NC clade. Both clades contain species currently from Argentina, Bolivia, Peru and Ecuador. The species from the NC clade are restricted to the eastern cordillera in the central Andes, while the species from SC clade occur at the eastern and western cordillera. The eastern cordillera received more precipitation than the western cordillera in the central Andes, and it is considered more mesic (Garreaud, 2009). Given this evidence, it is possible that the colonisation route from south to north of the common ancestor of the NC clade was restricted to the humid areas in the eastern cordillera.

CAM evolved in the early-divergent blue clade in the *P. alpestris*/*P. coerulea* species complex, particularly in populations from the species *P. alpestris* ssp. *zoellneri* (syn. *P. berteroniana*) and *P. venusta*. It is likely that constitutive CAM evolved as the ancestral populations colonised more xerophytic areas in the north of Chile from the more mesophytic and temperate areas of central Chile. There is a significant correlation between latitude, precipitation seasonality, aridity index and carbon isotope values in all

Chilean species (Supplementary Fig. 8). Constitutive CAM populations of *P. alpestris* ssp. *zoellneri* are present only at the southern limit of the Atacama Desert. This subspecies has been recently described by Ziska *et al.* (2013). Previously *P. alpestris* ssp. *zoellneri* was *P. berteroniana*, and *P. alpestris* ssp. *alpestris* was *P. alpestris*. In the recent taxonomic revision based on morphological and AFLP data Ziska *et al.* (2013) concluded that both taxa are part of the same species. It is remarkable because in the taxonomic assessment of *P. alpestris* species complex the photosynthetic pathway was not incorporated. Constitutive CAM populations are only present in *P. alpestris* ssp. *zoellneri*. The only described species that have populations showing constitutive CAM and perhaps constitutive C₃ are *P. chilensis* (Quezada *et al.*, 2014) and *P. alpestris* (the present study, Supplementary Fig. 7). It is necessary to investigate if this is a pattern unique in the species from Chile or other *Puya* species show such a wide range in their photosynthetic pathway at population level. The carbon isotope values of the populations from the mesophytic areas of central Chile are very negative, suggesting constitutive C₃ (supplementary Fig. 7). The photosynthetic pathway of the common ancestor of the blue clade is undetermined.

The evolution of CAM in the SC clade showed frequent reversals between C₃ and CAM. The photosynthetic pathway of the internal nodes in the SC clade showed a reversal from CAM to C₃, possibly as a consequence of colonisation of relatively less arid environments in the eastern cordillera (wet puna) given that the C₃ species from the SC clade occur in the eastern cordillera (Fig. 1). Colonisation of the Andes may have taken two routes; subclade A colonising northern Argentina, and subclade B colonising Bolivia, Peru and the south of Ecuador. The clade that colonised northern Argentina evolved CAM in high-elevation environments in the species *P. olivacea*, *P. castellanosi* and *P. harmsii*. The subclade B evolved constitutive CAM in species from mid-elevation and seasonally dry environments such as *P. stenothyrsa*, *P. claudiae* and *P. laxa*. The less speciose

and earlier-divergent clades contain constitutive CAM species occurring mainly at mid-elevations in the dry valleys of Peru and Bolivia. The high-elevation CAM species from Bolivia and Peru were not included in the current phylogeny and therefore it is not possible to speculate on whether CAM in high elevations evolved multiple times within the genus. The only high-elevation CAM species included in the current phylogenetic reconstruction are *P. harmsii* and *P. castellanosi*.

The possibility of neither constitutive CAM nor constitutive C₃ being the photosynthetic pathway of the last common ancestor of the SC clade must be considered. The common ancestor could be a weak CAM plant, showing the capacity for some nocturnal CO₂ fixation, but with this making a small or insignificant contribution to carbon gain by the plant. If so, constitutive CAM could evolve more readily, because weak CAM species would already possess the enzymatic and regulatory machinery to fix (or recycle) CO₂ at night time *via* phosphoenolpyruvate carboxylase. Furthermore, if the ancestral photosynthetic pathway was weak CAM, it might be comparatively easy to lose or maintain weak CAM in relatively humid areas, depending on the facultative nature of weak CAM. Once the weak CAM ancestor colonised relatively drier areas, the evolution of full, constitutive CAM could be triggered. Potentially, at least some of the C₃ species from the SC clade may indeed be weak CAM species, so more detailed physiological measurements should be performed in future to evaluate whether the *Puya* species from this clade with carbon isotope values more negative than -20‰ are constitutive C₃ or weak CAM plants.

The yellow and blue clades apparently experienced multiple hybridization events, as inferred from nuclear and chloroplast DNA sequence information (Jabaily and Sytsma, 2010; Schulte *et al.*, 2010). It is possible that CAM evolved independently in both species when they colonised the xeric environments at the southern limits of the Atacama Desert.

However, given the hybrid history of the two species, it is possible that CAM was transferred horizontally from the blue clade to *P. chilensis*, or *vice versa*. One possible scenario is that CAM populations from the blue clade hybridized with C₃ populations of *P. chilensis* in the northern part of Chile. Another possible scenario is that hybridization occurred between weak CAM populations of the blue clade and C₃ populations of *P. chilensis* before both colonised the northern part of Chile. It is unlikely that *P. chilensis* transferred CAM-specific genes to the blue clade because otherwise the yellow clade would presumably contain more constitutive CAM species. *P. casmichensis* is the other CAM species in the NC clade, this species occurring in sympatry with constitutive CAM species from the SC clade like *P. ferreyrae* and *P. macrura*. It is possible that *P. casmichensis* evolved CAM independently from *P. chilensis*, it is also possible that CAM could have had a hybrid origin. These hypotheses will need to be tested using chloroplast and nuclear markers.

No evidence was found for the evolution of constitutive CAM in the northern Andes clade of *Puya* species. The mesophytic ancestral niche of the northern Andes species and the high degree of niche conservatism could potentially prevent (during speciation) the colonisation of drier areas, which ultimately could prevent the clade evolving constitutive CAM. However, it is intriguing that one species (*Puya floccosa*) from the northern Andes has been reported to exhibit weak CAM.

Weak CAM in *Puya floccosa*

Puya floccosa is a species that could be identify as a non-constitutive CAM species given its carbon isotope values (−27.6 ‰, −28.5 ‰, −29.4 ‰, −30.5 ‰, Herrera *et al.*, 2010). *Puya floccosa* has populations that show nocturnal acidification (Herrera *et al.*, 2010). Given that information, it is likely that the contribution of nocturnal CO₂ uptake would be insignificant in *P. floccosa* (see for example *Werauhia sanguinolenta* in Pierce *et al.*, 2002a;

Beltrán *et al.*, 2013). The phylogenetic position of this species is unknown. Potentially *P. floccosa* might be assigned to the northern Andes clade *a priori* given the geographical range of this species and the limited geographical range found in the majority of *Puya* species. However, there are at least three reasons to be sceptical whether *P. floccosa* belongs *a priori* to the northern Andes clade: conservatism in elevational range, seed-dispersal limitation, and morphological dissimilarities.

While the majority of *Puya* species from the northern Andes are restricted to the paramo (above ~3000 m: Jabaily and Sytsma, 2013; Londoño *et al.*, 2014), *P. floccosa* is the species with widest range of elevation in the genus (Herrera *et al.*, 2010). This is unusual in *Puya*, in which the elevational range of most species is relatively restricted (Jabaily and Sytsma, 2013; Madriñan, 2015). This unusually wide range of elevations of *P. floccosa* is also found in *P. ferruginea*, the species with the second-widest range of elevations in the genus, which occurs in Bolivia, Ecuador and Peru. If *P. floccosa* belongs to the northern Andes clade, it means that the evolution of a wide elevation range occurred independently in *P. ferruginea* and in *P. floccosa*, given that the common ancestor of the northern Andes clade was a plant restricted to high-elevation environments (Jabaily and Sytsma, 2013).

Second, the majority of species in the northern Andes clade occupy very narrow and restricted areas, associated with high degrees of endemism (Madriñan, 2015), due to Pleistocene allopatric speciation and low seed-dispersal ability (Jabaily and Sytsma, 2013). However, *P. floccosa* occurs in the Andes, the Guiana shield and the Talamancan paramo (Costa Rica). This implies that long-distance seed dispersal in *P. floccosa* is effective. Additionally, *P. ferruginea* is an effective long-distance dispersal species. Moreover, the seed anatomy of *P. floccosa* is very similar to that of *P. ferruginea* (Varadarajan and Gilmartin, 1988). For example, the seeds of both species are elongate-triangular, falcate and cuneate.

The wings are permanent and the wing cells are mostly irregular (Varadarajan and Gilmartin, 1988). This morphology is different from the *P. hamata* group, which is found in the majority of northern Andean species. If *P. floccosa* belongs to the northern Andes clade, there has been convergence in seed morphology and dispersal in *P. ferruginea* and *P. floccosa*.

Morphologically, *P. floccosa* is more similar to Chilean and central Andes *Puya* species than to northern Andes species (Hornung-Leoni and Sosa, 2008). However, morphological characters have converged in different clades of *Puya* (i.e. sterile bracts; Jabaily and Sytsma, 2013).

These arguments do not necessarily mean that *P. floccosa* belongs to the central Andes clade or that it is closely related to *P. ferruginea*. *P. floccosa* may belong to the northern Andes clade and may have evolved numerous characteristics independently through convergence. On the other hand, if *P. floccosa* were closely related to *P. ferruginea* or belonged to the central Andes clade, the northern Andes would have been colonised twice by a high-elevation ancestor and an ancestor with more labile elevation range. Both scenarios are equally probable, and it will be necessary to include *P. floccosa* in a molecular-phylogenetic study to disentangle the origin of weak CAM in the northern Andes.

Conclusions

The monophyletic genus *Puya* is an excellent model to understand the ecological significance of CAM in tropical montane environments. The CAM and C₃ climatic niches have been formally defined and the aridity index and cloud cover seasonality found to be the most important variables, according to the Random Forest, in explaining the distribution of species with these two photosynthetic pathways. However, it is the combination of aridity index, cloud cover seasonality, mean annual cloud frequency and

precipitation seasonality that can classify the photosynthetic pathway despite the overlap. High-elevation CAM species experience lower temperatures and higher diel temperature range than low-elevation CAM species, suggesting that high-elevation CAM *Puya* species have adapted to operate full CAM at low temperatures. This topic is described and discussed further in Chapter 3.

The lack of CAM species in the northern Andes is explained by a selective biogeographic barrier preventing the colonisation of northern Andes by CAM species from the south. It is important to test whether there are northern Andes species of *Puya* showing weak CAM.

Constitutive CAM has evolved multiple times in *Puya*. The two largest clades of *Puya* (the southern and central Andes clade and the northern and central Andes clade) showed differences in the number of CAM species. This is probably a consequence of high degree of niche conservatism and the ancestral condition of weak CAM in the common ancestor of southern and central Andes clade. The photosynthetic pathway of the common ancestor of *Puya* is still undetermined. The present study demonstrates that changes in the climatic niche were correlated with changes in the photosynthetic pathway.

Finally, studying the evolution and expression of CAM genes in C₃ species from different clades of *Puya* and particularly in *P. alpestris* ssp. *zoellneri* and *P. alpestris* ssp. *alpestris*, will help to clarify the possible ancestral photosynthetic pathway of the common ancestor of *Puya* (Chapter 4).

Further discussion of limitations and implications of the ecological differences between CAM and C₃ are considered in Chapter 5.

Chapter 3. Physiological and biochemical adaptations of CAM species to high-elevation environments in the genus *Puya*

Abstract

Nocturnal CO₂ fixation has been documented in tropical high-elevation CAM species, but no studies have evaluated the physiological, biochemical, or molecular adaptations of tropical high-elevation CAM species to low temperatures. In particular, changes in phosphoenolpyruvate carboxylase (PEPC), the enzyme responsible for nocturnal CO₂ fixation. Adaptations to low temperatures were therefore evaluated in a common garden environment, using high- and low-elevation *Puya* species and contrasting them with equatorial bromelioids typically occurring at low elevations. The present study evaluated correlations between PEPC temperature optimum, thermal denaturation of PEPC at 50 °C, and relative activity reduction of PEPC at 5 °C, with elevation and latitude at which the various species are found. When cultivated under common-garden conditions, *Puya* species showed higher cold resistance than equatorial bromelioids. High-elevation *Puya* species showed higher cold resistance than low-elevation *Puya* species. PEPC assayed in crude extracts from high-elevation *Puya* species showed more rapid thermal denaturation at 50 °C and higher activity at 5 °C (relative to 25 °C) compared with low-elevation *Puya* and equatorial bromelioids. However, the temperature optimum for PEPC activity did not differ between high- and low-elevation *Puya* species. The present study demonstrates that the physiological and biochemical adaptations detected in these species, grown in a common-garden experiment, correlate well with the natural habitats

in which they occur. The molecular adaptation of PEPC in high-elevation CAM species of *Puya* is possibly explained by structural flexibility.

Introduction

CAM plants do not occur exclusively in warm environments (Chapter 2). The majority of CAM plants are not resistant to sub-zero temperatures, but a few high-latitude species occur in environments with prolonged periods of sub-zero temperatures (Nobel, 1996). Investigations of these species have included measurements of cold resistance (Nobel, 1980b; Nobel, 1996), titratable acidity (Wagner and Larcher, 1980; Larcher, 2010) and CO₂ exchange (Loik and Nobel, 1993a, b). However, tropical CAM species from high-elevation environments at equatorial latitudes have not been extensively studied for their cold resistance and physiology. The present study aims to evaluate the cold resistance and thermal lability of phosphoenolpyruvate carboxylase (PEPC; EC 4.1.1.31) in high-elevation CAM species of the genus *Puya*. PEPC is the key enzyme responsible for fixing nocturnal CO₂ into organic acids in CAM plants (Osmond, 1978; Winter *et al.*, 2015).

Cold resistance has been extensively investigated in CAM species that occur at high latitudes. For example, the cactus *Escobaria vivipara* found in the Mercury Valley in Nevada, USA, survives exposure to sub-zero temperatures during the winter (Nobel, 1981). Moreover, the cold resistance of *E. vivipara* increased after gradual cold acclimation (Nobel, 1981). The cacti *Opuntia humifusa* and *O. fragilis* are the world's most freezing-tolerant and most northerly cacti, found in Ontario and Alberta (CANADA), respectively (Loik and Nobel, 1991, 1993a, b). Furthermore, *O. erinacea*, from Washington State (USA) showed significant day–night titratable acidity changes in December when temperatures were close to 0 °C. However, day–night changes in the titratable acidity in the summer was 14 times higher than in winter (Littlejohn and Williams III, 1983). Despite the

survival of some CAM plants at low temperatures, CAM expression is reduced at low temperatures in high-latitude CAM species.

Other CAM species such as *Agave utahensis* and *A. parryi* experience temperatures close to 0 °C during the winter and show greater resistance to freezing compared with agaves from more equatorial regions (Nobel and Smith, 1983). The alpine and facultative CAM species *Sempervivum montanum* experiences nocturnal low temperatures even during the summer months (Larcher, 2010); in winter, this species exhibits C₃ photosynthesis (Wagner and Larcher, 1981). All these examples are plants from higher latitudes where winter temperatures impose limitations on the plant in terms of their growth and survival. It is likely such species do not fix significant amounts of CO₂ during winter. Moreover, these plants have time to acclimate to the winter by a process of cold hardening, given the progressive reduction of temperature during the autumn (Nobel, 1981; Nobel and Loik, 1993b). This is not the case for equatorial high-elevation CAM plants.

Tropical and subtropical high-elevation ecosystems (above the permanent tree-line) are characterized by experiencing greater diurnal temperature ranges than seasonal temperature ranges (Rundel *et al.*, 1994). Hedberg (1964) described the afroalpine environments as ecosystems that experience ‘a summer every day and a winter every night’ because of a high diurnal temperature range, which can approach 40 °C in tropical mountains at high elevation. Cold resistance has been demonstrated in afroalpine species genera *Senecio* and *Lobelia* (Beck *et al.*, 1982). This pattern is also seen in the equatorial and subequatorial Andes (Medina and Delgado, 1976; Keeley and Keeley, 1989). In these ecosystems, night frost is a relatively common occurrence. For example, in the central Andes (e.g. *Volcán el Misti* in Peru above 4000 m), 80–90% of the nights each year have temperatures below freezing (Rundel *et al.*, 1994). Such climatic characteristics are reflected in the diversification of a distinctive endemic flora (Sklenář *et al.*, 2011), which

contains species that are good models to test adaptations to low temperatures and night-time frost in the tropics.

The physiology of tropical CAM high-elevation species has been studied in terms of nocturnal CO₂ exchange, titratable acidity, carbon isotope values, and water potential. It is remarkable how species such as *Oroya peruviana* and *Tephrocactus floccosus* (syn. *Austrocylindropuntia floccosa*) show nocturnal titratable-acidity increases even when the night temperature drops below 0 °C (Keeley and Keeley, 1989). In contrast with CAM species from high latitudes, the equatorial high-elevation CAM species fix significant amounts of CO₂ at sub-zero temperatures. For example, *Puya meiziana* (4000 m) has a carbon isotope value of −12.0 ‰, indicating that the majority of CO₂ was fixed at night by PEPC (Crayn *et al.*, 2015; Chapter 2). Therefore, it is likely that there are molecular adaptations of the enzymes involved in the nocturnal CO₂ fixation at low temperatures in equatorial high-elevation CAM species. Despite the fact that studies have described nocturnal CO₂ fixation at temperatures close to 0 °C in equatorial high-elevation CAM species, there have been no studies of the thermal lability of the key enzyme PEPC, which is responsible for fixing CO₂ at night in such species.

Temperature affects the kinetic energy of proteins, reaction rates, the flexibility or strength of molecular interactions, and other chemical and physical properties of enzymes (Elias *et al.*, 2014). Psychrophile organisms (bacteria, archaea and eukaryote) can flourish at temperatures close to the freezing point of water (Feller and Gerday, 2003). Psychrophilic enzymes, from psychrophile organisms, tend to be more flexible near sub-zero temperatures compared with mesophilic enzymes (Somero, 2004; Chiuri *et al.*, 2009). This flexibility compensates for the extra energy cost of activation at low temperatures (Elias *et al.*, 2014). Consequently, the temperature optimum of psychrophilic enzymes is generally shifted towards lower temperatures than their mesophilic homologs (Feller,

2003, 2010, 2013). Therefore, cold-adapted enzymes tend to be inactive at higher temperatures (Georlette *et al.*, 2004; Elias *et al.*, 2014) and perform relatively better at lower temperatures (Fields and Somero, 1998) in comparison with mesophilic enzymes. High-elevation equatorial plants are not psychrophilic organisms because they do not permanently occur at temperatures close to sub-zero; they only experience sub-zero temperatures at night, for few hours but, as explained above, almost every night. Since CO₂ fixation occurs at sub-zero temperatures in high-elevation CAM species, it is possible that the enzymes involved exhibit characteristics of psychrophilic enzymes compared with low-elevation CAM species.

The genus *Puya* contains CAM species occurring at low and high elevations (Chapter 2). Therefore, *Puya* is an excellent model to test whether high-elevation CAM species have molecular adaptations at lower temperatures, which are usually experienced at night-time. The present study evaluated the differences between the cold resistance of high- and low-elevation species, the temperature optimum for PEPC activity, and the thermal lability of PEPC at high (50 °C) and low (5 °C) temperatures. These characteristics were measured in a common-garden experiment to establish whether latitude and elevation could explain the variation between all plants included in the present study. Given that the sister clade of *Puya* is the subfamily Bromelioideae (Schulte *et al.*, 2009; Givnish *et al.*, 2014), and that the majority of low-elevation species of *Puya* occur in Chile (Chapter 2), lowland equatorial species from the subfamily Bromelioideae were included in the present study to investigate the physiological differences between high- and low elevation CAM species. The non-equatorial bromelioid species *Fascicularia bicolor* and *Ochagavia carnea* were included in the comparison of the cold resistance of various species, but they were not used in other experiments because neither of them is a constitutive CAM plant.

Materials and Methods

Plant material and growing conditions

The species included in the present study were: *Aechmea nudicaulis* (1 plant); *Ananas comosus* (3 plants); *Bromelia humilis* (2 plants); *Fascicularia bicolor* (3 plants); *Ochagavia carnea* (2 plants); *Puya alpestris* ssp. *zoellneri* (3 plants); *P. chilensis* (1 plant); *P. harmsii* (3 plants); and *P. castellanosi* (2 plants). The plants were obtained from the University of Oxford Botanic Garden (*P. alpestris* ssp. *zoellneri* (accession numbers: 2010105.17 and 201010.17) and *P. chilensis* (accession number: 2010105.18)), the private collection of Professor J. Andrew C. Smith at the Department of Plant Sciences at University of Oxford (*A. nudicaulis*, *A. comosus*, *F. bicolor* and *O. carnea*) and *P. castellanosi* and *P. harmsii* were purchased from Tropical Britain Ltd, Addlestone, Surrey, UK. *F. bicolor* and *O. carnea* were collected by Professor Georg Zizka in southern and central Chile, respectively. The plants were initially at the *Botanischer Garten Frankfurt am Main* under the accession numbers: CL-0-FRP-16850 and CL-0-FRP-16847 (*F. bicolor*) and CL-0-FRP-18378 and CL-0-FRP-16853 (*O. carnea*). The seeds of *P. castellanosi* were collected in north-west Argentina, Salta Province, locally in the uplands around Laguna del Brealito in Valle Calchaquí (3000 m). The seeds for *P. harmsii* were collected in north-west Argentina, Salta, Catamarca and Tucumán Provinces at elevations of 2000–3000 m. The seeds of both species were germinated in the UK. The species *P. chilensis* and *P. alpestris* ssp. *zoellneri* were collected in the northern part of Central Chile. All the experiments were performed with adult plants, if the rosette diameter were larger than 20 cm and the height over 20 cm.

The adult plants were maintained in common-garden conditions for more than two months before starting the experiments. The experiments were run from February 2016 to April 2017. The plants were maintained in the glasshouse on the rooftop of the Department of Plant Sciences at the University of Oxford (51.75 °N, 1.26 °W). The

photoperiod was controlled to 18 h of light and 6 h of dark with external lamps in winter months. The plants were watered every three days. The plants never showed any signs of water stress and never experienced temperatures below 0 °C during this period. During the day, maximum temperature was 28 °C in summer and in winter the temperature is controlled and fixed to 21° C and during the night never dropped below 10 °C (either winter or summer), for that reason acclimation conditions were labelled as non-cold acclimated.

Cold resistance and freezing damage

Electrolyte leakage is a standard procedure to measure freezing damage in plant cells (Prášil and Zámečník, 1998). In principle, ice nucleation in the protoplasm irreversibly damages the cell membranes. Consequently, the cell would no longer be able to retain the ions and other solutes in the protoplast, and the electrolytes would leak from the cell into the extracellular space and external medium. The degree of freezing damage is established by measuring the conductivity of different sections of the same type of tissue at different temperatures (temperature treatments) over a fixed period. This initial conductivity measurement is C_i . In order to avoid bias, C_i is normalised by measuring the conductivity after boiling the tissue (C_f ; conductivity of fully dead tissue). For each temperature treatment, relative electrolyte leakage (C_i/C_f) is the proportion of dead cells after each temperature treatment. From these measurements, the temperature that causes 50% of the damage in the tissue (T_{50}) is calculated. In the present study, T_{50} was used to compare cold resistance among different species and different groups of species according to the elevation and latitude where the species were originally from. The groups were defined as lowland equatorial bromelioids (*Aechmea nudicaulis*, *Ananas comosus*, and *Bromelia humilis*); non-equatorial, higher latitude bromelioids (*Fascicularia bicolor* and *Ochagavia*

carnea); low-elevation *Puya* species (*P. alpestris* ssp. *zoellneri* and *P. chilensis*); and high-elevation *Puya* species (*P. castellanosi* and *P. harmsii*).

Plant tissue, isolated from the middle part of a mature, healthy leaf, was collected between 10.00 h and 13.00 h when acid concentration is reasonably low in CAM species. This parameter was standardised because the acid content is correlated with the cold resistance in bromeliads (Neves-Chaves *et al.*, 2015). The shape of the leaf pieces was evaluated to establish whether there were differences in the electrolyte leakage using leaf discs (5 mm diam.) or leaf rectangles (5.0 mm × 3.5 mm) in *Ananas comosus*. After excision, tissue segments were rinsed in distilled water to remove electrolytes from wounded cells and excess water removed using absorbent paper. Tissue pieces were then kept at six different temperatures: 0 °C; -5 °C; -10 °C; -13 °C; -20 °C; -80 °C. To establish conductivity of a tissue with zero freezing damage, tissue pieces were maintained at 4 °C. After 3 h in each temperature treatment, five pieces of tissue of each individual leaf were immersed in 5 ml of distilled water in a 10 ml vial at room temperature. The conductivity of each vial was measured using a Conductivity/TDS stick meter, Primo5® (Hanna Instruments, Leighton Buzzard, Bedfordshire, UK) until the conductivity reached a constant value. This equilibration time was estimated using rectangular- and circular-shaped pieces of leaf tissue (*Ananas comosus*). After conductivity reached the equilibrium value, the initial conductivity (C_i) was recorded. Sealed vials were then boiled for 1 hour and allowed to cool to room temperature, when conductivity was measured again (C_f). This procedure was carried out for three technical replicates. Three different leaves for each individual plant were assessed each time. Each individual plant was assessed three times independently from February 2016 to April 2017.

The T_{50} was calculated for each individual plant using a quasi-binomial logistic regression after the C_i/C_f ratio was transformed (C_i/C_f at 4 °C = 0; C_i/C_f at -80 °C = 1).

The differences in T_{50} between all species, between high and low-elevation species, and between equatorial and non-equatorial bromelioid species, were tested using ANOVA. If the F -value was significant, the Tukey Honestly Significant Difference (HSD; Tukey, 1949) test was performed to establish which species or groups were significantly different in a pairwise comparison.

PEPC activity

Plant tissue for PEPC activity measurements was excised from the middle portion of mature, healthy leaves. Plant tissues were collected at dusk to avoid high concentration of malate that inhibit the activity of PEPC (Buchanan-Bollig *et al.*, 1984). PEPC activity was measured following the protocol of Freschi *et al.* (2010) with the following modifications. The extraction buffer contained 200 mM Tris (pH 8.0, HCl); 1 mM Na_2EDTA ; 5 mM dithiothreitol (DTT); 10 mM MgCl_2 ; 10% (v/v) ethylene glycol; 0.5% (w/v) bovine serum albumin (BSA). Leaf tissue samples were ground in liquid nitrogen using a pestle and mortar. The fine powder was then homogenised in 9 volumes of extraction buffer. The sample was filtered through a layer of Whatman® filter paper No.1 and glass wool and centrifuged at 1000 rpm for 1 minute at 4 °C. The volume of the supernatant was recorded, the sample was maintained on ice, and aliquots were immediately used for enzymatic assays. PEPC activity was assayed in 0.975 ml of reaction buffer containing 50 mM Tris (pH 8.0, HCl); 10 mM MgCl_2 ; 10 mM NaHCO_3 ; 0.2 mM NADH; 3.0 mM phosphoenolpyruvate (PEP); 1.0 mM DTT. The reaction was started by adding 25 μl of the extract to complete a final volume of 1 ml. The absorbance was monitored at 340 nm and recorded every 10 s for 5 to 10 minutes until a linear rate of absorbance decrease was observed. The activity was calculated every 50 s and the maximum value over the reaction time was registered as the enzyme activity at saturating (3.0 mM) substrate concentration.

Temperature optimum of PEPC

The activity of PEPC at saturating substrate concentration (3.0 mM PEP) was measured as described above, changing the temperature of the reaction buffer (from 10 °C to 60 °C). The temperature of the reaction was monitored and recorded with a Fisher Scientific™ Traceable™ Pocket-Size Thermometer (Fisher Scientific Loughborough, Leicestershire, UK) using a K-type thermocouple. For each reaction, a single temperature point was established when the maximum velocity was reached. The temperature points and PEPC activity were plotted against each to fit a smooth polynomial function for each species. The inflection point of the polynomial function was calculated for each curve, and the value was a proxy of the temperature optimum. The differences in the shape of the fitting curve were evaluated by repeated measures ANOVA test. The differences between the temperature optimum between high- and low- elevation species was evaluated by one way ANOVA test.

Thermal stability of PEPC at 50 °C

PEPC activity at saturating substrate concentration (3.0 mM PEP) was measured as described above. The same extract used at 25 °C was then incubated at 50 °C. An aliquot of the extract was taken to measure the PEPC activity at 25 °C after 5, 10, 15, 30, 60 and 180 minutes of incubation at 50 °C. The relative loss of PEPC activity with duration of incubation at 50 °C was defined as the ratio of PEPC activity before and after incubation. PEPC activity was measured for each individual plant in three leaves collected between November 2016 and April 2017. The differences in high-temperature inactivation of PEPC between species and between lowland equatorial bromelioids, low-elevation *Puya* species and high elevation *Puya* species were evaluated by repeated measures of the ANOVA test. The linear mixed-effect model was corrected by the autocorrelation of the relative loss of PEPC activity and time. The differences between

groups and species were also evaluated using an ANOVA test and a *post hoc* analysis at 5 minutes after incubation to test whether at that time there were significant differences between species and groups in their relative loss of PEPC activity.

Reduction in PEPC activity at 5 °C

PEPC activity at saturating substrate concentration (3.0 mM PEP) was measured at 25 °C as described above and the same extract was used to measure PEPC activity at 5 °C. The relative loss of PEPC activity at 5 °C was defined as the ratio between activity at 5 °C and activity at 25 °C (5:25 ratio). PEPC activity was measured for each individual plant in three leaves collected between November 2016 and April 2017. The differences in the 5:25 ratio between species and between equatorial bromelioids, low-elevation *Puya* species and high-elevation *Puya* species were evaluated by an ANOVA test, and the significance of pairwise comparisons was assessed by a Tukey HSD test.

Statistical analyses

All the statistical analyses were performed using R (R Core Team, 2017). The ANOVA models were performed using the functions `lme`, `annova`, `lsmeans`, in the packages *nlme*, *car*, *multcompView* and *lsmeans* (Fox and Weisberg, 2011; Graves *et al.*, 2015; Lenth *et al.*, 2016; Pinheiro *et al.*, 2017). Figures were plotted using the package *ggplot2* (Wickham, 2009).

Results

Cold resistance

Cold resistance was measured in order to test whether, in material maintained under the same environmental conditions in a common-garden experiment, the T_{50} was different in species native to different elevations. The shape and dimensions of the tissue used in these tests was evaluated to optimize the experimental conditions. Rectangular

and circular pieces did not show significant differences in their T_{50} values (Supplementary Fig. 1A). Rectangular pieces (1 h) reached equilibrium more quickly than circular pieces (1.5 h; Supplementary Fig. 1B). For that reason, all the measurements of T_{50} were taken using rectangular-shaped tissue.

The quasi-binomial logistic regressions used to determine T_{50} for each species are shown in Supplementary Fig. 2. The lowland equatorial bromeliads showed a lower cold resistance (less negative values of T_{50}) than any other bromeliads in this study (Fig. 1, Table 1). The high-elevation species showed higher cold resistance (more negative values of T_{50}). The ANOVA test showed that there were significant differences in the cold resistance between species ($F = 240.7$; $df = 8$; $p < 0.0001$). The Tukey HSD test showed that *Ananas comosus*, *Aechmea nudicaulis* and *Bromelia humilis* did not exhibit significant differences in their cold resistance (Table 2). There were no significant differences in species' cold resistance in the following pairwise comparisons: *Puya castellanosi* vs. *P. harmsii*; *P. alpestris* vs. *P. chilensis*; *P. alpestris* vs. *Ochagavia carnea*; and *P. chilensis* vs. *O. carnea*. *Ochagavia carnea* and *Fascicularia bicolor* did not show significant differences in their cold resistance, but the p -value for this comparison was very close to 0.05 ($p = 0.0523$; Table 2).

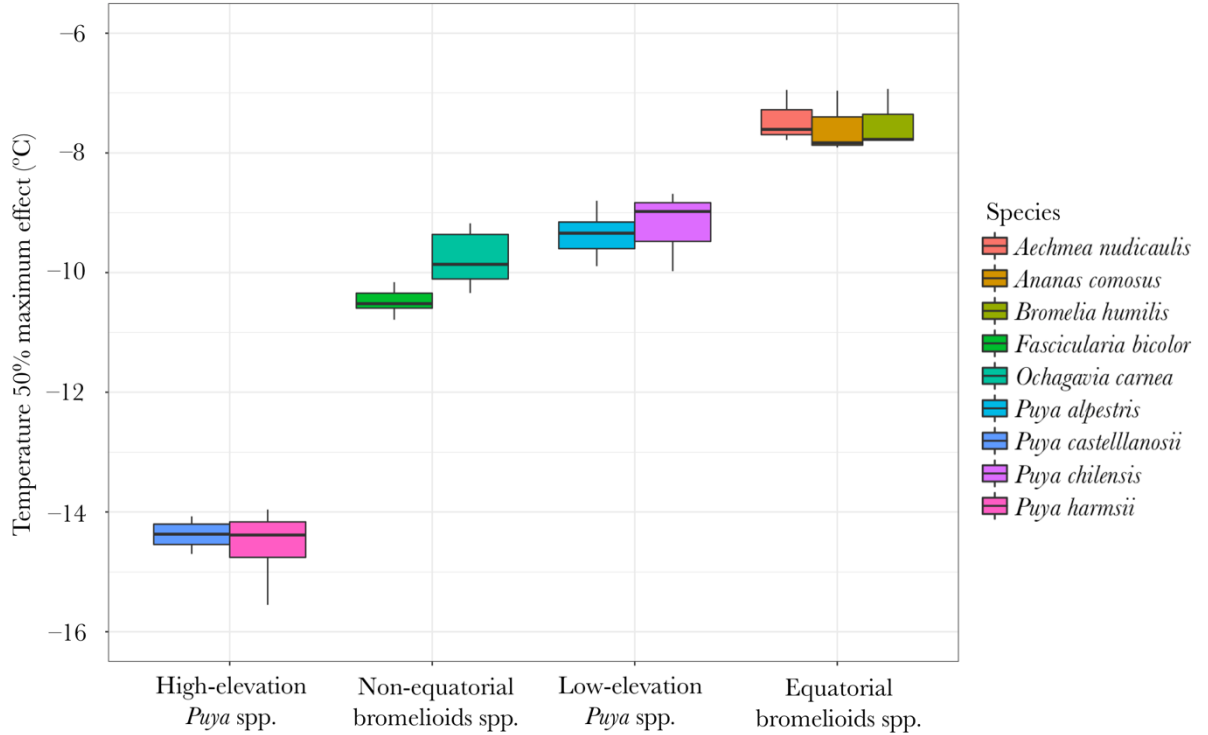


Figure 1. Temperature 50% maximum effect (T_{50}) for four species of *Puya* (including both high- and low-elevation species) and five species from the subfamily Bromelioideae (both equatorial and non-equatorial) in a common garden experiment. T_{50} was determined as the temperature that produces 50% of maximum electrolyte leakage using figures in Supplementary Fig. 2, after fitting using a quasi-binomial logistic regression. The ANOVA test suggested that there were significant differences between species ($F = 240.7$; $df = 8$; $p < 0.0001$) and between groups ($F = 563.4$; $df = 3$ $p < 0.0001$). The horizontal thick line in the boxplots indicates the median value. The whiskers represent data from the 25th or 75th percentile until the furthest data point that is within 1.5 times the inter-quartile range. For each individual plant (see the Materials and Methods section), the T_{50} was determined in three independent experiments at different times of year (Supplementary Fig. 2). n = independent experiments; $n = 3$ for *Aechmea nudicaulis* and *Puya chilensis*; $n = 6$ for *Bromelia humilis*, *Ochagavia carnea*, and *P. castellanosii*; $n = 9$ for *Ananas comosus*, *Fascicularia bicolor*, *P. alpestris* ssp. *zoellneri* and *P. harmsii*.

Table 1. Summary of the 50% maximum effect temperature (T_{50}) for each species. Plants were maintained for a minimum of 2 months in a common garden experiment. The T_{50} values were calculated using a quasi-binomial logistic regression of the relative electrolyte leakage at $-80\text{ }^{\circ}\text{C}$, $-20\text{ }^{\circ}\text{C}$, $-13\text{ }^{\circ}\text{C}$, $-10\text{ }^{\circ}\text{C}$, $-5\text{ }^{\circ}\text{C}$, $0\text{ }^{\circ}\text{C}$ and $4\text{ }^{\circ}\text{C}$ described in Supplementary Fig. 2. Values are means $\pm$ standard deviation. n = independent experiments; $n = 3$ for *Aechmea nudicaulis* and *Puya chilensis*; $n = 6$ for *Bromelia humilis*, *Ochagavia carnea*, and *P. castellanosii*; $n = 9$ for *Ananas comosus*, *Fascicularia bicolor*, *P. alpestris* ssp. *zoellneri* and *P. harmsii*.

Species	T_{50} ($^{\circ}\text{C}$)
<i>Aechmea nudicaulis</i>	-7.45 ± 0.44
<i>Bromelia humilis</i>	-7.50 ± 0.50
<i>Ananas comosus</i>	-7.57 ± 0.53
<i>Puya chilensis</i>	-9.21 ± 0.68
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	-9.37 ± 0.37
<i>Ochagavia carnea</i>	-9.77 ± 0.48
<i>Fascicularia bicolor</i>	-10.48 ± 0.22
<i>Puya castellanosii</i>	-14.38 ± 0.24
<i>Puya harmsii</i>	-14.48 ± 0.49

An ANOVA test was conducted to test differences between the groups of lowland equatorial bromelioids, non-equatorial bromelioids, low-elevation *Puya* species, and high-elevation *Puya* species. The ANOVA showed that there were significant differences between groups ($F = 563.4$; $df = 16$; $p < 0.0001$). According to the Tukey HSD test, all groups exhibited significant differences in their T_{50} values (Table 3). The present results demonstrate that high-elevation *Puya* species exhibit the highest degree of cold resistance of any group evaluated in this study (Fig. 1, Table 3).

Table 2. Summary of the Tukey HSD pairwise differences in T_{50} values between species. The p -values were corrected for multiple comparisons, and in bold indicate significant differences in T_{50} ($p < 0.05$) between the paired species.

Species	Mean difference (°C)	p-value
<i>P. castellanosii</i> vs. <i>A. comosus</i>	-6.8076	< 0.0001
<i>P. harmsii</i> vs. <i>A. comosus</i>	-6.9114	< 0.0001
<i>P. castellanosii</i> vs. <i>A. nudicaulis</i>	-6.929	< 0.0001
<i>P. harmsii</i> vs. <i>A. nudicaulis</i>	-7.0328	< 0.0001
<i>P. castellanosii</i> vs. <i>B. humilis</i>	-6.873	< 0.0001
<i>P. harmsii</i> vs. <i>B. humilis</i>	-6.9768	< 0.0001
<i>P. castellanosii</i> vs. <i>F. bicolor</i>	-3.8969	< 0.0001
<i>P. harmsii</i> vs. <i>F. bicolor</i>	-4.0007	< 0.0001
<i>P. castellanosii</i> vs. <i>O. carnea</i>	-4.6053	< 0.0001
<i>P. harmsii</i> vs. <i>O. carnea</i>	-4.7091	< 0.0001
<i>P. castellanosii</i> vs. <i>P. alpestris</i>	-5.0073	< 0.0001
<i>P. harmsii</i> vs. <i>P. alpestris</i>	-5.1111	< 0.0001
<i>P. chilensis</i> vs. <i>P. castellanosii</i>	5.1634	< 0.0001
<i>P. harmsii</i> vs. <i>P. chilensis</i>	-5.2672	< 0.0001
<i>F. bicolor</i> vs. <i>A. nudicaulis</i>	-3.0321	< 0.0001
<i>F. bicolor</i> vs. <i>B. humilis</i>	-2.9762	< 0.0001
<i>F. bicolor</i> vs. <i>A. comosus</i>	-2.9108	< 0.0001
<i>O. carnea</i> vs. <i>A. nudicaulis</i>	-2.3236	< 0.0001
<i>O. carnea</i> vs. <i>B. humilis</i>	-2.2677	< 0.0001
<i>O. carnea</i> vs. <i>A. comosus</i>	-2.2023	< 0.0001
<i>P. alpestris</i> vs. <i>A. nudicaulis</i>	-1.9217	< 0.0001
<i>P. alpestris</i> vs. <i>B. humilis</i>	-1.8658	< 0.0001
<i>P. alpestris</i> vs. <i>A. comosus</i>	-1.8004	< 0.0001
<i>P. alpestris</i> vs. <i>F. bicolor</i>	1.1104	< 0.0001
<i>P. chilensis</i> vs. <i>A. nudicaulis</i>	-1.7655	0.0002
<i>P. chilensis</i> vs. <i>B. humilis</i>	-1.7096	0.0003
<i>P. chilensis</i> vs. <i>A. comosus</i>	-1.6442	0.0005
<i>P. chilensis</i> vs. <i>F. bicolor</i>	1.2666	0.0012
<i>O. carnea</i> vs. <i>F. bicolor</i>	0.7085	0.0523
<i>P. chilensis</i> vs. <i>O. carnea</i>	0.5581	0.6122
<i>P. alpestris</i> vs. <i>O. carnea</i>	0.4019	0.654
<i>P. chilensis</i> vs. <i>P. alpestris</i>	0.1562	0.9997
<i>P. harmsii</i> vs. <i>P. castellanosii</i>	-0.1038	0.9999
<i>A. nudicaulis</i> vs. <i>A. comosus</i>	0.1213	1
<i>B. humilis</i> vs. <i>A. comosus</i>	0.0654	1
<i>B. humilis</i> vs. <i>A. nudicaulis</i>	-0.0559	1

Table 3. Summary of the Tukey HSD pairwise differences in T₅₀ values between equatorial bromelioid species, low-elevation *Puya* species and high-elevation *Puya* species. The *p*-values were corrected for multiple comparisons, and in bold indicate significant differences in T₅₀ (*p* < 0.05) between the paired species groups.

Groups	Mean	
	difference	<i>p</i> -value
	(°C)	
Low-elevation <i>Puya</i> spp. <i>vs.</i> High-elevation <i>Puya</i> spp.	5.1086	< 0.0001
Non-equatorial bromelioid spp. <i>vs.</i> High elevation <i>Puya</i> spp.	4.2425	< 0.0001
Equatorial bromelioid spp. <i>vs.</i> High elevation <i>Puya</i> spp.	6.9322	< 0.0001
Non-equatorial bromelioid spp. <i>vs.</i> Low elevation <i>Puya</i> spp.	−0.8661	< 0.0001
Equatorial bromelioid spp. <i>vs.</i> Low-elevation <i>Puya</i> spp.	1.8236	< 0.0001
Equatorial bromelioid spp. <i>vs.</i> Non-equatorial bromelioid spp.	2.6896	< 0.0001

Temperature optimum of PEPC

The temperature optimum of PEPC activity was measured at saturating substrate concentration (3.0 mM PEP) to establish whether there were significant differences in enzyme kinetics between the high- and low-elevation species of *Puya* (Fig. 2; Supplementary Table 1). *Puya alpestris* had the lowest temperature optimum (39.8 °C) for PEPC activity and *P. castellanosii* the highest temperature optimum (44.0 °C). However, there were no significant differences in the shape of the polynomial fitting curve between the four species ($F = 0.3080$; $df = 3$; $p = 0.8190$) and the temperature optimum did not differ between high- and low-elevation species ($F = 0.1460$; $df = 1$; $p = 0.7040$).

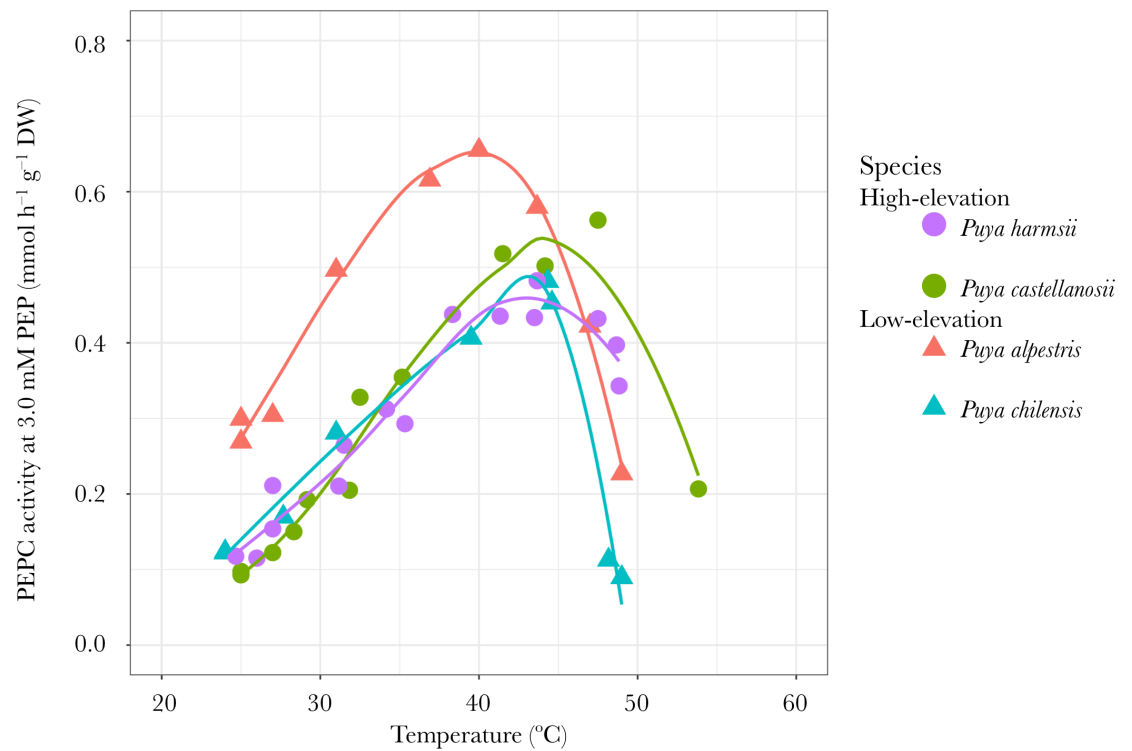


Figure 2. Temperature curves of PEPC activity recorded at 3.0 mM PEP activity of four species of *Puya*. High- and low-elevation species are represented by triangles and circles, respectively. Lines represent the smooth function obtained by fitting a polynomial surface for *P. alpestris* (red), *P. chilensis* (blue), *P. castellanosii* (green), and *P. harmsii* (purple). The temperature curves of the four species did not differ significantly ($F = 0.3080$; $df = 3$; $p = 0.8190$) and there were no significant differences between the T_{opt} of high- and low-elevation species ($F = 0.1460$; $df = 1$; $p = 0.7040$). The estimated T_{opt} are summarised in Supplementary Table.1. ($n = 1$ measurement per species)

Thermal stability of PEPC at 50 °C

The thermal stability of PEPC activity at 50 °C was measured to test whether high-elevation species show a higher lability of the enzyme at high temperatures compared with congeneric low-elevation species and lowland equatorial bromelioids. Thermal stability was calculated as the percentage of residual PEPC activity remaining after incubating the extract at 50 °C for a standardized time. In all species, the activity of PEPC assayed at saturating substrate concentration decreased after incubation at 50 °C

(Fig. 3A). The most dramatic loss of PEPC activity occurred in the high-elevation species *P. harmsii* and *P. castellanosi*, which after five minutes of incubation at 50 °C showed an average residual PEPC activity of 41% and 39%, respectively. In contrast, the lowland equatorial species showed the least reduction in PEPC activity, *Aechmea nudicaulis* in particular showing residual PEPC activity of 91% after incubation for 5 min at 50 °C (Fig. 3B).

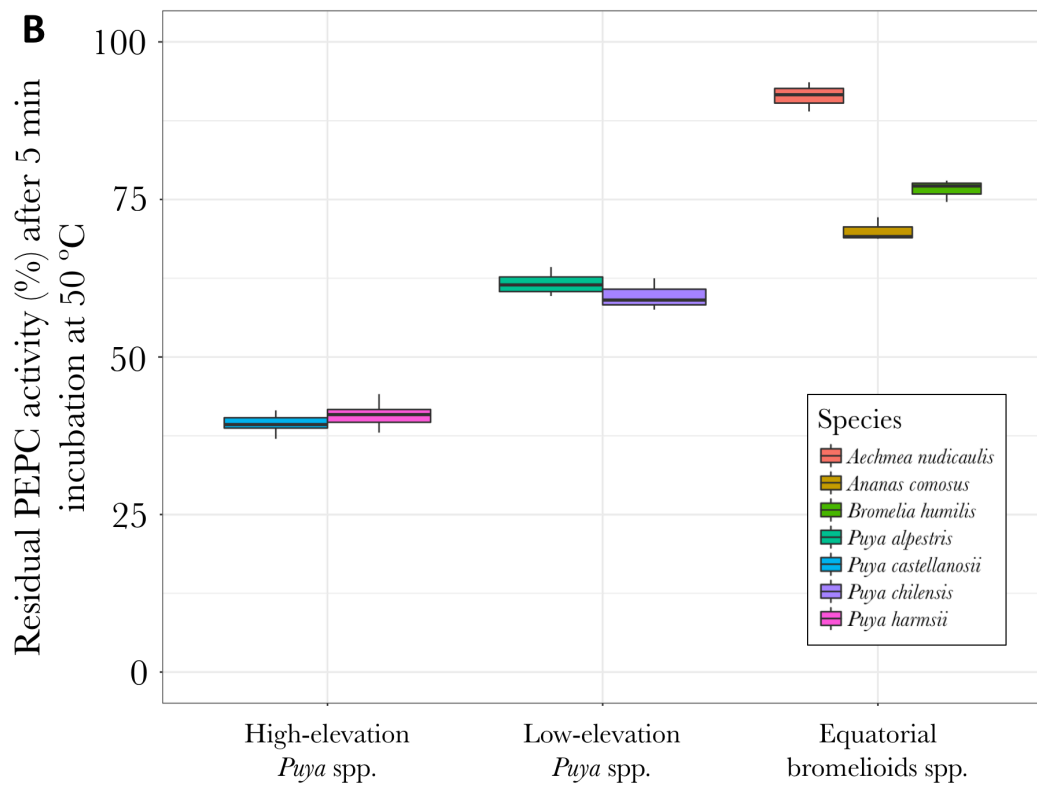
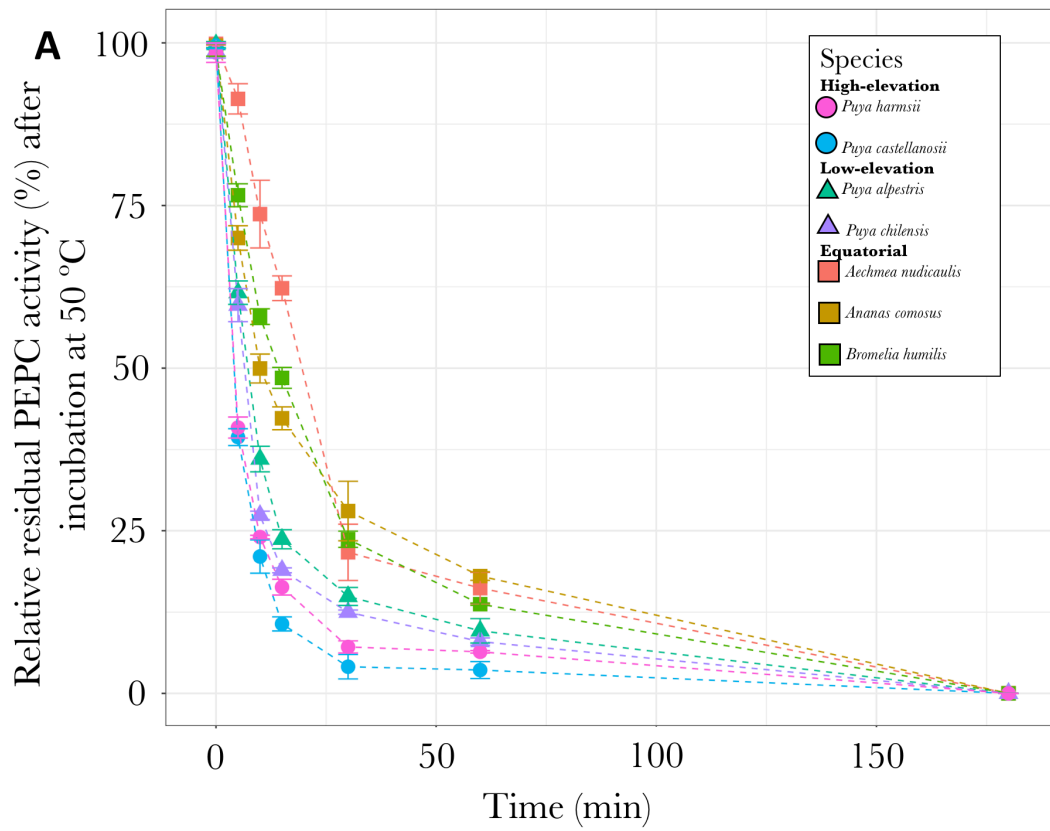


Figure 3A. Summary of the percentage of the residual PEPC activity remaining after incubation for various times at 50 °C of extracts from four species of *Puya* and three species of the subfamily Bromelioideae. The percentage of the residual PEPC activity was calculated as the ratio of the activity before and after incubation for 5, 10, 15, 30, 60 and 180 minutes at 50 °C. 100% means the activity at 25 °C without incubation. Circles represent high-elevation *Puya* species, triangles represent low-elevation *Puya* species, and squares represent equatorial bromelioid species. The ANOVA for repeated measures showed significant differences between species over time ($p = 0.0048$) and significant differences between equatorial bromelioid species, low-elevation *Puya* and high-elevation *Puya* species ($p = 0.0003$). Data are means $\pm$ standard deviation.

Figure 3B. Summary of the percentage of the residual PEPC activity remaining after incubation for 5 minutes at 50 °C of four species of *Puya* and three species of subfamily Bromelioideae. Percentage of residual PEPC activity after incubation at 50 °C was calculated as described in Fig. 3. The ANOVA test showed that there were significant differences in the residual activity between species ($F = 429.9$; $df = 6$; $p < 0.0001$) and between equatorial bromelioid species, low-elevation *Puya* species and high-elevation *Puya* species ($F = 563.4$; $df = 3$; $p < 0.0001$). The horizontal thick line in the boxplots indicates the median value. The whiskers represent data from the 25th or 75th percentile until the furthest data point that is within 1.5 times the inter-quartile range. n = independent experiments; $n = 3$ for *Aechmea nudicaulis* and *Puya chilensis*; $n = 6$ for *Bromelia humilis*, *Ochagavia carnea*, and *P. castellanosi*; $n = 9$ for *Ananas comosus*, *Fascicularia bicolor*, *P. alpestris* ssp. *zoellneri* and *P. harmsii*.

There were significant differences in the residual PEPC activity after treatment at 50 °C between species (Supplementary Table 2) and between lowland equatorial bromelioids, low-elevation *Puya* and high-elevation *Puya* species (Supplementary Table 3). There were also significant differences in residual PEPC activity according to the duration of treatment time at 50 °C in both groups and species. According to the *post hoc* analysis, there were significant differences in the pairwise comparisons between *A. nudicaulis* vs. *P. harmsii* and *A. nudicaulis* vs. *P. castellanosi* (Table 4). The *post hoc* analysis

revealed that there were significant differences in the residual PEPC activity between lowland equatorial *vs.* high-elevation *Puya* species, and between equatorial *vs.* low-elevation *Puya* species (Table 5). There were no significant differences between low-elevation *Puya vs.* high-elevation *Puya* species in their residual PEPC activity after high-temperature treatment.

Table 4. Summary of the *post hoc* analysis using Tukey-adjusted pairwise comparisons of the residual PEPC activity after incubation at 50 °C for 5 minutes. The pairwise differences between species are based on the least square means. The *p*-values in bold (*p* < 0.05) show significant differences over time in the relative residual PEPC activity between each pair of species. Degree of freedom = 29.

Species	Least-square mean	Standard error	<i>t</i>-statistic	<i>p</i>-value
<i>Puya castellanosii vs. Aechmea nudicaulis</i>	0.2530	0.0698	3.6236	0.0168
<i>Puya harmsii vs. Aechmea nudicaulis</i>	0.2197	0.0658	3.3383	0.0335
<i>Puya alpestris vs. Aechmea nudicaulis</i>	0.1879	0.0658	2.8545	0.0981
<i>Puya castellanosii vs. Bromelia humilis</i>	0.1733	0.0698	2.4819	0.2029
<i>Puya chilensis vs. Aechmea nudicaulis</i>	0.1759	0.0806	2.1819	0.3354
<i>Puya harmsii vs. Bromelia humilis</i>	0.1400	0.0658	2.1274	0.3642
<i>Puya castellanosii vs. Ananas comosus</i>	0.1252	0.0698	1.7929	0.563
<i>Puya alpestris vs. Bromelia humilis</i>	0.1082	0.0658	1.6436	0.6566
<i>Aechmea nudicaulis vs. Ananas comosus</i>	0.1278	0.0806	1.5854	0.6922
<i>Puya harmsii vs. Ananas comosus</i>	0.0919	0.0658	1.3966	0.7991
<i>Puya castellanosii vs. Puya alpestris</i>	0.0651	0.0520	1.2509	0.868
<i>Puya chilensis vs. Bromelia humilis</i>	0.0962	0.0806	1.1932	0.8911
<i>Puya chilensis vs. Puya castellanosii</i>	0.0771	0.0698	1.1042	0.9217
<i>Bromelia humilis vs. Aechmea nudicaulis</i>	0.0797	0.0806	0.9887	0.9525
<i>Puya alpestris vs. Ananas comosus</i>	0.0601	0.0658	0.9128	0.9675
<i>Puya harmsii vs. Puya alpestris</i>	0.0318	0.0465	0.6842	0.9925
<i>Puya harmsii vs. Puya chilensis</i>	0.0438	0.0658	0.666	0.9935
<i>Puya harmsii vs. Puya castellanosii</i>	0.0332	0.0520	0.6389	0.9948
<i>Bromelia humilis vs. Ananas comosus</i>	0.0481	0.0806	0.5967	0.9964
<i>Puya chilensis vs. Ananas comosus</i>	0.0481	0.0806	0.5965	0.9964
<i>Puya chilensis vs. Puya alpestris</i>	0.0120	0.0658	0.1823	1

Table 5. Summary of the *post hoc* analysis using Tukey-adjusted pairwise comparisons of the residual PEPC maximum activity after incubation at 50 °C. The pairwise differences between equatorial bromelioid species, low-elevation *Puya* and high-elevation *Puya* species are based on the least square means. The *p*-values in bold ($p < 0.05$) show significant differences over time in relative residual PEPC activity between each pair of species groups. Degree of freedom = 33.

Groups	Least-square mean	Standard error	<i>t</i>-statistic	<i>p</i>-value
Low-elevation <i>Puya</i> spp. <i>vs.</i> High-elevation <i>Puya</i> spp.	0.0481	0.0379	1.2705	0.4215
Equatorial bromelioid spp. <i>vs.</i> High elevation <i>Puya</i> spp.	0.1639	0.0412	3.9723	0.0010
Equatorial bromelioid spp. <i>vs.</i> Low-elevation <i>Puya</i> spp.	0.1157	0.0431	2.6823	0.0297

There were also significant differences in residual PEPC activity after 5 min of incubation at 50 °C between species (Fig. 3B; $F = 429.9$; $df = 6$; $p < 0.0001$). The *post hoc* analysis showed that all species differed significantly except *P. castellanosi* *vs.* *P. harmsii* ($p = 0.7720$) and *P. alpestris* *vs.* *P. chilensis* ($p = 0.7473$). The ANOVA test showed significant differences in the residual PEPC activity between lowland equatorial bromelioids, low-elevation *Puya* species and high-elevation *Puya* species ($F = 563.4$; $df = 3$; $p < 0.0001$). High-elevation *Puya* species exhibited less stability of PEPC activity at 50 °C than any other species evaluated in the present study.

Relative reduction in PEPC activity at 5 °C

The ratio between the PEPC activity at 25 °C and 5 °C (at saturating substrate concentration) was compared between species in order to establish whether the PEPC activity retained at the lower temperature was higher in high-elevation species than in low-elevation species and lowland equatorial bromelioids. All species showed lower PEPC activity at 5 °C than at 25 °C, expected from standard enzyme kinetics. The species that exhibited the lowest 5:25 ratio of PEPC activity were *A. nudicaulis*, *A. comosus*, and *B. humilis* (Supplementary Table 4 and Fig. 4). There were significant differences between species

in their 5:25 PEPC activity ratio ($F = 61.7$; $df = 6$; $p < 0.0001$). The Tukey HSD test showed that species belonging to the same group (lowland equatorial bromelioids, low-elevation *Puya* species, or high-elevation *Puya* species) did not differ significantly in their mean 5:25 PEPC activity ratios (Table 6).

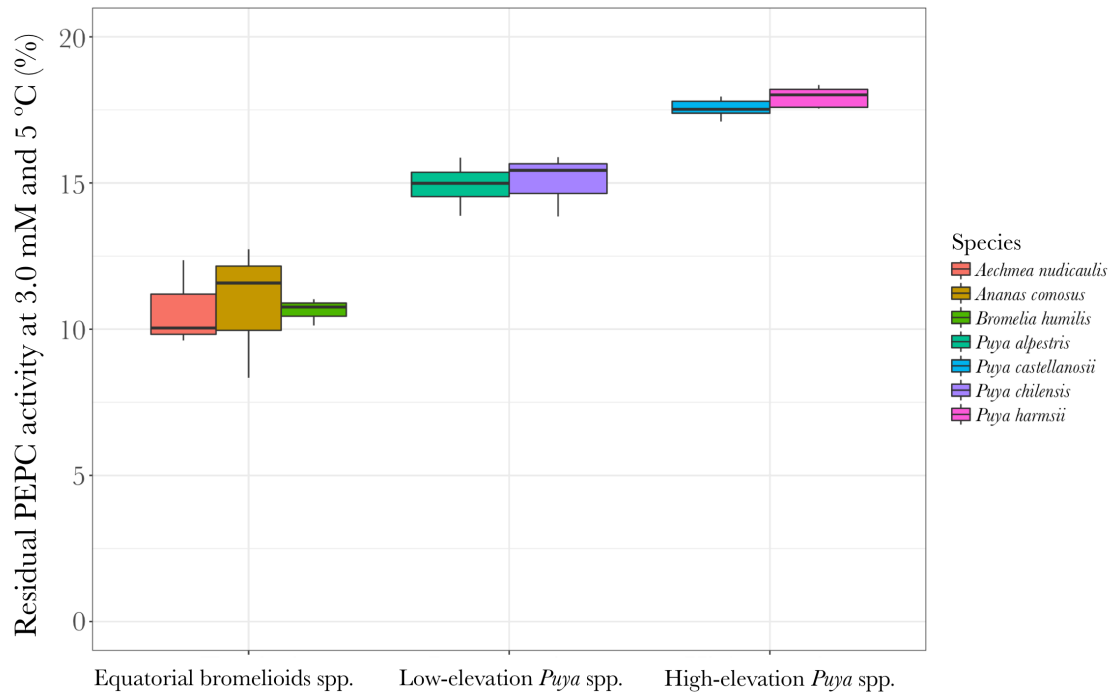


Figure 4. Low-temperature activity of PEPC, measured at 5°C, for four species of *Puya* and three species of the subfamily Bromelioideae, expressed as a percentage of the activity recorded in the same extract at 25 °C. According to the ANOVA test, there were differences in the relative low-temperature PEPC activity between the species ($F = 61.7$; $df = 6$; $p < 0.0001$) and between equatorial bromelioid species, low-elevation *Puya* species and high-elevation *Puya* species ($F = 204.6$; $df = 2$; $p < 0.0001$). The horizontal thick line in the boxplot indicates the median value. The whiskers represent data from the 25th or 75th percentile until the furthest data point that is within 1.5 times the inter-quartile range.

Table 6. Summary of the Tukey HSD pairwise differences in the relative PEPC activity at 5°C compared with 25°C between species. The *p*-values in bold (*p* < 0.05) indicate significant differences of between the paired species.

Species	Mean difference (%)	<i>p</i>-value
<i>Puya alpestris</i> vs. <i>Aechmea nudicaulis</i>	4.34	< 0.0001
<i>Puya castellanosi</i> vs. <i>Aechmea nudicaulis</i>	6.89	< 0.0001
<i>Puya chilensis</i> vs. <i>Aechmea nudicaulis</i>	4.39	< 0.0001
<i>Puya harmsii</i> vs. <i>Aechmea nudicaulis</i>	7.27	< 0.0001
<i>Puya alpestris</i> vs. <i>Ananas comosus</i>	4.12	< 0.0001
<i>Puya castellanosi</i> vs. <i>Ananas comosus</i>	6.67	< 0.0001
<i>Puya chilensis</i> vs. <i>Ananas comosus</i>	4.17	< 0.0001
<i>Puya harmsii</i> vs. <i>Ananas comosus</i>	7.06	< 0.0001
<i>Puya alpestris</i> vs. <i>Bromelia humilis</i>	4.37	< 0.0001
<i>Puya castellanosi</i> vs. <i>Bromelia humilis</i>	6.92	< 0.0001
<i>Puya chilensis</i> vs. <i>Bromelia humilis</i>	4.42	< 0.0001
<i>Puya harmsii</i> vs. <i>Bromelia humilis</i>	7.31	< 0.0001
<i>Puya harmsii</i> vs. <i>Puya alpestris</i>	2.94	< 0.0001
<i>Puya castellanosi</i> vs. <i>Puya alpestris</i>	2.55	0.0001
<i>Puya harmsii</i> vs. <i>Puya chilensis</i>	2.89	0.0005
<i>Puya chilensis</i> vs. <i>Puya castellanosi</i>	-2.50	0.0060
<i>Puya harmsii</i> vs. <i>Puya castellanosi</i>	0.39	0.9779
<i>Bromelia humilis</i> vs. <i>Ananas comosus</i>	-0.25	0.9998
<i>Ananas comosus</i> vs. <i>Aechmea nudicaulis</i>	0.21	0.9999
<i>Bromelia humilis</i> vs. <i>Aechmea nudicaulis</i>	-0.03	1.0000
<i>Puya chilensis</i> vs. <i>Puya alpestris</i>	0.05	1.0000

There were significant differences in the mean of the 5:25 PEPC activity ratio between the groups mentioned above ($F = 204.6$; $df = 2$; $p < 0.0001$). The Tukey HSD test showed that there were significant differences between all of the groups (Table 7). The present results suggest that the retention of PEPC activity at low temperature is superior in high-elevation *Puya* species to that of any other group evaluated in the present study.

Table 7. Summary of the Tukey HSD pairwise differences in the relative PEPC activity at 5°C compared with 25°C between equatorial bromelioid species, low-elevation *Puya* species and high-elevation *Puya* species. The *p*-values in bold ($p < 0.05$) indicate significant differences between the paired species groups.

Groups	Mean	
	difference (%)	<i>p</i>-value
Low-elevation <i>Puya</i> spp. <i>vs.</i> High-elevation <i>Puya</i> spp.	−2.77	< 0.0001
Equatorial bromelioid spp. <i>vs.</i> High-elevation <i>Puya</i> spp.	−7.06	< 0.0001
Equatorial bromelioid spp. <i>vs.</i> Low-elevation <i>Puya</i> spp.	−4.29	< 0.0001

Discussion

The purpose of measuring cold resistance, PEPC temperature optimum and PEPC thermal lability at high (50 °C) and low (5 °C) temperatures in species from different origin, in a common-garden experiment, was to understand whether high-elevation CAM species showed heritable molecular adaptations to lower temperatures. The following points will be discussed by reference to these results: (1) cold resistance in *Puya*; (2) PEPC temperature optimum at saturated conditions in different *Puya* species; (3) thermal stability at high temperatures of PEPC; (4) relative reduction of activity at 5 °C; and (5) the evolution of cold resistance in *Puya*.

The common-garden design facilitated understanding of the components of possible adaptation or exaptation (McKay *et al.*, 2001). Such investigations are valuable because they allow the genetic basis of phenological, physiological and allometric traits to be understood and because they reduce environmental noise, which could lead to plasticity in plants (de Villemereuil *et al.*, 2016). Moreover, the common-garden design permits genetically encoded adaptations reflecting the longer-term evolutionary histories of the organisms or lineages concerned to be distinguished from shorter-term acclimatory responses to the local environment. In this case the acclimation conditions did not include

low temperatures. For this reason, the present results demonstrate that there is a genetic component of cold resistance in *Puya* species. This component is not lost by plasticity in common-garden conditions where the plants did not experience low temperatures. The results obtained in the present study therefore demonstrate the physiological and biochemical evidence linking CAM with high-elevation species in *Puya*. In Chapter 5 other studies of non-CAM plants from high elevation environments are reviewed and discussed. As well as a new hypothesis of evolution of CAM in tropical alpine environments.

Cold resistance

Tropical plants may be highly damaged when exposed to freezing conditions (Körner, 2012). However, even equatorial plants can increase their cold resistance by acclimation to low temperatures (Nobel *et al.*, 1998). A similar process occurs in high-latitude species (e.g. *Allium × proliferum*; Wisniewski *et al.*, 2014), in which acclimated plants suffer less damage at low temperatures than non-acclimated plants. Consequently, it is notable that sub-equatorial bromelioids and *Puya* species can be shown to possess higher cold resistance than equatorial bromelioids, even when plants have never been cold acclimated, as this implies the existence of a genetically based adaptations in these distinct lineages to their distinct temperature regimes.

The T_{50} reported in the present study of *Aechmea nudicaulis* (-7.45 ± 0.44 °C; Table 1) is similar to the T_{50} of *Aechmea blanchetiana* reported in Neves-Chaves *et al.* (2015) in field conditions (day/night temperatures: 23.1–12.4 °C) in the distal part of the leaf (-7.72 ± 0.69 °C). Neves-Chaves *et al.* (2015) showed that the basal part of the leaf in *Aechmea blanchetiana* exhibit the greatest T_{50} in the plant. CAM does not operate in the basal part of the leaf in Bromeliaceae (Freschi *et al.*, 2010; Ming *et al.*, 2015). Therefore, the evaluation of cold resistance in the basal part of the leaf is not relevant for understanding

the adaptations of CAM to high-elevations. However, the base of the leaf could have provided protection to the youngest meristem in the centre of the rosette as shown in other paramo and afro alpine species such as *Espeletia schultzei* and *Lobelia keniensis* (Beck, 1994).

High-elevation *Puya* species have a higher cold resistance than all other species evaluated in the present study. Nobel (1982) demonstrated that cacti from Argentina and Chile higher mid-elevation (~2000 m) showed higher cold resistance than their counterparts distributed 100–400 m below in cold-acclimated conditions (4-6 weeks, day/night temperatures 10°/0° C). The present study is the first report that demonstrates that high-elevation tropical CAM species exhibit higher cold resistance than their close relatives from low-elevations in non-cold acclimated conditions (difference of elevation range ~3000 m). The T_{50} values for *Puya castellanosi* (25°18'S north Argentina; 3000 m; T_{50} : -14.38 ± 0.24 °C) and *P. harmsii* (22°S–23°S north Argentina; 2000–3000 m; T_{50} : -14.48 ± 0.49 °C) are higher than the highest values reported by Nobel (1982) for high-elevation cacti in Argentina and Chile; *Eriosyce ceratistes* (32°54'S central Chile; 1940 m; T_{50} : -10.1 ± 0.8 °C) and *Denmoza rhodacantha* (32°26'S northcentral Argentina, 2260 m; T_{50} : -10.4 ± 0.8 °C). The cacti reported in Nobel (1982) were cold acclimated as mentioned above. This means that if the cold resistance also has a plastic component, under such conditions *P. castellanosi* and *P. harmsii* will show even higher cold resistance. For future research the cold resistance and thermal lability of PEPC (at high and low temperatures) should be evaluated under cold-acclimated conditions to measure the plasticity degree of cold resistance in *Puya*.

Given the low nocturnal temperatures that high-elevation species experience on a daily basis, it is likely that high-elevation *Puya* species exhibit a higher cold resistance than equatorial plants in low-temperature acclimated conditions. This is because the

diurnal temperature range is higher than the annual temperature range in equatorial and subequatorial mountain environments (Rundel, 1994). Coastal areas in Chile also experience a wide diel temperature range because at night oceanic currents cool down the coast (Falvey and Garreaud, 2009). This could explain why low-elevation *Puya* species (from central and northern Chile) and non-equatorial bromelioids (from central Chile) showed higher cold resistance than equatorial bromelioids in non-cold acclimated environments. *Ochagavia carnea* showed a significantly higher T_{50} (less cold resistant) than *Fascicularia bicolor*. This result could be explained by the differences in latitude, as *O. carnea* has a more northerly distribution (32 °S–38 °S) than *F. bicolor* (34 °S–42 °S) (Zizka *et al.*, 1999, 2002, 2009).

Cold tolerance or supercooling?

The factor which prevents plants from colonising areas which experience frost is not the low temperature stress but rather the freezing damage to plant tissue (Körner, 2012; Zanne *et al.*, 2014). Plants have evolved mechanisms to cope with temperatures below zero and to avoid intracellular ice formation (Zanne *et al.*, 2014). These mechanisms are classified as escaping, avoiding and tolerating sub-zero temperatures (Körner, 2012).

Escaping from sub-zero temperatures is a common strategy in annual plants of higher latitudes, in which their life cycle is completed before they experience the freezing conditions. It is also found in deciduous trees of temperate latitudes, which shed their leaves before winter to prevent freezing damage in the leaves (Körner, 2012). However, deciduous trees have other mechanisms to protect buds, phloem and roots from freezing damage (deep super cooling). Plants can also avoid ice formation in the protoplast by depressing the freezing point without a solute concentration mechanism. These plants keep their protoplasm in a gel-like state, which prevents ice nucleation, a mechanism

called freezing avoidance or supercooling (Rada *et al.*, 2001). Finally, cold-tolerant plants prevent intracellular ice formation. Ice formation within the protoplast causes irreversible damage to the cell and is considered universally lethal (Guy, 1990; Körner, 2012). If the formation of ice occurs in the intercellular spaces, this could lead to dehydration. However, cold-tolerant plants have mechanisms, such as more fluid membrane and protective proteins and sugars, which coat the three-dimensional structure of the cell (Körner, 2012).

The present study used freezers at fixed temperatures to test the effects of low temperature rather than gradual cooling (Sklenář *et al.*, 2010, 2012), because the main goal was to measure freezing resistance rather than cold hardening (Larcher, 2010; Sierra-Almedia *et al.*, 2016). Furthermore, the present study did not specifically investigate what mechanisms *Puya* species have evolved to cope with low temperatures. Based on the investigations of cold resistance in the paramo flora of Ecuador, Sklenář *et al.* (2012) showed that species with a non-equatorial origin were more likely to be cold-tolerant plants, while the species of equatorial origin were likely to be supercooling plants. Therefore, it is likely that the cold-resistance mechanism in *Puya* would be cold tolerance rather than supercooling, given that *Puya* is have a non-equatorial origin (Jabaily and Sytsma, 2013; Chapter 2). However, it is necessary to measure the nucleation temperature to determine whether *Puya* species are cold-tolerant or supercooling plants in non-cold acclimated and cold acclimated plants. This is relevant because in central Boliva the high-elevation species *Polylepis tarapacana* exhibit supercooling in the wet warm season while in the cold dry season exhibit cold tolerance (Rada *et al.*, 2001).

Temperature optimum of PEPC

There is a preponderance of cold-adapted enzymes that have lower temperature optima than non-cold adapted enzymes in psychrophile bacteria, archaea and yeasts

(Georlette *et al.*, 2004; Siddiqui and Cavicchioli, 2006). This is not the case with the temperature optimum of PEPC in high-elevation species of *Puya* compared with the low-elevation species of bromeliads tested (Fig. 2, Supplementary Table 1). *Puya alpestris* spp. *zoellneri* showed the lowest value of T_{opt} of PEPC activity, while the T_{opt} of the high-elevation species was higher and the one-way ANOVA test showed that there are no significant differences between high- and low-elevation species of *Puya* in their PEPC T_{opt} (Supplementary Table 1). This result contradicts the hypothesis that PEPC of high-elevation *Puya* species would have a lower temperature optimum than low-elevation species of *Puya*. The kinetic differences of the PEPC from high *vs.* low-elevation CAM species is not associated with the temperature optimum, but it could be related to the relative performance of PEPC at lower temperatures. The temperature optimum of PEPC under saturating substrate conditions found in the present study are similar to the results reported for *Kalanchoe daigremontiana* (Buchanan-Bollig and Kluge, 1981; Buchanan-Bollig *et al.*, 1984).

The temperature of the assay influences the kinetics of the response of the enzyme to PEP (substrate), G6P (positive effector) and malate (negative effector) (Buchanan-Bollig and Kluge, 1981; Buchanan-Bollig *et al.*, 1984). The concentration of G6P can increased the temperature optimum of PEPC, while higher concentration of malate (20 mM) can reduce the temperature optimum of PEPC, this was evaluated in *Kalanchoe daigremontiana* (Buchanan-Bollig *et al.*, 1984). Therefore, it is possible that the temperature optimum of PEPC from high-elevation species might differ from that of low-elevation species under conditions of non-saturating substrate concentrations. In the future, this hypothesis could be evaluated, measuring the temperature dependence of PEPC in non-saturated conditions, different pH and different concentration of activators and inhibitor effectors, that might be closer to the conditions experienced by the plants.

Thermal lability of PEPC at high temperature

The thermal lability of PEPC at 50 °C was evaluated in constitutive CAM plants from equatorial and non-equatorial environments, including those occurring at both low elevations (<1000 m) and high elevations (>3000 m). The general trend was that there was a progressive reduction in PEPC activity over time when the extract was incubated at 50 °C (Fig. 3B). However, the lowland equatorial bromelioids showed a greater stability of PEPC at high temperature than either the high- or low-elevation *Puya* species (Fig. 3A). The general trend can be explained by the negative effect of high temperatures on the tetrameric conformation of PEPC. PEPC in the tetrameric state is more active than the dimer state, or monomer state as demonstrated in *Crassula argentea* by Wu and Wedding (1987).

Cold-adapted enzymes are generally more flexible and indeed tend to unfold into an inactive configuration at moderately high temperatures (Fields and Somero, 1998; Somero, 2004; Siddiqui and Cavicchioli, 2006). This is interpreted as a trade-off between enzymatic flexibility and temperature performance. The higher the flexibility of an enzyme, the higher will be the activity of the enzyme at lower temperatures, but the enzyme will also be more easily inactivated by high temperatures (Siddiqui and Cavicchioli, 2006). The fact that the PEPC activity of *Puya* species was impaired more rapidly than that of lowland equatorial bromelioids might therefore imply that the flexibility of the enzyme in equatorial species is lower than that of *Puya* species. This interpretation would correlate well with the relatively high cold resistance observed in *Puya*.

After high-temperature treatment at 50 °C, it was possible to separate not only the temperature lability of PEPC from lowland equatorial bromelioids from *Puya* species, but also to resolve differences in PEPC stability in high- *vs.* low-elevation *Puya* species.

This result implies that PEPC of high-elevation species has more flexibility than the low-elevation species, presumably as a consequence of local adaptation to high-elevation environments. It is necessary to investigate more deeply whether there are amino-acid differences between the CAM PEPC isoforms in equatorial bromelioids and *Puya* species that might be involved in the contact region of the monomers of PEPC, or possibly in the reaction centre as is the case in other cold-adapted enzymes (Siddiqui and Cavicchioli, 2006). It is also necessary to investigate more deeply the amino acid differences between high- and low-elevation *Puya* species. For future research the kinetics can be studied by cloning the protein from high- and low-elevation *Puya* into a bacterial vector to clarify whether the enzyme that is more flexible is PEPC or malate dehydrogenase or both. This is very important to test because the PEPC activity protocol is a coupled enzyme protocol and therefore the two enzymes are involved in the reaction.

Relative reduction in PEPC activity at low temperature

All species considered in this study showed a reduction in PEPC activity at 5 °C, as would be expected from basic enzyme-kinetic theory. However, *Puya* species showed a lesser reduction in PEPC activity at 5 °C (relative to its activity at 25 °C) than was the case with the lowland equatorial bromelioids (Fig. 4). This implies that the flexibility of PEPC in *Puya* species, suggested initially by the experimental result of the thermal lability at 50 °C, might also be responsible for the relatively superior performance of PEPC at 5 °C in *Puya* species. The relative performance of PEPC was even higher in high-elevation species than in low-elevation species of *Puya* (Fig. 4). This result demonstrates local adaptation of CAM species in high-elevation environments in the key enzyme PEPC. The reduction of PEPC activity at 5 °C in high-elevation species on average was 83%. This reduction in activity could decrease if the plants were acclimated at lower temperatures at night, as experienced in their natural environments, for example the as the cold-resistance in

Opuntia fragilis is increased in comparison to cold acclimation and non-cold acclimation (Loik and Nobel, 1993b). Yamori *et al.* (2014) demonstrated in *Kalanchoe daigremontiana* that during relative cold-acclimation (day/night: 20 °C/10 °C) showed a lower T_{opt} (16.1 °C) for the nocturnal CO₂ fixation rate than the T_{opt} (21.1 °C) for the nocturnal CO₂ fixation rate in warmer acclimation conditions (day/night: 30 °C/20 °C).

Evolution of *Puya* and colonisation of the Andes

The early-divergent genera of Bromelioideae, *Ochagavia* and *Fascicularia*, and the early-divergent species of *Puya* are distributed in Chile, towards the southern latitudinal limit of distribution of the family Bromeliaceae (Zizka *et al.*, 1999, 2002). The species included in the present study confirm that the cold resistance has a genetic component. Moreover, the fact that the early-divergent species of *Puya* and the early-divergent species of Bromelioideae show intrinsic cold resistance independent of prior acclimation suggests that the common ancestor of these two sister clades was a cold-resistant plant. This could explain why the great majority of *Puya* are montane species. The cold resistance of the subequatorial common ancestor may have allowed the diversification of the genus in relatively colder environments emerging with the rise of the Andes (Chapter 1). Further, the conformational flexibility in the PEPC protein from the low-elevation common ancestor could have allowed CAM species to colonise higher-elevation environments, and once there acquire even more flexibility in the conformation of PEPC as an adaptation to the low-temperature environments.

Conclusions

The present study, conducted with plants grown in a common-garden environment, demonstrates that the cold resistance observed in *Puya*, in comparison with other bromeliads, has a genetic, heritable component that reflects adaptation to the higher-latitude or higher-elevation environments in which these species typically grow.

All species of *Puya* showed a higher cold resistance than the lowland equatorial bromeliads, suggesting that the common ancestor of *Puya* was most likely a cold-resistant plant. This in turn could explain why the genus was so successful in colonizing the Andes and why the majority of present-day *Puya* species are montane plants. The typical habitats occupied by *Puya* are also reflected in the thermal lability of the key enzyme PEPC from these species at high temperature, but its retention of relatively greater activity at low temperatures compared with the enzyme from tropical lowland bromeliads. Based on precedents from a number of other enzymes studied from organisms living in low-temperature environments, these features of PEPC activity might be attributable to a higher conformational flexibility of the enzyme in high-elevation compared with low-elevation *Puya* species, although this remains to be investigated. Whatever the precise cause at the protein level of this difference in temperature dependence of PEPC activity, this is the first report of enzymatic differences between high- and low-elevation CAM species. Further discussion of the evolution of CAM in high-elevation environments are described in Chapter 5.

Chapter 4. Evolution of phosphoenolpyruvate carboxylase in monocots with emphasis on the origin of CAM in the genus *Puya*

Abstract

Crassulacean acid metabolism is a complex physiological plant trait that maximises water use efficiency by fixing CO₂ at night-time. The enzyme responsible for nocturnal CO₂ fixation is phosphoenolpyruvate carboxylase (PEPC). To the date there is no evidence of polymorphism in this enzyme to distinguish CAM and C₃ species. The genus *Puya* contains both C₃ and CAM species, but the ancestral photosynthetic pathway of the genus is currently undetermined. It is possible to clarify the ancestral photosynthetic pathway of the genus using transcriptome (RNA-seq) data and by investigating the evolution of key genes of CAM such as the PEPC gene (*ppc-1*). The present study uses RNA-seq data from eight *Puya* species (CAM and C₃ species) in order to identify differences in: (1) the expression of 24 genes involved in CAM; and (2) whether these genes evolved under positive selection. In addition, the sequences of *ppc-1* from five constitutive CAM species and seven closely related C₃ species (belonging to the orders Asparagales and Poales) were analysed to determine the type of selection under which *ppc-1* evolved. Furthermore, in addition to the 16 sequences of *Puya* (which were identified with the RNA-seq data), 45 sequences of the *ppc-1* (from 27 species) were analysed and incorporated into a phylogenetic analysis to establish where the most expressed isoforms of *ppc-1* gene of *Puya* belonged. According to the differences in CAM gene expression, the species *P. raimondii* is the only C₃ species that is statistically different from *P. chilensis* (CAM).

All other CAM and C₃ species did not show significant differences in the CAM gene expression. One third of genes involved in CAM evaluated in the present study evolved under positive selection, possibly as a consequence of rapid diversification across the Andes, but this is uncorrelated with CAM evolution in the genus. The PEPC CAM isoform evolved under positive selection in monocots and there are two residues (534 and 679) which differentiate CAM and C₃. However, only residue 679 had a higher posterior probability. This is the first report of a polymorphism in the PEPC sequence of constitutive CAM in monocots. CAM and weak CAM species contained an Arg679 or Asn679 residue, while C₃ species or C₃ isoforms from CAM species contained a His679 residue. The majority of C₃ *Puya* species evaluated in the present study showed the polymorphism. This suggests either a recent loss of CAM or weak CAM species. The only species that did not exhibit the polymorphism was *P. raimondii*, which belongs to a clade where the most likely ancestral photosynthetic pathway is C₃. Given the phylogenetic position of *P. alpestris*, *P. coerulea*, and *P. ferruginea*, the most likely ancestral photosynthetic pathway of the genus is weak CAM. The phylogenetic reconstruction of *ppc-1* isoforms suggests that CAM evolved in two different clades: one in Asparagales (*ppc-1Ma*); and the other in Poales (*ppc-1Mb*).

Introduction

Understanding the origin of complex traits has been a frequent question in evolutionary biology (Hancock and Edwards, 2014). Crassulacean acid metabolism (CAM) is a complex physiological trait that encompasses anatomical and physiological changes as well as a tightly regulated gene network operating in a circadian rhythm (Borland *et al.*, 2014). The evolution of CAM has been studied in families of angiosperms such as Orchidaceae (Silvera *et al.* 2009, 2010 a,b, 2014), Asparagaceae (Agavoideae; Heyduk *et al.*, 2016), Cactaceae (Edwards *et al.*, 2005), Clusiaceae (Gehrig *et al.*, 2003),

Crassulaceae (Mort *et al.*, 2001), Euphorbiaceae (Horn *et al.*, 2014) and Bromeliaceae (Crayn *et al.*, 2004; Givnish *et al.*, 2011, 2014). The majority of these studies used a molecular phylogenetic approach in addition to carbon isotope values to assess the evolution of the photosynthetic pathway. The carbon isotopic composition is a relatively quick and accurate approach to identify constitutive CAM species (Silvera *et al.*, 2010; Crayn *et al.*, 2015). It is also an excellent tool if the taxon-sampling is relatively large and the taxa are distributed in a wide range of latitudes and longitudes (Chapter 2). For example, determining the carbon isotopic composition of herbarium specimens is a convenient approach to evaluate the photosynthetic pathway in large-scale taxon-sampling in a relatively short time. However, carbon isotope values can only identify constitutive CAM species and not weak CAM species (Winter *et al.*, 2015).

The use of molecular phylogenies has helped to clarify the evolution of complex traits (Christin *et al.*, 2010b; Hancock and Edwards, 2014). However, it is necessary to gather further evidence from physiological, biochemical and ecological empirical data to support the models (Christin *et al.*, 2010a), especially to identify intermediate states. For example, Silvera *et al.* (2014) assessed the photosynthetic pathway of 11 species of the subtribe Oncidiinae (Orchidaceae) using titratable acidity and CO₂ gas exchange. This approach is optimal to evaluate very precisely the CAM expression in plants and to identify correctly constitutive CAM, constitutive C₃ and weak CAM species. However, in these types of experiments the taxon sampling is usually limited and requires a long time to test the induction of any expression of CAM. Therefore, it is desirable to find a molecular, physiological or biochemical signature to identify not only constitutive CAM but also weak CAM species. This signature can combine the large-scale advantages of carbon isotope values while also distinguishing weak CAM from C₃ species.

One of the key enzymes of CAM photosynthesis, and C₄ photosynthesis, is phosphoenolpyruvate carboxylase (PEPC; EC 4.1.1.31). PEPC is responsible for fixing nocturnal CO₂ into oxaloacetate (OAA). Molecular signatures have been documented in the PEPC of C₄ plants (Christin and Besnard 2009, Christin *et al.*, 2007). For example, in C₄ species the C₄-PEPC isoform has a serine residue in the position 780 (in maize C₄-PEPC numbering) while the C₃-isoform has an alanine residue. However, not all C₄ species in Chenopodiaceae have the molecular signature in the residue 780 (Rosnow *et al.*, 2014). C₃ species closely related to C₄ species do not have the Ser780 in their homolog isoform of C₄-PEPC. Therefore, this residue can help identify C₄ species relatively easily only in particular clades. Further, there are functional polymorphisms between C₃ and C₄ plants described in the genus *Flaveria*; Arg884 in C₃ to Gly884 in C₄ species (*Flaveria* numbering; Paulus *et al.*, 2013). These signatures are extremely valuable in order to recognise species with intermediate complex traits or species that have lost the complex trait. In CAM plants, the quest for evolutionary signatures has not been very successful, in particular in the key enzyme PEPC (Silvera *et al.*, 2014). Perhaps the reason for this lies in the limited length of the sequences and the narrow taxonomic range of species evaluated.

PEPC is present not only in CAM and C₄ plants; it is also found in bacteria, cyanobacteria, green algae and all land plants (Chollet *et al.*, 1996; O’Leary *et al.*, 2011). In land plants, PEPC has many functions. For example, PEPC controls the movement of guard cells (Asai *et al.*, 2000), maintains cellular pH (Shi *et al.*, 2015), is involved in fruit maturation (Guillet *et al.*, 2002) and seed germination (Blonde and Plaxton, 2003), supplies carbon to bacteria in the root nodules of species in the family Fabaceae (White *et al.*, 2007), among others. The biochemistry of PEPC in plants and bacteria has been investigated since its discovery in 1953 (Bandurski and Griner, 1953). There is a good

understanding of the kinetics (Chollet *et al.*, 1996), structure (Kai *et al.*, 2003; Izui *et al.*, 2004) and regulation (Nimmo, 2000) of PEPC. The characteristics of the CAM and C₄ PEPC isoforms are the result of neofunctionalisation, possibly involved with gene duplications, regulatory changes and amino acid substitutions (Zhang, 2003; Christin *et al.*, 2010a, 2010b; Silvera *et al.*, 2014).

The evolution of PEPC has been investigated over the last 22 years. Initially, only partial sequences of the *ppc* gene were used in phylogenetic reconstructions due to limitations in PCR (Gehrig *et al.*, 1995, 1998, 2001, 2005; Christin and Besnard, 2009; Christin *et al.*, 2010a, 2010b, 2013, 2014; Silvera *et al.*, 2014). With the partial sequences, it was possible to conclude that there were two main isoforms in land plants: first the plant type (highly expressed; *ppc-1*); and second the bacterial type isoform (poorly expressed; *ppc-2*) (O’Leary *et al.*, 2011). Christin *et al.* (2014) demonstrates that the CAM and C₄ isoforms evolved only in the *ppc-1* gene. The evolution of *ppc-1* has been complex in angiosperms, with at least two independent duplications: one in monocots (*ppc-1M1* and *ppc-1M2*); and the other in eudicots (*ppc-1E1* and *ppc-1E2*) (Christin *et al.*, 2014), as well as more duplications within grasses and orchids (Christin *et al.*, 2007; Silvera *et al.*, 2014; Deng *et al.*, 2016). More recently, the full-length sequences of the *ppc* gene were obtained by sequencing the genome and/or transcriptome (RNA-seq) data (Gross *et al.*, 2013; Cai *et al.*, 2014; Ming *et al.*, 2015; Deng *et al.*, 2016). These data have been used to reconstruct the evolution of the *ppc* gene in monocots with a particular emphasis in orchids (Deng *et al.*, 2016; Zhang *et al.* 2016). Zhang *et al.* (2016) concluded that the CAM and C₄ isoform in monocots evolved in the clade *ppc-1M1*. Full length *ppc-1M1* and *ppc-1M2* sequences from constitutive CAM species and close C₃ relatives would be useful to establish whether there are polymorphisms in the *ppc* gene that could distinguish CAM from C₃ species using a selection test. RNA-seq data can be used to generate full-length sequences of the

ppc genes and also to understand the relative expression of different genes involved directly in CAM photosynthesis (Zhang *et al.*, 2016).

The evolution of CAM in *Puya* was complex and the photosynthetic pathway of the common ancestor is still undetermined (Chapter 2). Further, the southern and central Andes clade (48% CAM spp.) was more likely to produce more CAM species than the northern and central Andes clade (5% CAM spp.). Given the uncertainty of the ancestral photosynthetic pathway using carbon isotope values, it is important to use other methods to clarify the origin of CAM in the genus *Puya*. For example, the use of transcriptome sequence (RNA-seq) data in CAM and C₃ species belonging to different clades in *Puya* could add information to the differences between different clades in their relative expression of genes involved in CAM. Moreover, it is possible to test whether or not the genes involved in CAM evolved under positive selection as result of neofunctionalisation or adaptive radiation (Silvera *et al.*, 2014; Nevado *et al.*, 2016). The synonymous *vs.* non-synonymous substitutions in key CAM enzymes would provide evidence of rapid evolution. More importantly, the expression of genes involved in CAM photosynthesis by C₃ species of *Puya* belonging to different clades will clarify the evolution of CAM in *Puya*.

The present study: (1) investigated the differences in the relative expression of genes involved in CAM between CAM and C₃ species of *Puya*; (2) evaluated the positive selection of genes involved in CAM in *Puya*; (3) investigated whether or not *ppc-1* isoforms of CAM species evolved under positive selection, identifying CAM polymorphisms in the PEPC sequence in monocots; and (4) investigated the phylogenetic position of the most expressed isoforms in *Puya*.

Materials and Methods

Plant material

The species included in the present study were: *Puya alpestris* ssp. *alpestris* (C₃, blue clade; accession number 107550), *P. alpestris* ssp. *zoellneri* (CAM, blue clade; accession number 020375), *P. chilensis* (CAM, NC clade; accession number 090299), *P. coerulea* (C₃, blue clade; accession number 120104), *P. ferruginea* (C₃, early divergent species of core *Puya* clade; accession number 100990), *P. mirabilis* (C₃, SC clade; accession number 130471), *P. raimondii* (C₃, NC clade; accession number 120001) and *P. spathacea* (C₃, SC clade; accession number 100741). The plant tissue was collected in Ventnor Botanic Garden in May 2015. The tissue was collected in the afternoon. The tissue was snap frozen in liquid nitrogen and stored at –80 °C until RNA extraction. The plant tissue consisted of the middle part of a mature and healthy leaf.

RNA extraction

CAM plants are usually rich in polysaccharides and phenols. These characteristics make it very difficult to extract good quality RNA, especially for RNA-seq purposes. In the present study, a new RNA extraction protocol was developed to obtain good quality and good yield RNA for species rich in polysaccharides and phenols. The protocol was an improved version of the protocol used by Gehrig *et al.* (2000). The RNA was extracted for the species described above. A detailed description of the RNA extraction protocol is provided in the supplementary section of this Chapter (Supplementary Material 1). RNA integrity was evaluated with the Agilent 2100 Bioanalyzer.

RNA-seq assemble and relative expression of genes involved in CAM in *Puya*

The transcriptome of each species described above was obtained by the Illumina Hi-seq sequencing platform with paired-end reads of size 100–125 bp. The reads were done at the University of Edinburgh. The assembly of the transcripts and relative

expression were completed as described in Nevado *et al.* (2016). Short sequence reads were quality-checked for each species using *fastqc* v 0.11 (available from <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Cutadapt v 1.4 (Martin, 2011) was used to remove illumina adaptors, and trimmomatic v 0.32 (Bolger *et al.*, 2014) to trim reads. Reads smaller than 36 bp after trimming were discarded. To assemble a *de novo* transcriptome for each individual species, trinity v r20140413p1 (Grabherr *et al.*, 2011) was used with default settings except for the minimum assembled contig length (set to 300 bp). The putative coding sequences (CDSs) within each transcript were identified using transdecoder (available from <https://transdecoder.github.io/>) using default settings.

A list of 19 key enzymes (25 genes and 44 isoforms; Table 1) participating in the nocturnal CO₂ fixation, import and export of malate to and from the vacuole, and day-time decarboxylation in CAM plants were used to screen for genes involved in CAM in *Puya*. Based on the *Ananas comosus* genome and Table 1, all the CAM and non-CAM isoforms were used to annotate the transcriptome of each *Puya* species. The homology of the genes found in *Puya* species with *A. comosus* genes was assessed by constructing a phylogenetic tree of each gene containing more than two isoforms. The relative expression of the genes in Table 1 was evaluated using Trinity Transcript Quantification (Grabherr *et al.*, 2011) and quantified by RSEM v1.3.0 (Li and Dewey, 2011). The gene expression was approximated by the relative abundance of transcripts. The relative abundance of transcripts was calculated by Transcripts Per Million (TPM). The differences in the TPM between CAM and C₃ species in *Puya* was evaluated using an ANOVA test. A post-hoc analysis was conducted to establish pairwise differences. The model incorporated interactions between isoforms, species and photosynthetic pathways.

The photosynthetic pathway was determined according to the carbon isotope values of the species described above (Chapter 2).

Table 1. List of the 25 genes involved in CAM photosynthesis included in the present study. The orthologous genes were identified in *Ananas comosus* after Ming *et al.* (2015).

Gene name	<i>Ananas comosus</i> orthologous gene
Aluminium-activated malate transporter	Aco003023.1
Aluminium-activated malate transporter	Aco010725.1
Beta carbonic anhydrase 2	Aco005402.1
Beta carbonic anhydrase 2	Aco006181.1
Beta carbonic anhydrase 5	Aco002732.1
Malate dehydrogenase	Aco004349.1
Malate dehydrogenase	Aco004996.1
Malate dehydrogenase	Aco006122.1
Malate dehydrogenase	Aco007734.1
Malate dehydrogenase	Aco008626.1
Malate dehydrogenase	Aco010232.1
Malate dehydrogenase	Aco013935.1
Malate dehydrogenase	Aco014690.1
Malate dehydrogenase	Aco017525.1
Malate dehydrogenase	Aco017526.1
Malate dehydrogenase	Aco017527.1
Malate dehydrogenase	Aco017528.1
Malate dehydrogenase	Aco019631.1
NAD-dependent malic enzyme	Aco016569.1
NAD-dependent malic enzyme 2	Aco007622.1
NADP-malic enzyme 1	Aco009967.1
NADP-malic enzyme 3	Aco005989.1
NADP-malic enzyme 4	Aco005631.1
Phosphoenolpyruvate carboxykinase 1	Aco017762.1
phosphoenolpyruvate carboxylase 1	Aco010025.1
phosphoenolpyruvate carboxylase 1	Aco018093.1
phosphoenolpyruvate carboxylase 2	Aco016429.1
Phosphoenolpyruvate carboxylase kinase 1	Aco010095.1
Phosphoenolpyruvate carboxylase kinase 1	Aco013938.1
Phosphoenolpyruvate carboxylase-related kinase 1	Aco022525.1
Phosphoenolpyruvate carboxylase-related kinase 2	Aco001261.1
PPDK regulatory protein	Aco014488.1
Pyrophosphate:fructose-6-phosphate 1-phosphotransferase	Aco017133.1
Pyrophosphate:fructose-6-phosphate 1-phosphotransferase	Aco018526.1
Pyruvate, orthophosphate dikinase	Aco024818.1
Tonoplast dicarboxylate transporter	Aco000795.1
V-ATPase A subunit	Aco008300.1
V-ATPase B subunit	Aco015226.1

V-PPiase (K ⁺ -dependent)	Aco008167.1
Vacuolar hexose transporter	Aco005379.1
Vacuolar invertases	Aco017533.1
Vacuolar invertases	Aco023030.1
Vacuolar sucrose transporter	Aco011916.1
Vacuolar sucrose transporter	Aco015779.1

Ppc-1 sequences in other plants

Multiple sequences of *ppc-1* from monocots (58 seq.; 34 spp.), basal angiosperms (2 seq.; 2 spp.) and conifers (1 seq.; 1 spp.) were retrieved from complete transcriptomes or genomes libraries available in the 1KP project (Matasci *et al.*, 2013), GeneBank, Phytozome (Ouyang *et al.*, 2007; Droc *et al.*, 2013). In addition, available data from studies of the genome and transcriptome of agaves, bromeliads, and orchids (Cai *et al.*, 2014; Gross *et al.*, 2013; Palma-Silva *et al.*, 2016). The *ppc-1* copies found in *Puya* were used as a query of a BLAST search against the transcriptomes mentioned above. Sequences were incorporated only if they aligned with the reference sequence along more than 1500 bp. The species and codes of *ppc-1* sequences are listed in Supplementary Table 1. The sequences were aligned using the MUSCLE function (Edgar, 2004) in SeaView (Gouy *et al.*, 2010) and manually refined.

Phylogenetic analyses

The phylogenetic analyses were performed in order to determine the orthologous genes of the *ppc-1a* and *ppc-1b* of *Puya*. Phylogenetic trees were inferred by maximum parsimony (MP), neighbour joining (NJ), maximum likelihood (ML), and Bayesian inference (BI) methods using the nucleotides and amino acid alignment. The MP, NJ and ML trees were constructed using 1000 bootstrap replicates. The MP and NJ trees were constructed in MEGA 7.0.14 (Kumar *et al.*, 2016). The ML analyses were completed using RAxML (Stamatakis, 2006) using the JTT+G (Jones *et al.*, 1992) substitution model

for the protein alignment and GTR+G+I (Tavaré, 1986) substitution model for the DNA alignment. The BI analyses were completed using Mr Bayes 3.2.6 (Ronquist and Huelsenbeck, 2003). Two parallel analyses each composed of twelve chains were run for 20,000,000 generations, sampling a tree every 1000 generations after a burn-in period of 10,000,000 generations (50%). A consensus tree was computed from all trees sampled after the burn-in period for both parallel analyses. Convergence of the analyses and adequacy of the burn-in period were determined using Tracer (Rambaut and Drummond, 2007).

Selection test in genes involved in CAM photosynthesis from eight *Puya* species

Orthologous protein differences are the result of nucleotide sequence changes that modify the codon and therefore the amino acid residue. The rate of change of the nucleotides in the coding sequence are classified as either synonymous (dS) or non-synonymous (dN) substitutions. In synonymous substitutions, the change in the codon does not change the amino acid, while in non-synonymous substitutions the amino acid is altered. Based on the ratio of dN to dS ($\omega = dN/dS$) it is possible to establish under which type of selection the proteins evolved. The proteins were more likely to evolve under neutral evolution when $\omega = 1$. The proteins were more likely to evolve under purifying selection when $\omega < 1$. The proteins were more likely to evolve under positive selection when $\omega > 1$ (Yang, 1998).

The program CODEML from the package PAML v 4.9 (Yang, 2007) was used to test positive selection of genes involved in CAM in *Puya* species. The ML tree for the multiple analyses was obtained with all 4-fold degenerate sites across the entire transcriptomes. That tree was compared with a ML tree for each gene separately. The

gene was excluded from the analysis when the tree based on the transcriptomes disagreed strongly with the ML tree for each gene.

For each gene the ‘sites models’ M7 and M8 (Yang *et al.*, 2000) were compared. M7 split the sites in an alignment into two classes: neutrally evolving sites ($\omega_1 = 1$) or negatively selected sites ($\omega_0 < 1$). The model M8 includes the categories of M7 and additionally includes a class of positive selection sites ($\omega_s > 1$). The two models were compared with the likelihood ratio test (LRT) with 2 degrees of freedom, and corrected resulting *p*-values as in Nevado *et al.* (2016). The Bayes Empirical Bayes (BEB) approach was implemented in PAML to identify the amino acids under positive selection according to their posterior probability of $\omega > 1$ (Yang *et al.*, 2000, 2005).

Selection test of *ppc-1* in monocots

The selection test was evaluated in sequences of *ppc-1* from CAM and C₃ monocots to establish whether PEPC evolved under positive selection in CAM species. An unrooted maximum likelihood tree was constructed using RAxML, using the most expressed *ppc-1* sequences from five CAM species (*Agave tequilana*, *Aloe vera*, *Ananas comosus*, *Phalaenopsis equestris* and *Vanilla planifolia*) and six C₃ species (*Brocchinia reducta*, *Disporopsis pernyi*, *Ludisia discolor*, *Masdevallia yungasensis*, *Phormium tenax* and *Pitcairnia albiflos*). These species represent five independent origins of CAM. The most expressed *ppc-1* isoforms were identified previously in the original transcriptome studies for CAM species (Gross *et al.*, 2013; Cai *et al.*, 2014; Ming *et al.*, 2015). Only CAM isoforms and their orthologues from the C₃ species were used in this analysis. The program CODEML from the package PAML v 4.9 (Yang, 2007) was used to test positive selection in the evolution of *ppc-1* in monocots. A series of models were evaluated from the simplest model M0 to the most complex M8. M0 assumes ω constant across the phylogeny. M1a (nearly neutral) allows ω be variable at different sites and to vary between 0 and 1. M2a allows ω to have values

higher than 1. The model M8 was described above. The performance of the models M1a *vs.* M2a and M8 *vs.* M7 were compared using the LRT. The BEB approach was implemented in PAML to identify the amino acids under positive selection according to their posterior probability of $\omega > 1$ (Yang *et al.*, 2005). To evaluate whether *ppc-1* evolved under positive selection in the CAM branches of the tree, branch-site model tests were conducted. These tests evaluated whether ω showed different values on CAM *ppc-1* (foreground branches). The null model (Model A0) forced $\omega_s < 1$ in the CAM branches, while the alternative model (Model A) allowed $\omega_s > 1$ in the foreground branches. The performance of the model was evaluated by using LTR test.

Results

Expression of genes involved in CAM in *Puya*

The expression of 25 genes involved in CAM photosynthesis (Table 1) were evaluated in eight species of *Puya* (Supplementary Table 2). Malate dehydrogenase was not included because of the multiple copies in *Ananas comosus* and *Puya* spp. For all the identified isoforms (31 in total), it was possible to establish the relative expression of these genes in CAM and C₃ species of *Puya* according to the associated carbon isotope values (Chapter 2). The results of the expression are summarised in Table 2. The most expressed and the lowest expressed isoform of *ppc-1* was labelled *ppc-1a* and *ppc-1b*, respectively. For all species, the gene with highest expression was beta carbonic anhydrase 2, in particular the isoform orthologous to Aco006181 (*Ananas comosus*). The two isoforms of *ppc-1* differed substantially in terms of expression in the majority of *Puya* species (Fig. 1). For example, the isoform *ppc-1a* was expressed 15 times more than the isoform *ppc-1b* in *Puya chilensis*. The species *P. raimondii* showed the lowest ratio of *ppc-1a* to *ppc-1b* expression (Fig. 1).

The differences in the expression of genes involved in CAM listed in Table 2 was evaluated in CAM and C₃ *Puya* species. The genes Aco006181 and Aco017762 were not

considered in the analysis given their high expression. The ANOVA model evaluated the Transcripts Per Million (TPM) differences between: (a) CAM *vs.* C₃ species; (b) all species; (c) all isoforms; (d) all isoforms per species; and (e) all isoform per CAM *vs.* C₃ species. The test showed that there were no significant differences between CAM *vs.* C₃ species in the overall expression of the genes involved in CAM ($p = 0.7458$; Table 3). However, there were significant differences in the overall gene expression between all species ($p = 0.0034$; Table 3). There were significant differences in the expression between the interaction of species and the isoform ($p < 0.0001$; Table 3). The post-hoc analysis showed that the majority of differences were driven by the significant dissimilarities between *P. chilensis* (CAM) and *P. raimondii* (C₃). Significant differences in pairwise comparison were found in *P. ferruginea vs. P. raimondii* ($p = 0.0309$) and *P. spathacea vs. P. raimondii* ($p = 0.0368$) (Table 3). The biological meaning of these results is that CAM and C₃ is a continuum in *Puya* and only constitutive CAM is represented by *P. chilensis* and *P. alpestris*. While constitutive C₃ is more likely to be represented by *P. raimondii*. The rest of species are part of the CAM–C₃ continuum.

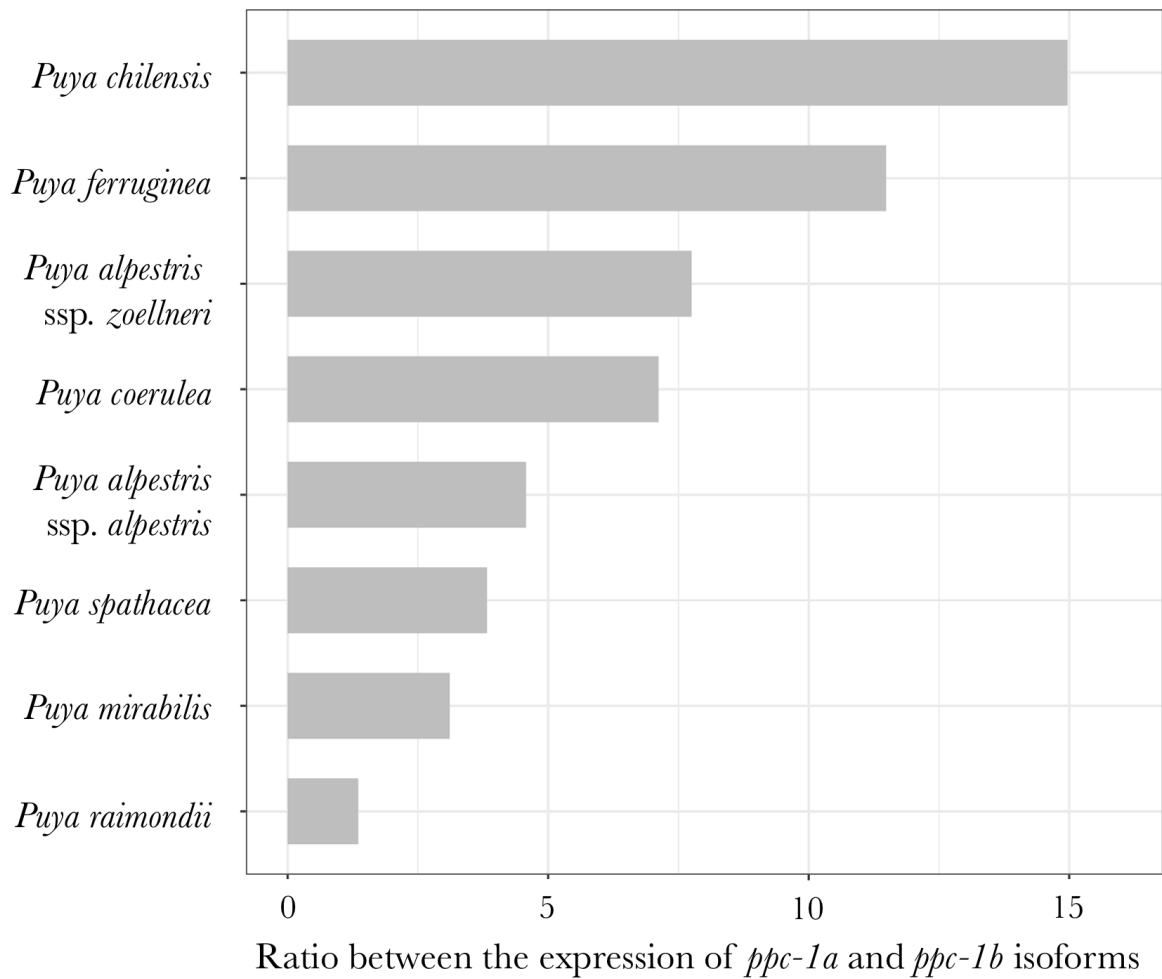


Figure 1. Ratio between the expression (TPM) of *ppc-1a* and *ppc-1b* isoforms in eight species of *Puya*. *P. chilensis* and *P. alpestris* ssp. *zoellneri* are the only species with a carbon isotope value less negative than -20 ‰ and therefore classified as constitutive CAM species.

Table 2. Expression of 23 genes involved in CAM in each species evaluated in the present study. The gene expression is approximated as the relative abundance of transcripts calculated as Transcripts Per Million (TPM). TPM values are comparable between gene and samples because they are already normalised. The photosynthetic pathway is based on carbon isotope values from Chapter 2.

Genes involved in CAM (gene ID <i>Ananas comosus</i>)	Transcripts Per Million							
	<i>P. alpestris</i> ssp. <i>alpestris</i> (C ₃)	<i>P. alpestris</i> ssp. <i>zoellneri</i> (CAM)	<i>P. chilensis</i> (CAM)	<i>P. coerulea</i> (C ₃)	<i>P. ferruginea</i> (C ₃)	<i>P. mirabilis</i> (C ₃)	<i>P. raimondii</i> (C ₃)	<i>P. spathacea</i> (C ₃)
Aluminium-activated malate transporter								
Aco003023.1	12.02	10.99	9.41	27.69	11.55	9.33	4.22	8.18
Aco010725.1	6.27	15.33	3.21	24.02	3.01	5.69	33.76	13.16
Beta carbonic anhydrase 2								
Aco005402.1	55.69	37.61	10.74	62.56	45.59	37.74	136.12	39.55
Aco006181.1	2752.97	3527.44	6555.02	280.03	3654.33	2925.71	2141.41	3017.5
Beta carbonic anhydrase 5								
Aco002732.1	22.09	15.61	27.17	11.16	18.7	8.63	47.19	14.15
NAD-dependent malic enzyme								
Aco016569.1	19.35	21.05	24.71	14.41	8.74	24.67	16.43	15.3
NAD-dependent malic enzyme 2								
Aco007622.1	17.06	17.59	21.18	33.88	14.77	13.72	18.4	32.23
NADP-malic enzyme 1								
Aco009967.1	60.96	39.4	135.24	92.95	279.72	80.77	175.4	44.66

NADP-malic enzyme 4								
Aco005631.1	203.12	234.85	211.45	199.96	226.37	196.97	83	328.45
Phosphoenolpyruvate carboxykinase 1								
Aco017762.1	720.18	56.58	1012.74	4094.85	1100.61	600.27	224.29	749
Phosphoenolpyruvate carboxylase 1 (<i>ppc-1</i>)								
Aco010025.1	283.94	113.68	569.84	497.43	652.69	231.41	50.16	232.29
Aco018093.1	62.06	14.66	38.07	69.88	56.82	74.37	37.05	60.7
Phosphoenolpyruvate carboxylase 2 (<i>ppc-2</i>)								
Aco016429.1	28.17	13.62	37.51	38.92	22.11	49.88	16.91	15.69
Phosphoenolpyruvate carboxylase kinase 1								
Aco010095.1	0	3.1	2.45	12.78	7.31	47.23	8.2	1.7
Aco013938.1	6.24	20.75	32.94	7.19	20.46	42.12	9.33	17.02
Phosphoenolpyruvate carboxylase-related kinase 1								
Aco022525.1	23.86	23.82	11.83	14.17	21.65	16.07	20.72	42.03
Phosphoenolpyruvate carboxylase-related kinase 2								
Aco001261.1	38.2	18.01	19.9	21.36	12.89	17.93	10.19	29.57
PPDK regulatory protein								
Aco014488.1	16.45	10.2	13.76	8.44	12.89	25.39	17.05	9.43

Pyrophosphate:fructose-6-phosphate 1-phosphotransferase								
Aco017133.1	27.64	20.76	15.05	36.76	3.25	18.99	14.79	17.69
Aco018526.1	11.2	61.02	60.74	51.04	58.7	79.42	67.65	70.67
Pyruvate, orthophosphate dikinase								
Aco024818.1	64.53	46.44	58.29	29.58	49.21	124.5	56.63	55.5
Tonoplast decarboxylate transporter								
Aco000795.1	1.73	0	3.1	3.92	0	1.46	1.89	0
V-ATPase A subunit								
Aco008300.1	191.18	210.05	215.87	257	313.23	305.34	204.5	324.98
V-ATPase B subunit								
Aco015226.1	209.89	268.99	238.84	314.24	566.95	240.97	299.46	346.42
V-PPiase (K⁺-dependent)								
Aco008167.1	265.36	265.45	337.69	46.19	87.15	351	101.71	162.52
Vacuolar hexose transporter								
Aco005379.1	32.72	38.98	65.95	55.54	97.05	47.05	43.81	36.71
Vacuolar invertases								
Aco017533.1	2.16	6.21	56.87	237.41	72.05	14.71	18.88	24.97
Aco023030.1	26.51	6.64	10.12	7.08	5.01	29.94	7.76	2.48
Vacuolar sucrose transporter								
Aco011916.1	34.42	13.93	56.87	32.54	11.92	76.72	269.17	46.84
Aco015779.1	19.19	13.23	48.06	20.7	18.41	25.73	8.21	22.32

Table 3. Summary of the ANOVA test of the differences in the gene expression of CAM and C₃ species of *Puya*. Significant differences are shown in bold.

Factors	Degrees of freedom	Mean Sq.	F-value	p-value
Photosynthetic pathway	1	140	0.1050	0.7458
Species	7	4272	3.2180	0.0034
Isoform	28	76070	57.3020	< 0.0001
Isoform * species	163	5402	4.0690	< 0.0001
Photosynthetic pathway * isoform	28	1339	1.0090	0.4621
Residuals	144	1328		

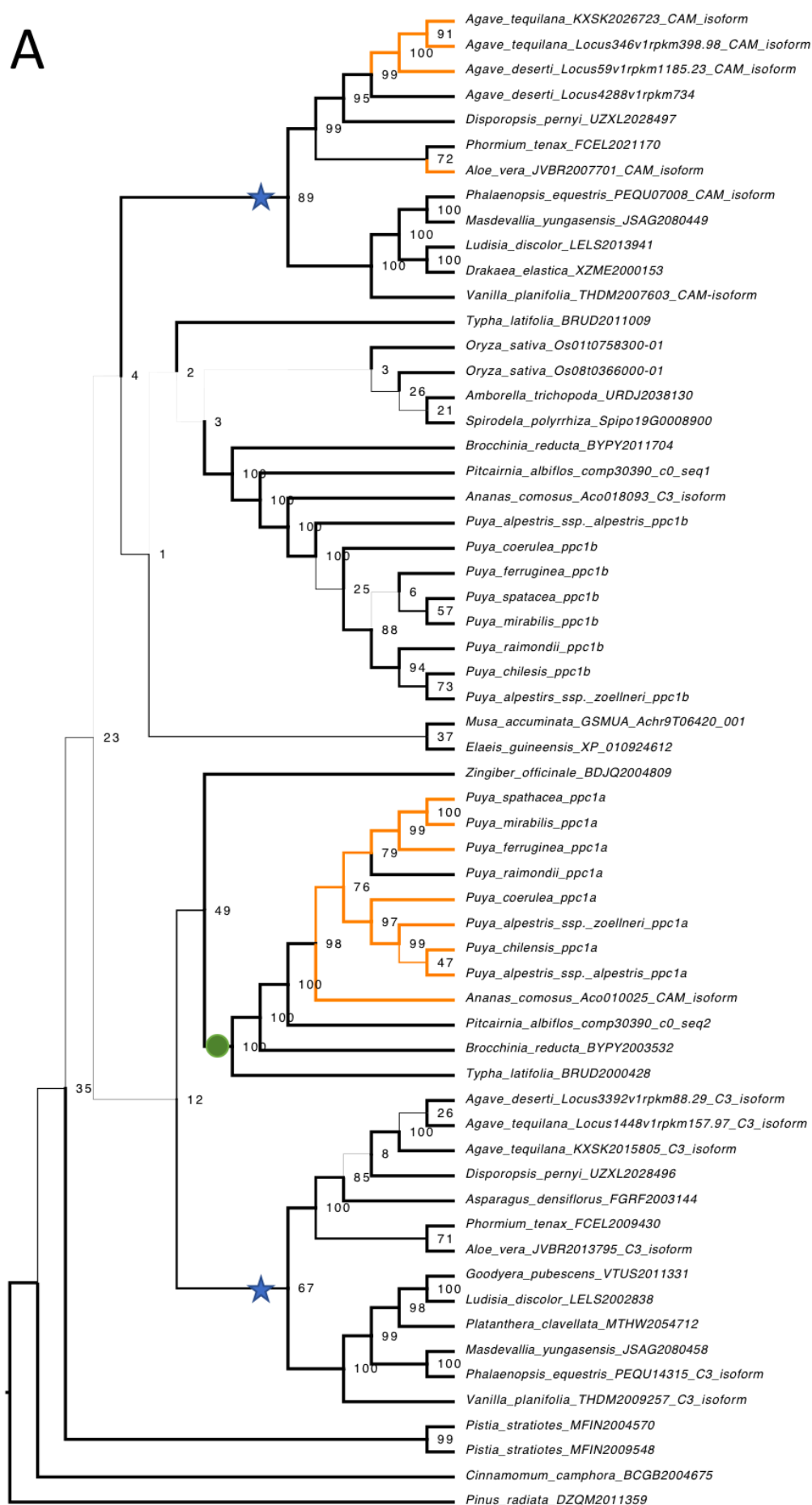
The post-hoc test showed that the only significant differences in the pairwise comparison between species were in *Puya chilensis* vs. *P. raimondii* ($p < 0.001$); *P. ferruginea* vs. *P. raimondii* ($p = 0.0309$); and *P. spathacea* vs. *P. raimondii* ($p = 0.0368$).

* Indicates interaction between factors

Evolution of *ppc-1* in monocots plants

A series of phylogenetic analyses were completed using a representative species of CAM and C₃ monocots to understand the phylogenetic position of the *ppc-1* isoforms in *Puya* (Fig. 2–5). The branches of the phylogenetic trees in Fig. 2–5 have been transformed to illustrate the support of internal nodes. The non-transformed trees are shown in Supplementary Fig. 1–4

A



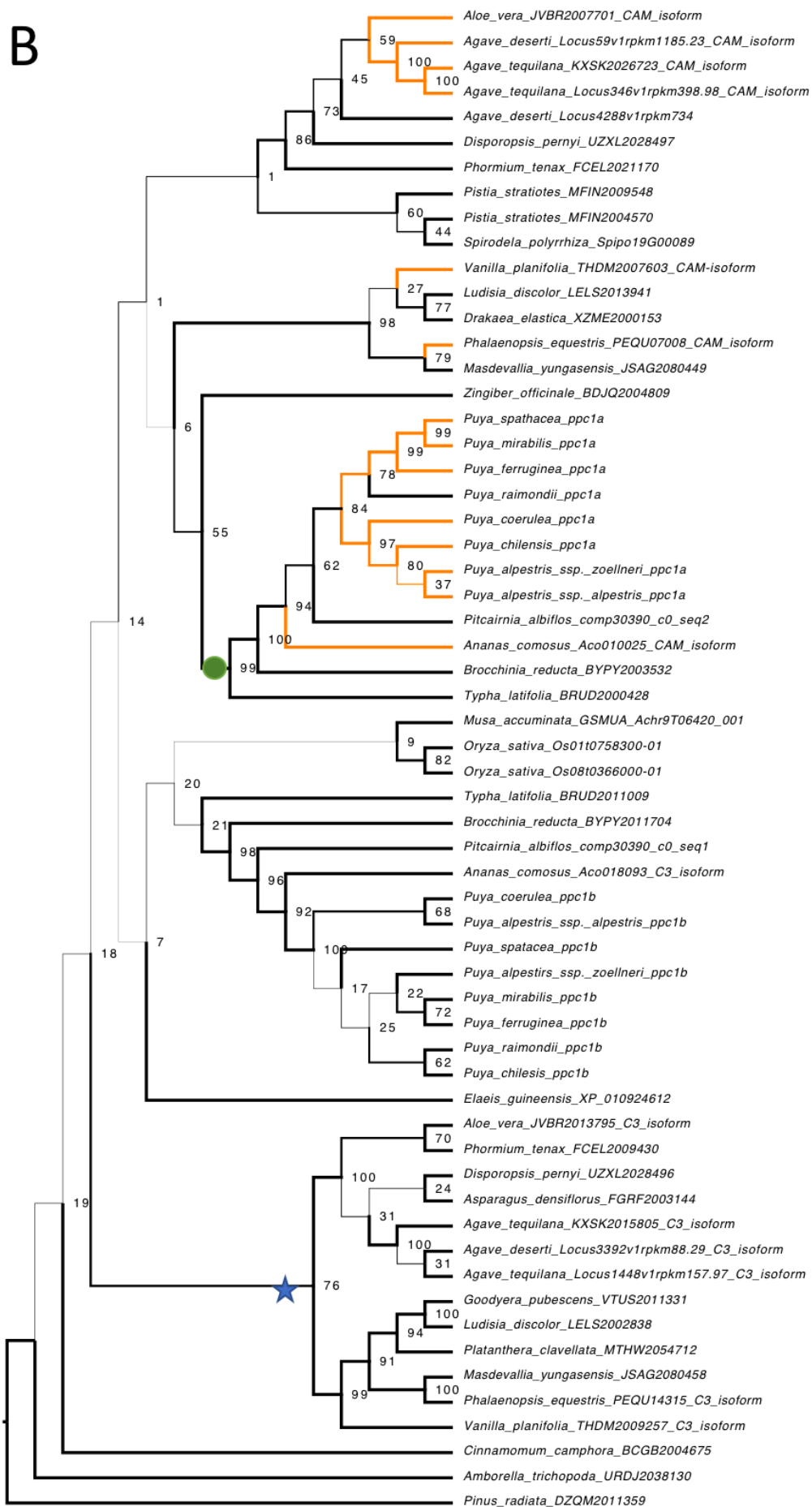
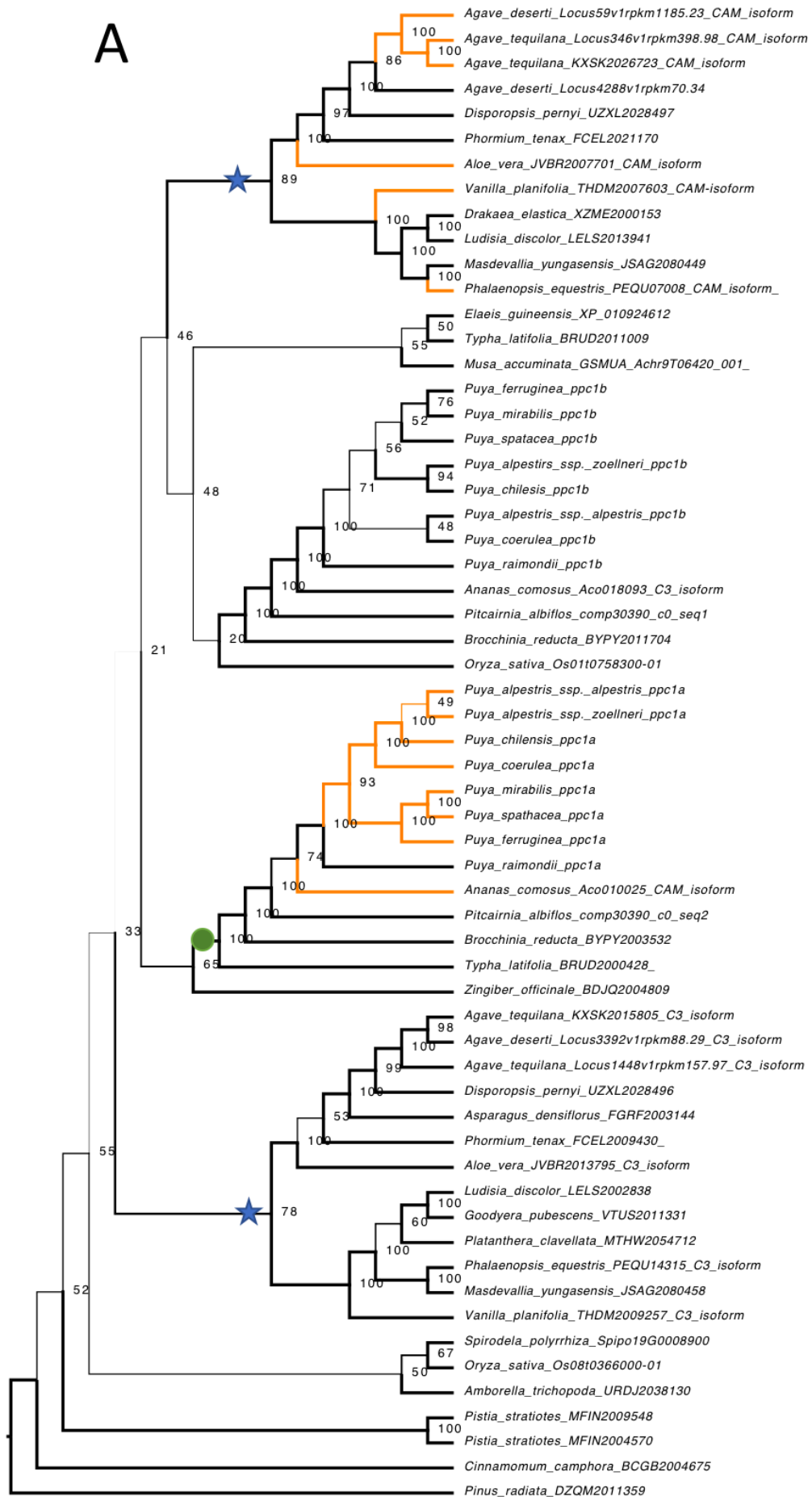


Figure 2. Maximum parsimony (MP) trees showing the evolution of *ppc-1* genes in monocots. Branch support was based on 1000 bootstrap replicates. The branch thickness is directly proportional with the branch support. **(A)** Phylogenetic relationships among *ppc-1* using the nucleotide alignment. **(B)** Phylogenetic relationships among *ppc-1* using the amino acid alignment. Orange branches are taxa showing the residue Arg679 or Asn679 (*A. comosus* CAM isoform numbering) in the PEPC sequence. Black branches are taxa exhibiting the residue His679 (*A. comosus* CAM isoform numbering) in the PEPC sequence. The blue star and the green circle only occur when the monophyly of the order Asparagales and the order Poales were recovered, respectively. Branch lengths were altered to better visualize the branch support. The original branch length is found in Supplementary Fig. 1. The species used in the analysis are described in Supplementary Table 1.



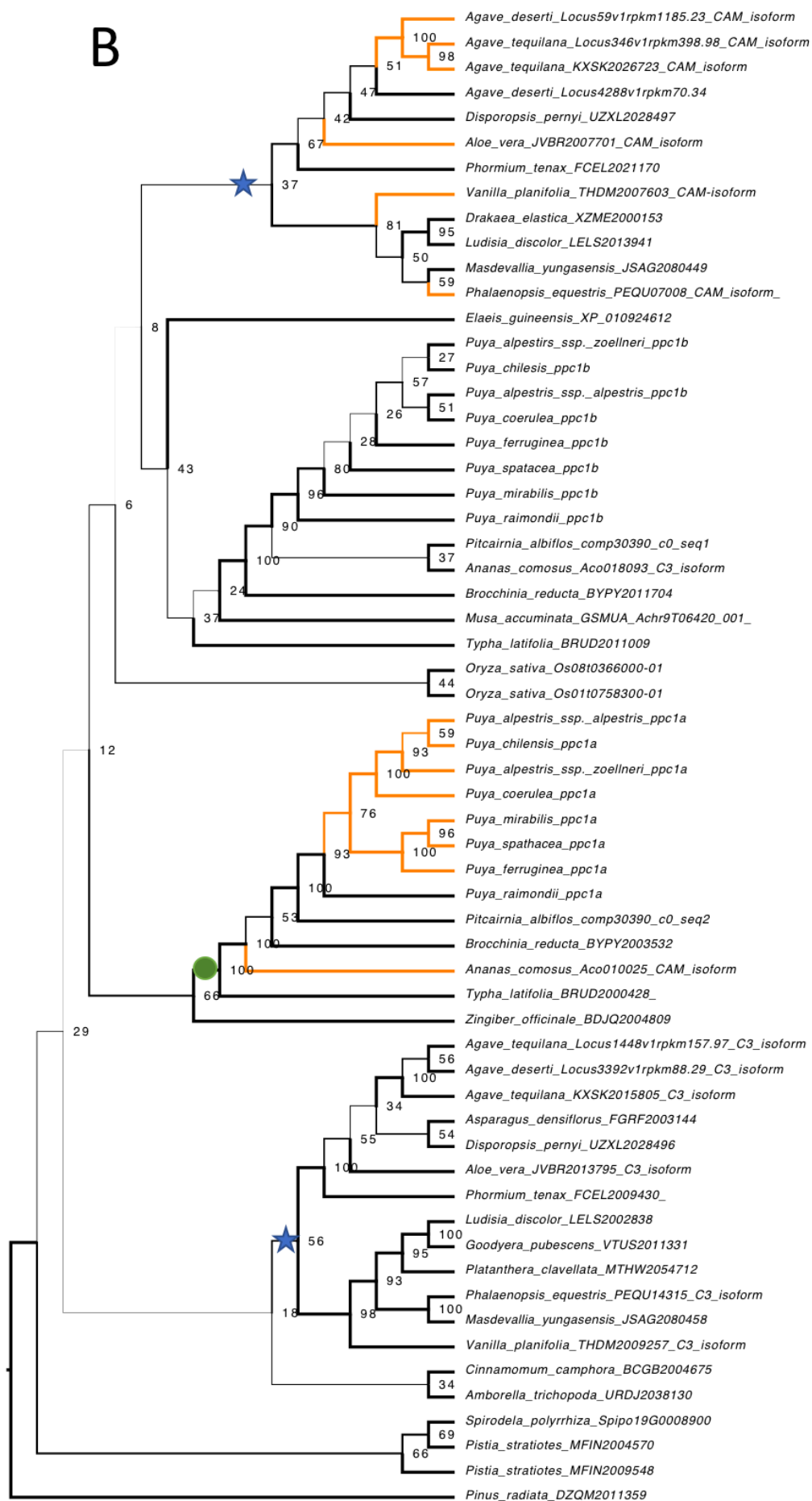
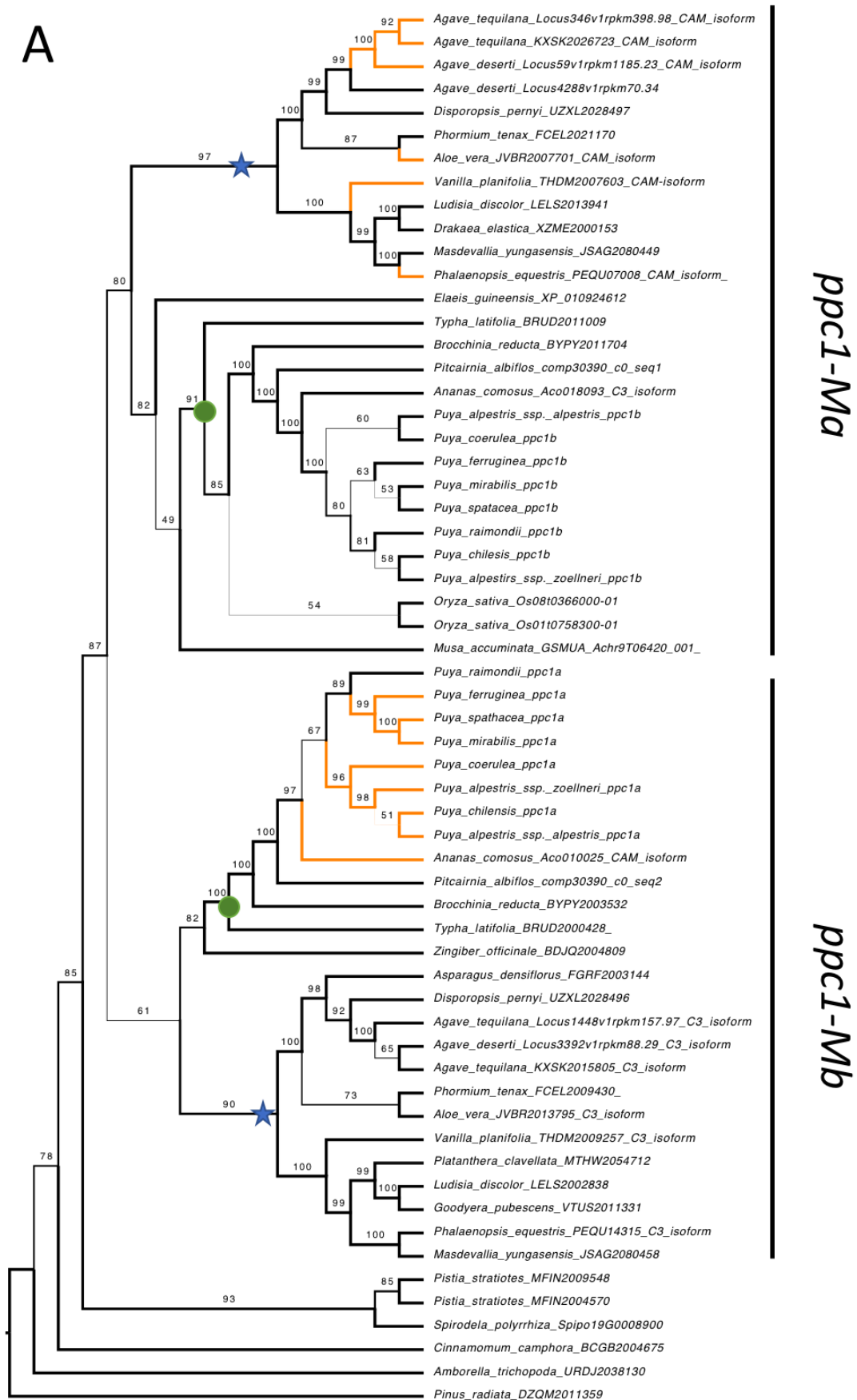


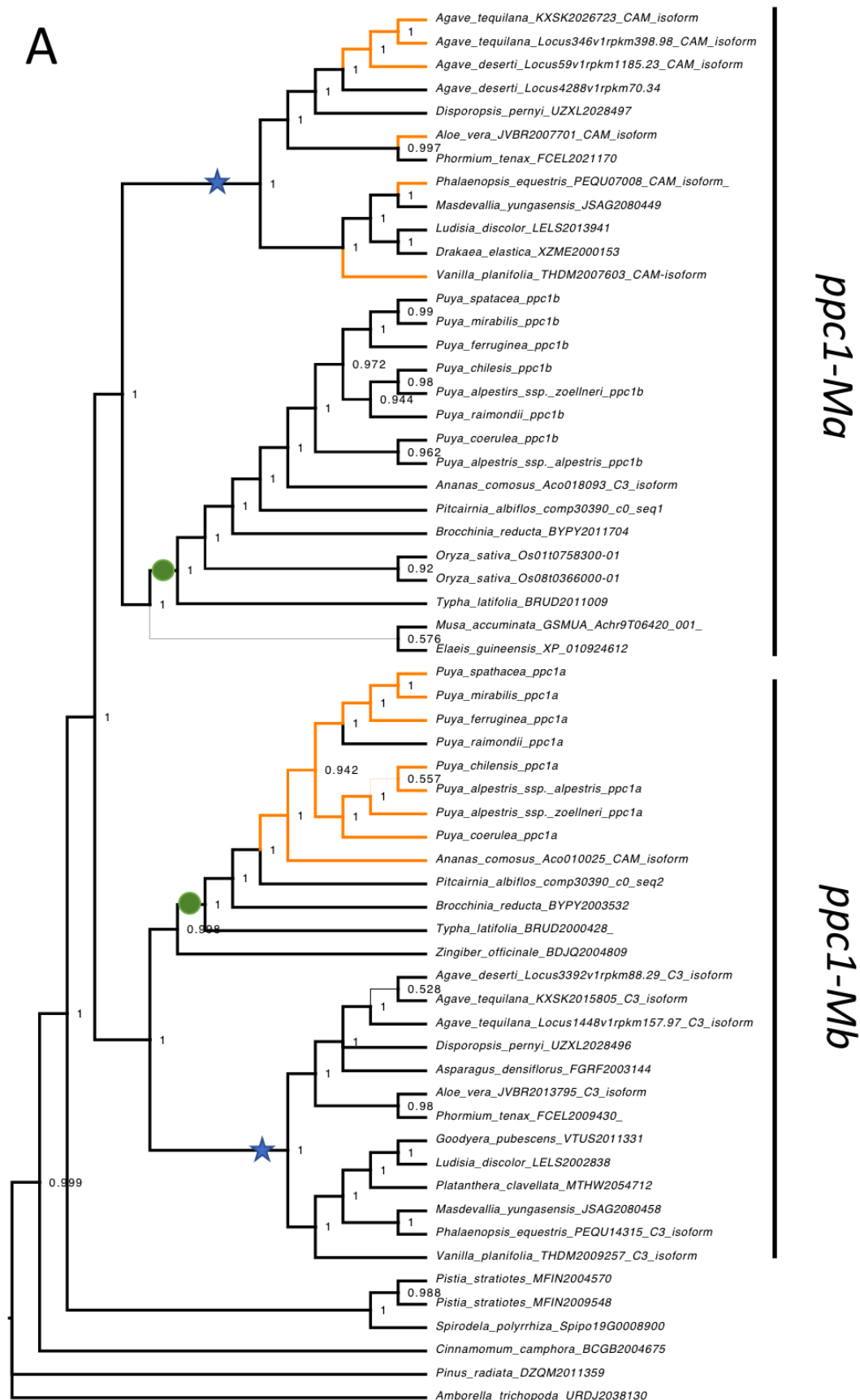
Figure 3. Neighbour-joining (NJ) trees showing the evolution of *ppc-1* genes in monocots. Branch support was based on 1000 bootstrap replicates. The branch thickness is directly proportional with the branch support. **(A)** Phylogenetic relationships among *ppc-1* using the nucleotide alignment. **(B)** Phylogenetic relationships among *ppc-1* using the amino acid alignment. Orange branches are taxa showing the residue Arg679 or Asn679 (*A. comosus* CAM isoform numbering) in the PEPC sequence. Black branches are taxa exhibiting the residue His679 (*A. comosus* CAM isoform numbering) in the PEPC sequence. The blue star and green circle only occur when the monophyly of the order Asparagales and the order Poales were recovered, respectively. Branch lengths were altered to better visualize the branch support. The original branch length is found in Supplementary Fig. 2. The species used in the analysis are described in Supplementary Table 1.



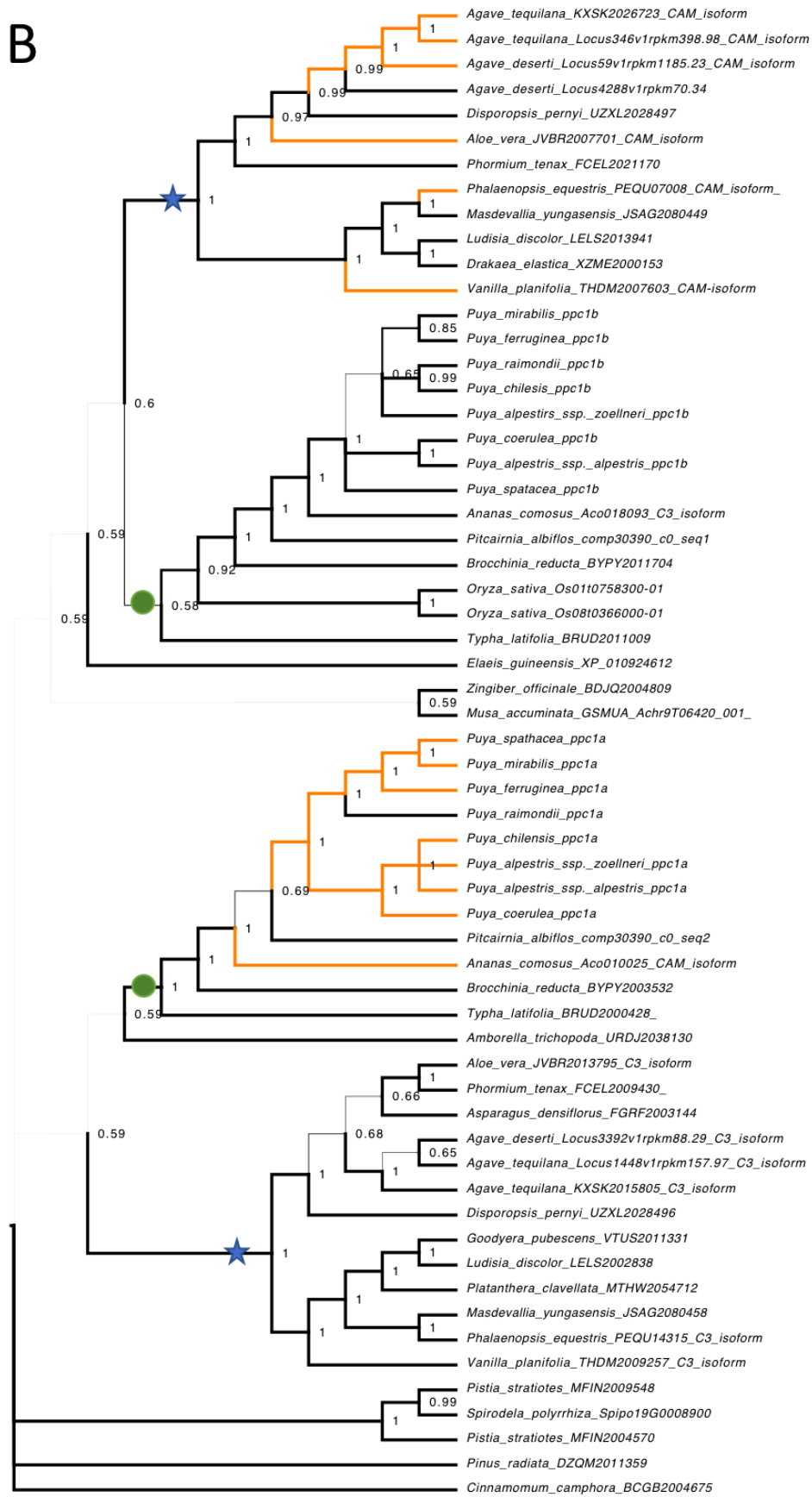
B



Figure 4. Maximum likelihood (ML) trees showing the evolution of *ppc-1* genes in monocots. The clades *ppc-1Ma* and *ppc-1Mb* correspond to *ppc-1M1* and *ppc1-M2* in Christin *et al.* (2014), respectively. Branch support was based on 1000 bootstrap replicates. The branch thickness is directly proportional with the branch support. **(A)** Phylogenetic relationships among *ppc-1* using the nucleotide alignment. **(B)** Phylogenetic relationships among *ppc-1* using the amino acid alignment. Orange branches are taxa showing the residue Arg679 or Asn679 (*A. comosus* CAM isoform numbering) in the PEPC sequence. Black branches are taxa exhibiting the residue His679 (*A. comosus* CAM isoform numbering) in the PEPC sequence. The blue star and the green circle only occur when the monophyly of the order Asparagales and the order Poales were recovered, respectively. Branch lengths were altered to visualize better the branch support. The original branch length is found in Supplementary Fig. 3. The species used in the analysis are described in Supplementary Table 1.



B



ppc1-Ma

ppc1-Mb

Figure 5. Bayesian inference (BI) trees showing the evolution of *ppc-1* genes in monocots. The clades *ppc-1Ma* and *ppc-1Mb* correspond to *ppc-1M1* and *ppc-1M2* in Christin *et al.* (2014), respectively. Branch support was based on posterior probability. The branch thickness is directly proportional with the branch support. **(A)** Phylogenetic relationships among *ppc-1* using the nucleotide alignment. **(B)** Phylogenetic relationships among *ppc-1* using the amino acid alignment. Orange branches are taxa showing the residue Arg679 or Asn679 (*A. comosus* CAM isoform numbering) in the PEPC sequence. Black branches are taxa exhibiting the residue His679 (*A. comosus* CAM isoform numbering) in the PEPC sequence. The blue star and the green circle only occur when the monophyly of the order Asparagales and the order Poales were recovered, respectively. Branch lengths were altered to visualize better the branch support. The original branch length is found in Supplementary Fig. 4 The species used in the analysis are described in Supplementary Table 1.

The phylogenetic analyses using the DNA alignment showed an ancestral duplication of *ppc-1* (clades *ppc-1Ma* and *ppc-1Mb*) in monocots, with high support for the ML and BI reconstructions (Fig. 2A and 3A; Supplementary Fig. 1A and 2A). However, it was less clear in MP and NJ trees (Fig. 4A and 5A; Supplementary Fig. 3A and Fig. 4A) due to the low support of the backbone of the phylogeny. In each clade (*ppc-1Ma* and *ppc-1Mb*) the orders Asparagales (represented by the genera *Agave*, *Aloe*, *Disporopsis*, *Ludisia*, *Masdevallia*, *Phalaenopsis*, *Phormium* and *Vanilla*) and Poales (represented by the genera *Ananas*, *Brocchinia*, *Oryza*, *Pitcairnia*, *Puya* and *Typha*) were monophyletic. According to the ML and BI analyses, the isoforms of PEPC that were highly expressed in CAM species of monocots did not show a monophyletic clade. For example, there were CAM isoforms from the order Asparagales in the clade *ppc-1Ma* (*ppc-1M1* in Deng *et al.*, 2016) and CAM isoforms from the order Poales in the clade *ppc-1-M2* (*ppc-1M2* in Deng *et al.*, 2016) (Fig. 2A and 3A). These results were less clear when using the protein alignment because

independently of the reconstruction method the backbone of the phylogeny was not well supported (Fig. 2B–5B).

Genes involved in CAM evolving under positive selection in *Puya*

The selection test in the different genes involved in CAM in eight species of *Puya* was evaluated to identify positive or neutral selection. Only 15 isoforms were tested due to multiple copies per species and the incongruence in the phylogeny of each isoform with respect to the phylogeny of the transcriptomes. There was a high probability that one third of the genes involved in CAM examined evolved under positive selection (Table 4). This means that the rate of non-synonymous substitutions was higher than the rate of synonymous substitutions in these genes. The genes that presented positive selection were Aluminium-activated malate transporter, NAD-dependent malic enzyme, PPDK regulatory protein, Pyrophosphate-fructose-6-phosphate 1-phosphotransferase, Phosphoenolpyruvate carboxylase-related kinase, and Pyruvate, orthophosphate dikinase (Table 4).

Table 4. Summary of genes involved in CAM evolving under positive selection in the genus *Puya*. The sites listed are those at which positive selection is detected with Bayes Empirical Bayes (BEB) analysis with posterior probability >0.95; those with pp >0.99 are in bold.

Genes involved in CAM (gene ID <i>Ananas comosus</i>)		Gene	Number of codons	Probability of $\omega > 1.0$	Positive selected sites
Aco003023		Aluminium-activated malate transporter	590	0.0011	48 , 83, 147, 342, 406, 519, 578
Aco007622		NAD-dependent malic enzyme	605	0.0003	13, 230, 262 , 318, 433, 485
Aco022525		Phospho eno lpyruvate carboxylase-related kinase	513	0.0097	448, 499
Aco014488		PPDK regulatory protein	405	0.0144	83
Aco018526		Pyrophosphate-fructose-6-phosphate 1-phosphotransferase	616	0.0164	551
Aco024818		Pyruvate, orthophosphate dikinase	844	< 0.0001	47, 98 , 268 , 382 , 387 , 497, 519, 819

Positive selection on amino acid residues of *ppc-1* in CAM species

To establish whether PEPC evolved under positive selection in CAM species of monocots, a set of models was used to evaluate the rate of non-synonymous *vs.* synonymous substitution (ω). The most likely model indicated that there was a different value of ω for each residue (M0 *vs.* M1). However, there were no significant differences between the model of neutral selection (M1a and M7) and the model of positive selection (M2a and M8). The model for positive selection along branches leading to CAM (Model A) had a significantly higher likelihood than the likelihood of the null model (A1; p -value < 0.001 ; Table 5). Therefore, the majority of residues in the sequence of PEPC in monocots (CAM and C₃) evolved under nearly neutral selection (Table 5). However, during the evolution of CAM in monocots eight residues evolved under positive selection (Table 5 and 6). Only two of them (residues 534 and 679, numbered by the CAM isoform of *Ananas comosus*) were found in all CAM species. The residue 679 had the highest posterior probability to have evolved under positive evolution in monocots. This polymorphism in the majority of monocot species did not represent a change in the type of amino acid (Table 6). The functional significance of this polymorphism is currently being tested (Supplementary Material 2)

Table 5. Comparison between amino acid substitution models to identify residues under positive selection in the most expressed *ppc-1* isoform of 11 monocots species (5 CAM species; 6 C₃ species).

Model with positive selection					Null model			LRT	
Model ¹	Log-likelihood	Parameters ²	Positive sites ³	selected	Model	Log-likelihood	Parameters	2L	<i>p</i> -value
Analysis for positive selected sites common for C ₃ and CAM monocot species									
M1a	-14894.12	$\kappa = 1.97,$ $p_0 = 0.91,$ $\omega_0 = 0.08$	None		M0	-15097.70	$\kappa=1.78,$ $\omega=0.08$	407.16	< 0.001
M2a	-14894.12	$\kappa = 1.97,$ $p_0 = 0.91,$ $\omega_0 = 0.05,$ $p_s = 0.00,$ $\omega_s = 24.28$	None		M1a	-14894.12	$\kappa = 1.97,$ $p_0 = 0.91,$ $\omega_0 = 0.08$	0	1.0000
M8	-14790.27	$\kappa = 1.85,$ $p_0 = 0.98,$ $p = 0.25,$ $q = 2.63,$ $\omega_s = 1.00$	None		M7	-14791.61	$\kappa = 1.84,$ $p = 0.22,$ $q = 1.98$	2.69	0.2605

Analysis for positive selection along branches leading to CAM clades

		$\kappa = 1.95,$				$\kappa = 1.78,$		
		$p_0 = 0.86,$				$p = 0.06,$		
		$\omega_0 = 0.04,$				$\omega_0 = 0.13$	48.60	<0.001
		$p_s = 0.08,$	232, 533, 534, 573,					
A	-14849.19	$\omega_s = 1.50$	659, 679, 687, 856	A0	-14873.49			

¹ The models used in the present study are: M1a (nearly neutral), M2a (positive selection), M7 (beta) and M8 (beta & v) for PAML site models; and A0 and A are PAML branch site models.

² The parameters calculated in the model are: κ (transition/transversion rate ratio); ω (dN/dS ratio); ω_s (dN/dS ratio in a class under putative positive selection); p_0 and p_s are the proportion of codons with $\omega_s < 1$ and $\omega_s > 1$, respectively; p and q are parameters of beta distribution in the range (0, 1).

³ The sites listed are those at which positive selection is detected with Bayes Empirical Bayes (Yang *et al.*, 2005; BEB) analysis with posterior probability >0.95; those with pp >0.99 are shown in bold

Table 6. Summary of amino acid replacements under positive selection in the CAM lineages of monocots.

AA no. ¹	AA change (C ₃ to CAM)	Type of change ²	Probability of $\omega > 1$	Note
232	T → A T → Q	P → A P → P	0.992	Polymorphism not present in <i>Agave tequilana</i> and <i>Vanilla planifolia</i>
533	A → S	A → P	0.972	Polymorphism not present in <i>Ananas comosus</i> and <i>Phalaenopsis equestris</i>
534	P → S P → T	A → P A → P	0.986	Polymorphism presents in all CAM species included in the present study
573	A → S A → E	A → P A → D	0.960	Polymorphism not present in <i>Ananas comosus</i> and <i>Phalaenopsis equestris</i>
659	H → N H → L H → E	R → P R → A R → D	0.968	Polymorphism is also present in the C ₃ species <i>Masdevallia yugasensis</i>
679	H → R H → N	R → R R → P	0.999	Polymorphism presents in all CAM species included in the present study
687	F → Y	A → P	0.994	Polymorphism not present in <i>Agave tequilana</i>
856	A → N A → E	A → P A → D	0.965	Polymorphism not present in <i>Ananas comosus</i> , <i>Agave tequilana</i> and <i>Phalaenopsis equestris</i>

In bold are the polymorphisms that are present in all CAM species and are not present in C₃ species.

¹Number of amino acid residues based on the numeration of *Ananas comosus* CAM isoform (Aco010025; Ming *et al.* 2015).

²Amino acid type: A (non-polar aliphatic); P (polar uncharged); D (polar negatively charged); and R (polar positively charged).

The distinctive residue 679 was evident in the alignment of the two copies of *ppc-1* from six constitutive CAM species (*Agave deserti*, *A. tequilana*, *Aloe vera*, *Ananas comosus*, *Phalaenopsis equestris*, *Vanilla planifolia*) (Fig. 6). Moreover, the residue 679 (numbered by *A. comosus* CAM isoform) of the C₃ isoforms contained a histidine residue while all CAM isoforms contained arginine residues (Arg679), except for *Phalaenopsis* which contained an asparagine residue. Closely related C₃ species for each CAM species showed His679 (numbered by *A. comosus* CAM isoform or 680 numbered by *A. comosus* C₃ isoform; Table 5). This is the first report of a polymorphism in the PEPC enzyme between CAM and C₃ species in monocots. The *ppc-1* sequences of the non-CAM species included in the phylogenetic analyses have a His679 in both *ppc1-M1* and *ppc-1Mb*. The *ppc-1* sequence of *Amborella trichopoda*, *Cinnamomum camphora* and *Pinus radiata* exhibited the His67

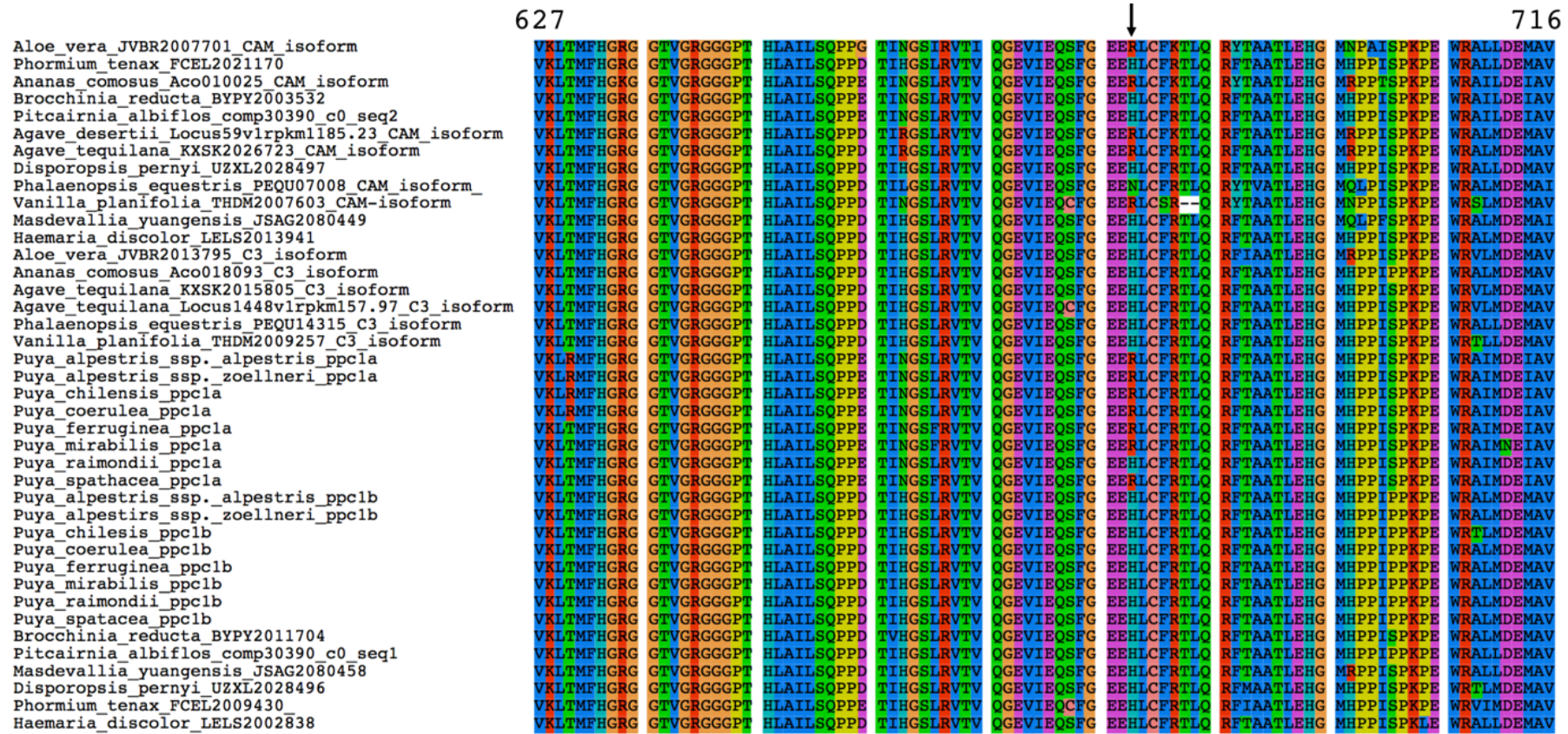


Figure 6. Fragment of the alignment of the *ppc-1* sequences in the order Asparagales and Poales. The black arrow shows the position of the residue Arg679 in CAM species and His679 for C₃ species. The numeration is based on the CAM PEPC isoforms in *Ananas comosus* identified in Ming *et al.* (2015) called Aco010025. More details about the species can be found in Supplementary Table 1.

The two isoforms of *ppc-1* from the eight species of *Puya* were then aligned with the rest of the CAM and C₃ species. Given that CAM evolution of *Puya* has been complex (Chapter 2), the polymorphism could give insights about CAM evolution. All *Puya* species in the less expressed isoform of *ppc-1* (*ppc-1b*) contained His679. In the most expressed isoform (*ppc-1a*) only *P. raimondii* had a His679 residue. The rest of species had an Arg679 residue, even the non-constitutive CAM species (Fig. 6).

Discussion

To improve our understanding of CAM evolution in the genus *Puya*, it is important to consider data from: the relative expression of genes involved in CAM in constitutive and non-constitutive CAM species of *Puya*; the type of selection of genes involved in CAM; evolutionary signatures in the isoforms of *ppc-1*; and the phylogenetic position of the highest expressed isoforms of *ppc* of *Puya* with respect to other CAM and C₃ monocots. The following points will be discussed: (1) the significance of relative expression of genes involved in CAM in non-CAM species in *Puya*; (2) the evolutionary implications of positive selection in genes involved in CAM; (3) the evolutionary implications of the polymorphism Arg679 in *Puya*; and (4) the significance of evolution of CAM in two independent isoforms of *ppc-1* in monocots.

The significance of relative expression of genes involved in CAM in non-CAM species in *Puya*

The relative expression of genes involved in CAM between CAM and C₃ species of *Puya* showed no significant differences (Table 3). In contrast, the relative expression of the genes involved in CAM showed significant difference between species. This result could be explained because the constitutive CAM species and the non-constitutive CAM species represent a continuum rather than two binary categories (Silvera *et al.*, 2010a). Based on

the differential gene expression of the two isoforms of *ppc-1*, the pattern seems to be a continuum rather than two categories (Fig. 1). Indeed, the two species which represented the extremes of the continuum were *Puya chilensis* (CAM) and *P. raimondii* (C₃). It is very likely that the most recent common ancestor of *P. raimondii* and other species from the Northern and Central Andes clade was a C₃ species (Chapter 2). Therefore, it is more likely that *P. raimondii* behaves more similar to a C₃ species than to a constitutive CAM species. In contrast, there are C₃ species (according to the carbon isotope values) that are more similar to CAM species. This is the case for *P. ferruginea* and *P. coerulea* (Fig. 1). These two species belong to clades where their common ancestor photosynthetic pathway is likely to be CAM or undetermined (Chapter 2). Moreover, recently an undergraduate student in the laboratory of Professor's Smith at University of Oxford demonstrated that *P. ferruginea* is able to acidify; but this is not reflected in the carbon isotope value (Chapter 2).

The samples for RNA-seq data in the present study were taken in a single moment of the day. Therefore, it is not possible to establish the circadian rhythm exhibited by some genes involved in CAM (Ming *et al.*, 2015) and also add limitations to the statistical analysis. The expression of *ppc-1* CAM isoform is strong in the green tip of *Ananas comosus* leaves which show CAM photosynthesis, while in the base of the leaves the expression is lower (Ming *et al.*, 2015). These differences are shown within species and between species (Christin *et al.*, 2014; Silvera *et al.*, 2014; Ming *et al.*, 2015). For example, in *Portulaca oleracea* the *ppc-1* isoforms involved in CAM exhibited a circadian rhythm when CAM is activated (Christin *et al.*, 2014). However, the CAM and C₄ isoforms are always more highly transcribed than the housekeeping isoforms (Christin *et al.*, 2014). In CAM plants the activation of PEPC is modulated by the regulation of phosphoenolpyruvate carboxylase kinase (PPCK) (Nimmo, 2000). So, even if the samples were taken in a single moment,

the differences in the *ppc-1a* isoform are correlated with CAM expression. Therefore, given that *Puya raimondii* exhibited the lowest ratio between the expression of *ppc-1a* and *ppc-1b*, this is another indication that *P. raimondii* is perhaps a constitutive C₃ species. In the same way, it is very likely that *P. ferruginea* and *P. coerulea* are weak CAM species (See Fig. 2 in Chapter 5).

The evolutionary implications of positive selection in genes involved in CAM

A third of the genes involved in CAM showed that at least one residue evolved under positive selection. It is important to note that the taxa sample size is limited and therefore to test molecular adaptation of CAM it is necessary to improve the taxon sampling. It is likely that the C₃ species, according to the carbon isotope values, represent points but not extremes in the C₃–CAM continuum. Therefore, finding positive selection only in constitutive CAM species could limit the finding of molecular signatures in the genes involved in CAM. However, the positively selected sites found in Table 4 are not related with CAM and are therefore more likely to reflect speciation, as shown in the Andean *Lupinus* species (Nevado *et al.*, 2016). This pattern could be the result of rapid radiation in the Andes, but not necessarily ecological speciation (Nevado, *et al.*, 2016). *Puya* is an excellent model to study speciation and diversification. In the future, positive selection tests could be applied to high- and low- elevation CAM species to investigate further molecular adaptations to the high elevation environments.

The evolutionary implications of the polymorphism Arg679 in *Puya*

The present study is the first report of a CAM polymorphism in monocots. CAM species contain an Arg679 or Asn679 residue rather than a His679 residue typical of C₃ species. The functional implications of this polymorphism are currently being investigated (Supplementary Material 2). In the most recent divergent clades of orchids, the polymorphism seems to be Asn679 rather than Arg679. For example, *Vanilla planifolia*

belongs to an early divergent clade in Orchidaceae and has the Arg679 residue, while in the species from the tribe Oncidiinae and *Phalaenopsis equestris* the residue is Asn679. This residue is evident in the protein alignment showed in Silvera *et al.* (2014). Species showing ‘weak-CAM’ such as *Rossioglossum insleayi* has the Asn679 residue. In the sequences reported by Silvera *et al.* (2014) the C₃ species have the His679 residue in all their isoforms, except for one isoform in the species *Oncidium sotoanum*. *O. sotoanum* is a constitutive C₃ species. The least expressed *ppc-1* isoform in this species contains the Asn679 residue, (4% in Silvera *et al.*, 2014). Moreover, the metabolic contribution of this isoform is relatively low. Therefore, the most expressed and important isoform of *O. sotoanum* is the isoform that contains His679. This polymorphism has the highest posterior probability according to the BEB analysis. The fact that that polymorphism is present in weak-CAM species (see alignment of Silvera *et al.*, 2014) strongly suggests that this residue could be a possible marker for weak and constitutive CAM plant, something that carbon isotope values could not do.

The taxon sampling of the positive selection test on the monocots PEPC in the present study was not very large (n = 11). However, the taxon sampling represents five independent origins of CAM in monocots. Therefore, the fact that the majority of the *ppc-1* CAM isoforms showed a polymorphism is less likely to be by chance. Moreover, the C₃ isoforms do not contain the polymorphism. The neofunctionalisation of *ppc-1* CAM isoform also occurred in two distinctive clades (Fig. 4–5). Therefore, this is not only the first polymorphism found in the sequence of PEPC; it also shows the convergence in the residue 679 in CAM species of monocots.

It was not possible to determine the photosynthetic pathway of the common ancestor of *Puya* using $\delta^{13}\text{C}$ values and AFLP phylogeny (Chapter 2). However, it was established that CAM was more likely to have evolved in the southern and central Andes

clade than in the northern and central Andes clade (Chapter 2). In the present study, the *ppc-1* sequences of C₃ species of *Puya* representing all the clades were analysed. *Puya raimondii* was the only C₃ species that did not have the residue Arg679. This species belongs to the northern and central Andes clade and its common ancestor was more likely to be a C₃ species (Chapter 2). In contrast, the other C₃ species (perhaps more likely to be part of the C₃–CAM continuum) belong to clades (SC and yellow) where the ancestral condition was more likely to be CAM or undetermined. These species are sister species of constitutive CAM species (Fig. 7 in Chapter 2). These results suggest that the common ancestor of these species was a weak CAM species with an Arg679 residue and that CAM evolved in their sister species when there was a change of climatic niche (Chapter 2). The phylogenetic position of *P. alpestris*, *P. coerulea* and *P. ferruginea*, suggests that the common ancestor of all *Puya* has a *ppc-1* isoform which contains the Arg679 residue rather than His679 (showed in Chapter 5). It is necessary to include *ppc-1* data from basal Bromelioideae species from the genus *Ochagavia* and *Fascicularia* to understand much better the evolution of CAM in the subfamilies Bromelioideae and Puyoideae.

The significance of evolution of CAM in two independent isoforms of *ppc-1* in monocots

It has been established that *ppc-1* had an ancestral duplication early in the diversification of monocots (Deng *et al.*, 2016; Zhang *et al.*, 2016). This is corroborated in the present study (Fig. 4–5). Also, the history of *ppc-1* in monocots has been complex because of multiple gene duplications in grasses and orchids, or gene losses (Christin *et al.*, 2010a; Silvera *et al.*, 2014). The current hypothesis is that all C₄ and CAM isoforms belong to the *ppc-1* and that the CAM isoform evolved only in the *ppc-1Ma* clade (Zhang *et al.*, 2016). The evidence provided in the present study contradicts the last part of this hypothesis. Despite the fact that CAM isoforms were present only in *ppc-1*, it is not true

that all the CAM isoforms in monocots belong to the clade *ppc-1Ma* (Fig. 4–5). According to the data, the isoforms belonging to the *ppc-1Ma* clade evolved CAM in the order Asparagales, while the *ppc-1Mb* evolved CAM in the order Poales. The phylogenetic analysis of PEPC in Zhang *et al.* (2016) showed that one isoform of *Ananas comosus* called AcoPPC-1 was contained in the same clade as the CAM isoforms in orchids. In contrast, the isoform called AcoPPC-2 was contained with the clade of the C₃ isoforms. It is not clear if the AcoPPC-2 is the same as Aco10025 (CAM) or Aco18093 (C₃) isoform identified in Ming *et al.* (2015). The differences between the results of present study and the phylogenetic hypothesis from Zhang *et al.* (2016) could be the result of using the DNA alignments rather than the protein alignment. Natural selection acts directly in the proteins and therefore proteins are more likely to converge (Yang, 2007). Zhang *et al.* (2016) used protein alignments: these are more likely to exhibit convergences. In contrast, the DNA is less likely to exhibit protein convergences. As a result of this convergence, the phylogenies based on protein alignment show very low support in the backbone. This is also showed in phylogenetic trees in Zhang *et al.* (2016) and this could lead to misinterpretation. The nomenclature problem of different isoforms of PEPC has been noticed before (Rosnow *et al.*, 2014). For the future, it is necessary to unify the nomenclature based on the phylogenetic position rather than the expression of each particular isoform. Neutral names such as *ppc-1Ma* and *ppc-1Mb* rather than *ppc-1M1* and *ppc-1M2* are recommended because they avoid confusion about the evolutionary implications. In the future, the evolution of more CAM related genes should be explored to understand in more detailed the evolution of CAM.

Conclusions

The present study demonstrates that the expression of genes involved in CAM in species such *Puya ferruginea* and *P. coerulea* do not differ from constitutive CAM species such

as *P. chilensis* and *P. alpestris* ssp. *zoellneri*. The differences in gene expression suggest that *P. raimondii* is a constitutive C₃ species. The genes involved in CAM that evolved under positive selection is consistent with the recent origin of *Puya* and the speedy diversification across the Andes but not with the evolution of CAM in the genus. This is the first report of a polymorphism in the PEPC sequence that could identify CAM and weak CAM species from C₃ species in monocots. In combination with carbon isotope values, this finding could increase the understanding of the evolution of CAM. The polymorphism is located in the residues 679 (*Ananas comosus* numbering) of the most expressed *ppc-1* isoform. The C₃ isoforms and both isoforms of C₃ species contains a His679 while the CAM-isoform of CAM plants contains either Arg679 or Asn679. C₃ species of *Puya* from the central and southern Andes clade contain the Arg679 residue in the most expressed *ppc-1* gene. This result suggests that in these species the CAM expression was lost recently or that the species exhibit weak CAM. The neofunctionalisation of *ppc-1* in CAM monocots evolved in two different clades. This result contradicts the prevailing view that CAM isoforms evolved only in the *ppc-1Ma* clade.

Chapter 5. General discussion

The present chapter is divided in seven sections. The first section discusses the climatic niche differences between constitutive CAM and non-constitutive CAM species in *Puya*, with an emphasis on the limitations of Chapter 2. The second section discusses the evolution of CAM in the genus *Puya*, based on the results found in Chapters 2 and 4. A new hypothesis about the hybrid origin of CAM in *P. chilensis* is also proposed. The third section discusses the colonisation of the Andes by *Puya*, based on the results from Chapters 2 and 3. A new hypothesis is proposed to explain the evolution of CAM and the cold resistance of *Puya* in the Andes. The fourth section discusses the possible ecological advantages of weak CAM. The fifth section discusses the physiological and biochemical differences between high- and low-elevation CAM species of *Puya* with other studies in the literature. Further avenues for future research are also explored. The sixth section summarises the results of the entire thesis by reference to the main hypotheses of the thesis. Finally, the seventh section presents the main conclusions of this thesis.

Climatic niche differences between constitutive CAM and non-constitutive CAM species in *Puya*

Identifying the climatic niche differences between CAM and C₃ species is not only important to understand the ecology and evolution of plants with complex physiological traits; it is also important to identify areas for cultivation of CAM species that are used for biofuels such as *Ananas comosus*, *Opuntia ficus-indica* and *Agave tequilana* (Owen and Griffiths, 2014; Owen *et al.*, 2016). The genus *Puya* is an outstanding model to study the ecology and evolution of CAM. The climatic variables which drive the differences between CAM and non-constitutive CAM species are aridity index, cloud cover seasonality, mean annual cloud frequency and precipitation seasonality (Fig. 1). These

variables can be useful to differentiate the niche of CAM and C₃ species from other taxa which are different from *Puya*.

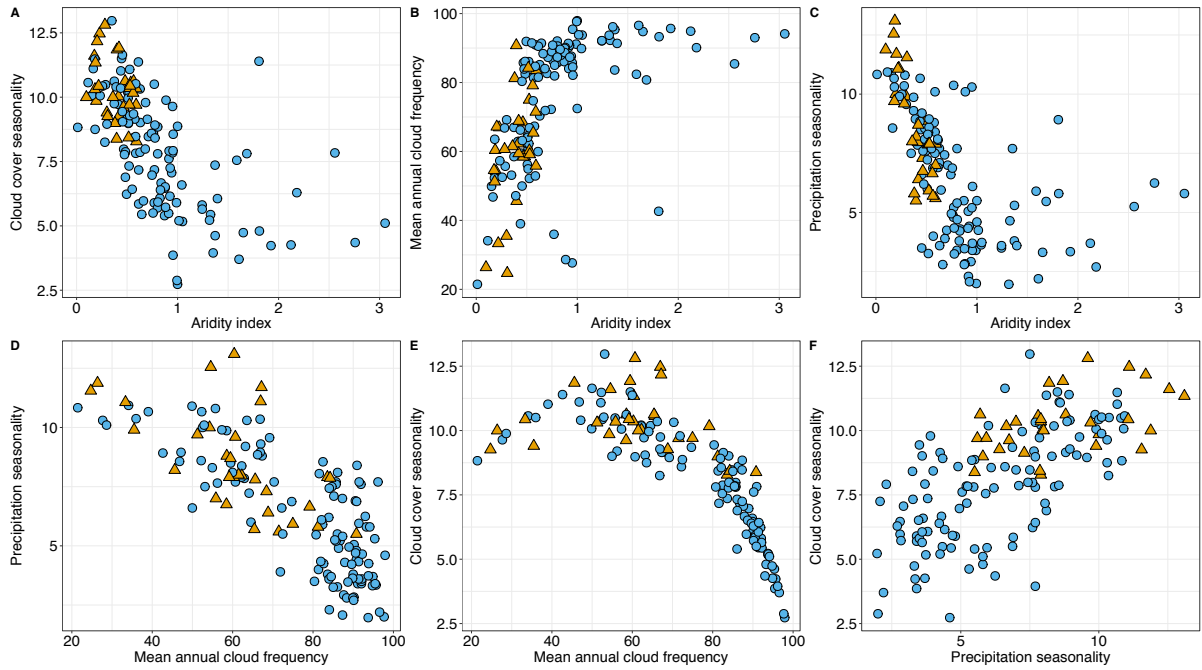


Figure 1. Relationships between the four climatic variables that discriminate more between the 31 constitutive CAM and the 116 non-constitutive CAM species in *Puya* found in Chapter 2. The panels show the relationship between: **(A)** aridity index (mm mm⁻¹) and cloud cover seasonality (SD %); **(B)** aridity index (mm mm⁻¹) and mean annual cloud frequency (%); **(C)** aridity index (mm mm⁻¹) and precipitation seasonality; **(D)** mean annual cloud cover frequency (%) and precipitation seasonality; **(E)** mean annual cloud cover frequency (%) and cloud cover seasonality (SD %); and **(F)** precipitation seasonality (SD %) and cloud cover seasonality (SD %). The orange triangles represent the CAM species. The blue circles represent the non-constitutive CAM species.

Limitations

The present study evaluated the climatic niche of CAM and C₃ species of *Puya* based on herbarium records. The use of natural history collections to evaluate species distribution models is common practice due to the quantity of information available in museums (Franklin, 2010). However, natural history collections can introduce bias into

an analysis (Araújo and Guisan, 2006). For example, collection of the specimens might have followed roads or urbanisation, rather than be a random pattern that could reflect the “real” distribution of the species (Hastie and Fithian, 2013); the collector might have been interested in particular taxa; or the georeferenced information might have been wrong (Araújo and Guisan, 2006, Franklin, 2010). There are three strategies to reduce the bias, and discount these problems. Firstly, use specimens from different collectors and different museums. Secondly, remove poor quality data that adds more noise than signal to a model. Thirdly, use a broad scale (for example, continental scale) in the model. All three approaches were used in the present study to increase confidence in the results and the climatic niche differences between constitutive CAM and non-constitutive CAM species in *Puya*.

The majority of measurements and tests for niche overlap take into account the niche model projected onto a map (Broennimann *et al.* 2012). Such an approach is valuable to test hypotheses about suitability for colonisation by invasive species (Konowalik *et al.*, 2017) or the evaluation of niche convergence (Alvarado-Cardenas *et al.*, 2013). This could be problematic when the model is constructed using presence-only data given the uncertainty of unsuitable areas (Segurado and Araújo, 2004, Segurado *et al.*, 2006; Hastie and Fithian, 2013). Consequently, overlap in the present study was evaluated by comparing the realised niche (bootstrap comparisons) and a proxy using the difference between the model projections of CAM and non-constitutive CAM niche (MaxEnt models). The present study suggests that the high degree of niche overlap could be a consequence of microhabitats occupied differentially by (constitutive and non-constitutive CAM species) that are contained within 1 km².

Future studies could include a distribution model incorporating presence-absence data and relative abundance data of constitutive CAM and non-constitutive CAM species

of *Puya* to evaluate more precisely the overlap between the two niches. Future studies should be restricted in particular to the areas of the central Andes where constitutive and non-constitutive CAM species occur, separated by relatively short distances. The restricted area is fundamental to obtain relative abundance within a reasonable period. The study areas should ideally include dry puna and wet puna environments. A careful annotation of the microenvironments where *Puya* species are reported could help understand whether the overlap described in the present study is the result of microhabitats that are contained within 1 km².

Evolution of CAM in *Puya*

AFLP Phylogeny limitations

The present study constructed a phylogenetic hypothesis based on AFLP data from Jabaily and Sytsma (2013). This phylogeny, whilst similar to those based on chloroplast and nuclear markers, differs particularly in the phylogenetic position of the yellow clade (Jabaily and Sytsma, 2010, 2013). The generic backbone has been maintained and monophyly of the southern and central Andes clade and the northern Andes clade are robust. However, the current phylogeny has relatively low support for the internal branches. This is perhaps a consequence of a very low mutation rate in Bromeliaceae in general (Gaut *et al.*, 1992; Givnish *et al.*, 2007; Smith and Donoghue, 2008). Dating an AFLP phylogeny is impossible because of the lack of an appropriate substitution rate model. Furthermore, outgroups were not included because of the likelihood of non-homologous fragments in *Ananas* and *Deuterocohnia* (Jabaily and Sytsma, 2013). However, branch lengths could be time proxies in comparison with other clades in the same phylogeny, after ultrametric transformation. It is necessary to increase the confidence level of the current phylogenetic hypothesis by including different markers. These markers have to allow the incorporation of outgroups and accurate dating. The

incorporation of more species will allow to estimate more precisely the relationships between CAM and weak CAM species in the central Andes.

Dating the origin of *Puya*

Dating the *Puya* phylogeny is relevant to understand the correlation between the evolution of CAM and geological events. Given the limitations, the ALFP phylogeny was not used to estimate the date of the origin of *Puya*. The most recent phylogeny of the whole family Bromeliaceae (Givnish *et al.*, 2014) was used to estimate the origin of the core *Puya* clade and the central Andes clade. This helped to corroborate whether CAM was associated with the aridification periods described in Chapter 2 and identified by Amidon *et al.* (2017). According to Givnish *et al.* (2014) the crown age of the genus is 9.4 Ma, while the crown age of the core *Puya* is 4.56 Ma. Amidon *et al.* (2017) identified two periods of aridification for the Andes: ~6.1–5.2 Ma and ~3.6–3.3 Ma. These periods correspond with moments before and after the origin of the central Andes and its subsequent diversification. However, it is necessary, to improve the phylogenetic hypothesis using a marker that allows us to date the phylogeny more precisely to test whether the aridity periods suggested by Amidon *et al.*, (2017) triggered the evolution of constitutive CAM in *Puya*.

Evolution of CAM in *Puya*

CAM evolved at least five independent times in the family Bromeliaceae (Givnish *et al.*, 2014). According to Crayn *et al.* (2004), one of these independent origins occurred in the common ancestor of the subfamilies Bromelioideae and Puyoideae, although Givnish *et al.* (2014) suggested that the photosynthetic pathway of the common ancestor of Bromelioideae and Puyoideae was not possible to determine with their data. In contrast, Silvestro *et al.* (2013) suggested that C₃ was the ancestral photosynthetic pathway of the common ancestor of the subfamily Bromelioideae. This finding initially suggests

that CAM evolved independently in *Puya* and the subfamily Bromelioideae. However, Silvestro *et al.* (2013) only considered carbon isotope values as a proxy of CAM and C₃. It is therefore more appropriate to say that the common ancestor of Bromelioideae was a non-constitutive CAM plant.

As demonstrated in this study, it is likely that the common ancestor of *Puya* contained the CAM polymorphism Arg679 in the highest expressed PEPC isoform (Chapter 4). The presence of the CAM polymorphism in non-constitutive CAM species from the early divergent blue clade suggests that the common ancestor of *Puya* was a weak CAM species. The loss of the CAM polymorphism in *P. raimondii* (Chapter 4) is consistent with the ancestral photosynthetic pathway (C₃) of the clade containing *P. raimondii* and the northern Andes clade (Fig. 7 in Chapter 2). Other northern Andes species of *Puya* are constitutive C₃, such as *P. aristeguitae* (pers. comm., Ana Herrera, Universidad Central de Venezuela). McWilliams (1970) also reported constitutive C₃ photosynthesis in an unidentified *Puya* species. Therefore, the fact that the common ancestor of *Puya* was a weak CAM species does not mean that all the species showing a carbon isotope values more negative than -20 ‰ are weak CAM, as mention above during the evolution of *Puya* there were reversals to constitutive C₃ photosynthesis.

Constitutive CAM evolved multiple times in *Puya*, in particular in the southern and central Andes clade, where the common ancestor was likely to contain an Arg679 residue (Fig. 2; Chapter 4). There are two explanations for the multiple gains and possible losses of constitutive CAM in *Puya*. First, weak CAM species have neofunctionalised genes for nocturnal CO₂ fixation (for example in *ppc-1a*, Chapter 4). This means that the transition between weak CAM to constitutive CAM was less complex than from constitutive C₃ to constitutive CAM. Therefore, the neofunctionalisation of CAM genes did not evolve *de novo* in constitutive CAM species. Secondly, the heterogeneous central

Andes could have imposed environmental pressures that triggered the change from weak CAM to constitutive CAM. This could have occurred when the weak CAM population colonised relatively drier areas that have high cloud cover seasonality; in other words, there was a change in the climatic (Chapter 2).

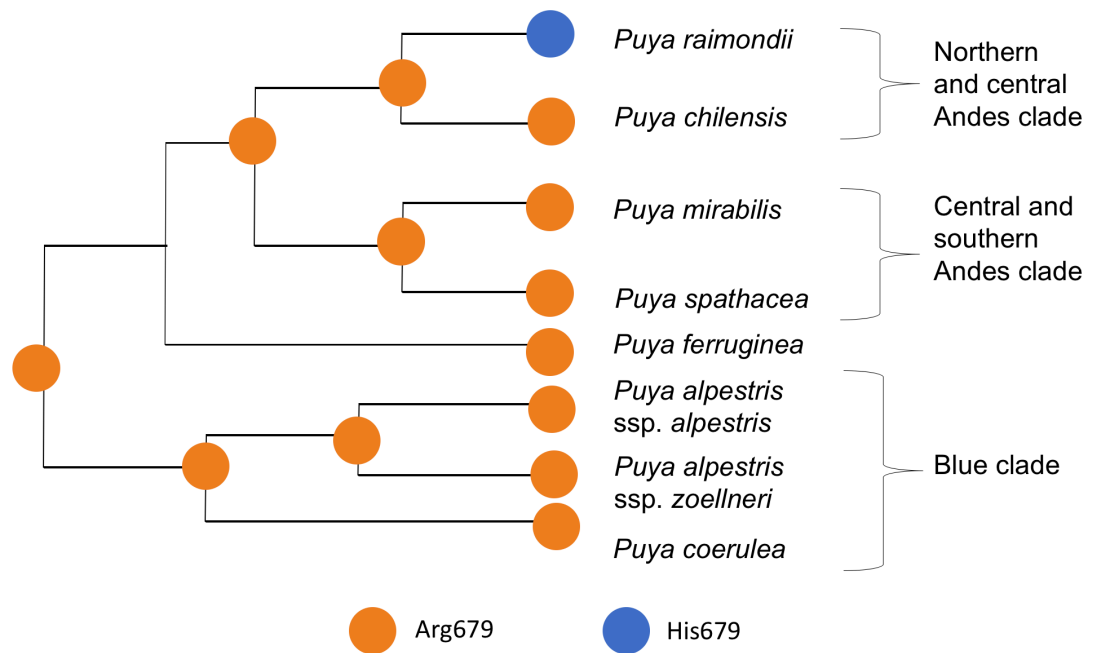


Figure 2. Ancestral state reconstruction by maximum parsimony of the CAM polymorphism in the residue 679 in the highest expressed PEPC isoform of eight species of *Puya*. The phylogeny is a simplified version of Fig. 7 in Chapter 2. The orange circles represent species with Arg679 and the blue circles represent species with His679. According to the reconstruction, the CAM polymorphism was lost in *Puya raimondii*. *P. raimondii* is the sister species of the northern Andes clade which does not contain constitutive CAM species. The limitation of this reconstruction is the limited sampled, more species of *Puya* should be included, particularly from the northern Andes.

Hybrid origin of CAM in *Puya chilensis*

The origin of constitutive CAM in *Puya chilensis* may be a consequence of hybridisation (Chapter 2). Hybridisation between the *P. chilensis* complex and the blue clade is proposed based on disagreements between nuclear and plastid phylogenies

(Jabaily and Sytsma, 2010; Schulte *et al.*, 2010). Jabaily and Sytsma's (2010) plastid phylogeny placed *P. chilensis* and *P. gilmartiniae* in the blue clade (with the *P. alpestris* and *P. coerulea* complexes). This clade was named "Chilean *Puya*". Jabaily and Sytsma (2010) also reconstructed the phylogeny of *Puya* using one nuclear marker (PHYC), where *P. chilensis* and *P. gilmartiniae* formed an independent clade, sister to *P. boliviensis* (yellow clade). The yellow and blue clades were not close relatives. However, *P. alpestris* ssp. *zoellneri* was not included in the nuclear analysis (Jabaily and Sytsma, 2010). In the combined phylogeny, *P. alpestris* ssp. *zoellneri* and *P. chilensis* formed a monophyletic clade.

To explain the discordance, Jabaily and Sytsma (2010) proposed a hypothesis whereby two hybridisation events and one introgression event occurred among the yellow clade, the blue clade and the "core *Puya*" clade. One hybridisation event occurred between the common ancestor of the blue clade and the common ancestor of the yellow clade. In the second event, the hybridisation occurred between *P. alpestris* complex and the common ancestor of *P. chilensis* and *P. gilmartiniae*. A final chloroplast introgression event occurred between *P. boliviensis* and the "core *Puya*" clade. One assumption of the Jabaily and Sytsma (2010) model is that the yellow, the core *Puya* and the blue *Puya* had almost the same age. Another assumption is that the tree clades originated in central Chile. This assumption is not valid according to the most recent and complete phylogeny from Jabaily and Sytsma (2013) as well as the phylogeny from the present study where the yellow clade is contained in the northern and central Andes clade.

There is other evidence that the Jabaily and Sytsma (2010) assumptions are invalid. First, sister species of the yellow clade are from central Peru (*P. angusta*, *P. casmichensis* and *P. casmichensis*, Fig.7 in Chapter 2), suggesting that the geographical origin of the yellow clade is different from the proposed origin in Jabaily and Sytsma (2010). It is likely that the clade's common ancestor originated in the central Andes rather than in

central Chile. Secondly, *P. boliviensis* is the northern-most species of the yellow clade and it is the early divergent species of the yellow clade (Schulte *et al.*, 2009; Jabaily and Sytsma, 2013; Chapter 2). The recently discovered and described species *P. hoxeyi* is likely to belong to the yellow clade given the colour of the flowers and sterile bracts of the inflorescence (Janeba, 2017). *P. hoxeyi* occurs at the Lomas formation in the Atacama Desert in Peru (17°15'S to 17°30'S) and, if it belongs to the yellow clade, it would be that clade's northern-most representative.

The present study therefore proposes that the route of colonisation of Chile by the yellow clade was from northern Peru to the south of central Chile. The yellow clade migrated from central and southern Peru to central Chile using the Lomas formations as stepping stones, and diversifying whilst colonising Chile and northern Peru.

The present study modified the Jabaily and Sytsma (2010) model of hybridisation based on the two different evolutionary gene trajectories of the nuclear gene of PEPC (*ppc-1a* and *ppc1-b*) and the considerations of the new understanding of the geographic origin of yellow clade. The present study proposes that there was only one hybridisation event between the common ancestor of *P. chilensis*–*P. gilmartiniae* and the common ancestor of *P. alpestris* spp. *zoellneri* (Fig. 3) rather than two hybridisation events, as described in Jabaily and Sytsma (2010).

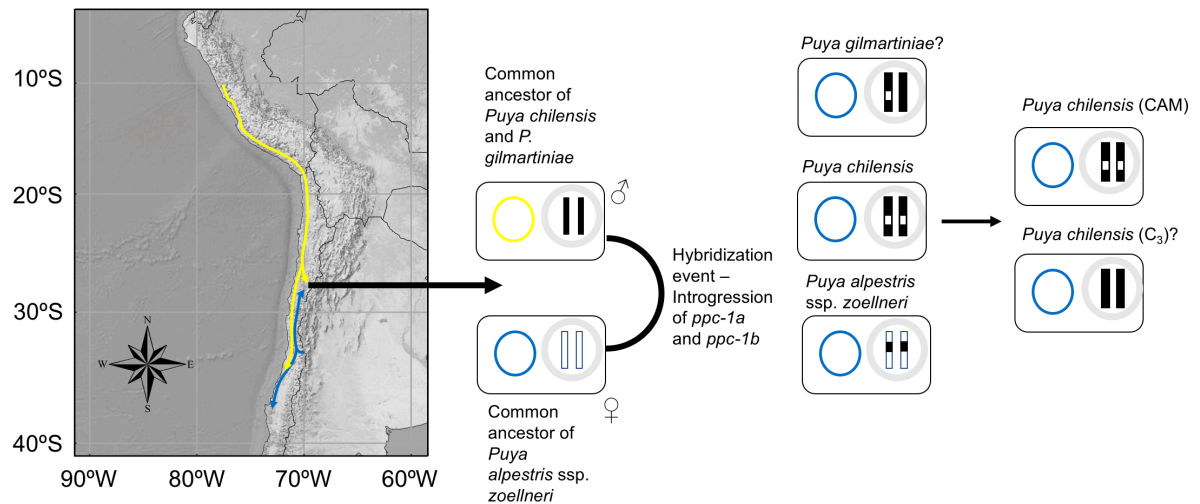


Figure 3. Schematic representation of the origin of CAM in *Puya chilensis* by hybridisation with *P. alpestris* ssp. *zoellneri* and its consequences for the genome evolution. In the left panel, the map represents the most likely biogeographic route by the yellow clade (in yellow) and by *Puya alpestris* (in blue). The colonisation of Chile by the yellow clade was from the north of Peru, while *P. alpestris* had a central Chilean origin. In the right panel, there is a schematic representation of the hybridisation event between the common ancestor of *P. chilensis* and *P. gilmartiniae* and the common ancestor of *P. alpestris* ssp. *zoellneri*. The maternally inherited chloroplast DNA is represented by the small and solid circle; yellow for the common ancestor of *P. chilensis* and *P. gilmartiniae*; and blue for the common ancestor of *P. alpestris* ssp. *zoellneri*. The nuclear DNA is represented by the larger circle containing two bars. After hybridisation, the common ancestor of *P. chilensis* and *P. gilmartiniae* inherited the cpDNA of *P. alpestris* ssp. *zoellneri*. That explains why using chloroplast markers *P. chilensis* and *P. gilmartiniae* are nested with the blue clade. The introgression of nuclear genes are represented by changes in the colour of the bars in the genomic DNA. Adaptive alleles such as *ppc-1a* were transferred from *P. alpestris* ssp. *zoellneri* to *P. chilensis*. Neutral alleles, such as *ppc-1b*, were transferred from *P. chilensis* to *P. alpestris* ssp. *zoellneri* as shown in Fig. 4

Since the yellow clade is contained in the northern and central Andes clade, Jabaily and Sytsma's (2010) introgression step between *P. boliviensis* and "core *Puya*" is no longer necessary. The first hybridisation event proposed by Jabaily and Sytsma (2010) implies an introgression of the chloroplast DNA (cpDNA) from the common ancestor of

the blue clade to the common ancestor of the yellow clade. This step was necessary to explain the “Chilean *Puya*” clade, and the low statistical support of the internal branches of the “Chilean *Puya*” clade. This hybridisation step assumed that the yellow and blue clades have the same geographic origin. As noted above, this assumption is no longer valid under the current phylogenetic hypothesis. The low support of the branches could be a consequence of the slow rate of substitution, which is typically found in the whole Bromeliaceae family (as mentioned above). Therefore, the “Chilean clade” can exist in the cpDNA phylogeny not only if the common ancestor of *P. chilensis* and *P. gilmartiniae* hybridised with the common ancestor of the blue clade, but also if the common ancestor of *P. chilensis* and *P. gilmartiniae* hybridised with the common ancestor of any species in the blue clade. Moreover, the yellow clade diversified relatively recently in comparison with the diversification of the blue clade (Fig. 7 in Chapter 2). As mentioned above, it was not possible to date accurately this tree but it is possible to undertake comparisons within the tree. Therefore, the first hybridisation step suggested by Jabaily and Sytsma (2010) is unlikely.

The second hybridisation step proposed by Jabaily and Sytsma (2010) occurred between the common ancestor of the *P. alpestris* complex and the common ancestor of *P. chilensis* and *P. gilmartiniae*. If this is true, then nuclear genes, such as *ppc-1a* and *ppc-1b*, could show *P. chilensis* in a different clade to *P. alpestris* (both subspecies). According to the phylogenetic hypothesis based on *ppc-1a*, *P. chilensis* is sister to *P. alpestris* ssp. *zoellneri*, and both of them are sisters of *P. alpestris* ssp. *alpestris* (Fig. 4). The clade formed by *P. alpestris* ssp. *zoellneri* and *P. chilensis* is not highly supported. This ultimately would indicate a polytomy between the two subspecies of *P. alpestris* and *P. chilensis* (Fig. 4). However, the phylogenetic hypothesis based on *ppc-1b* showed a different relationship between *P. alpestris* (both subspecies) and *P. chilensis*. First, the phylogenetic relationships of *Puya* using *ppc-1b* are

more similar to the AFLP phylogeny (Fig. 4; Fig. 7 in Chapter 2), showing that *P. chilensis* is closely related to *P. raimondii*. Secondly, *P. alpestris* ssp. *alpestris* and *P. alpestris* ssp. *zoellneri* do not form a monophyletic clade. This result suggests that the second hybridisation step in Jabaily and Sytsma (2010) was only between *P. alpestris* ssp. *zoellneri* and *P. chilensis*. The hybridisation step that is proposed in the present study is between the common ancestor of *P. chilensis*–*P. gilmartinae* and the common ancestor of *P. alpestris* ssp. *zoellneri* (Fig. 3). This hybridisation step includes cpDNA introgression from *P. alpestris* ssp. *zoellneri* to the common ancestor of *P. chilensis*, and the introgression of nuclear genes from *P. alpestris* ssp. *zoellneri* to the common ancestor of *P. chilensis*, and vice versa (Fig. 4).

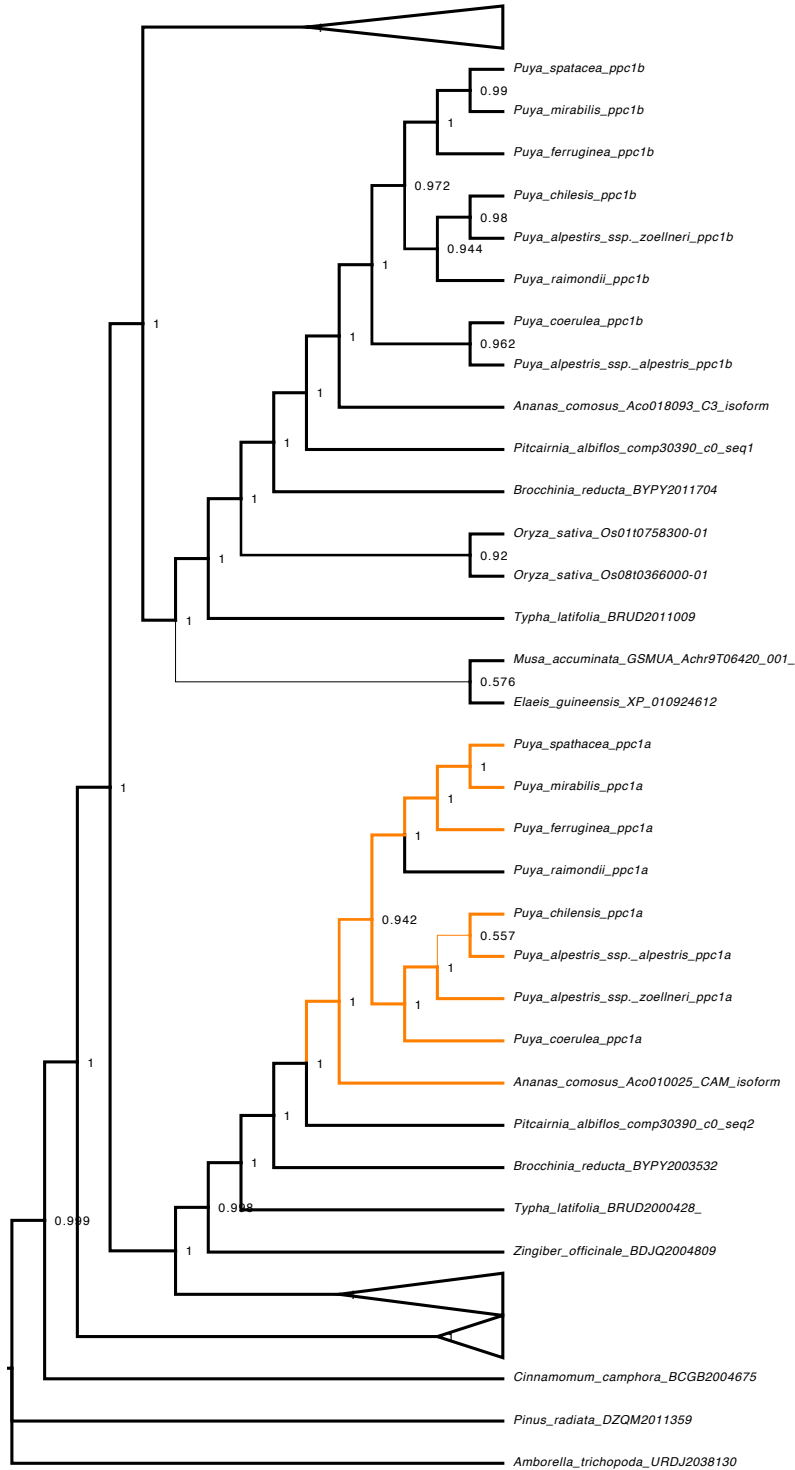


Figure 4. Bayesian inference (BI) trees showing the evolution of *ppc-1* genes in *Puya* using nucleotide alignment. Clades were collapsed to visualise the different topologies of *ppc-1a* and *ppc-1b*. Branch support was based on posterior probability; the branch thickness is directly proportional with the branch support. The phylogenetic relationships between the two isoforms are different and reflect the hybridisation event described in Fig. 3.

Early divergent species of the yellow clade (*Puya boliviensis* and *P. gilmartiniae*), and the new species *P. hoxeiyi*, occur in the most mesic part of the Lomas formation in the Atacama Desert (Rundel and Dillon, 1998, Janeba, 2017). The Atacama Desert is more arid than the central part of Chile. However, given the almost constant water supply by fog, it is likely that during the colonisation of Chile by the common ancestor of *P. chilensis*–*P. gilmartiniae*, the populations experienced a transition between a “mesic” environment (the most mesic part of Lomas formation) to a xeric environment in the north of central Chile. In this region, there is evidence that *P. alpestris* ssp. *zoellneri* populations show constitutive CAM. The *ppc-1* isoform of *P. chilensis* could be gained by introgressive hybridisation from *P. alpestris* ssp. *zoellneri*, perhaps among other CAM genes that could give populations of *P. chilensis* constitutive CAM photosynthesis. The hybridisation also modifies *P. alpestris* ssp. *alpestris* because the less expressed *ppc-1b* was transferred by introgressive hybridisation from *P. chilensis* to *P. alpestris* ssp. *alpestris*. The present study suggests that the hybridisation between *P. alpestris* ssp. *zoellneri* and *P. chilensis* took place in northern central Chile (Fig. 3).

Introgressive hybridisation allows species to incorporate genes that may have an adaptive value (Arnol and Kunte, 2017). Indeed, introgressive hybridisation could be differential, transferring “adaptive” alleles, or be neutral with neutral alleles (Harrison and Larson, 2014). This could explain why the *ppc-1a* gene from *P. alpestris* ssp. *zoellneri* was transferred to *P. chilensis*, given its relevance in CAM, and also explains why the less expressed *ppc-1b*, perhaps a neutral allele, was transferred from *P. chilensis* to *P. alpestris* ssp. *zoellneri*. Introgressive hybridisation has been documented in several organisms, including, birds, plants and humans (Arnol and Kunte, 2017). One example is the evolution of petal colour in the *Mimulus aurantiacus* species complex (Stankowski and Streisfeld, 2015). Based on genome-wide variation across the species complex, Stankowski

and Streisfeld (2015) found “two” independent origins of the red colour of the petals in the species complex in *M. aurantiacus* ssp. *puniceus* and *M. aurantiacus* ssp. *flemingii*. The red colour of the petal is due to a *cis*-regulatory mutation of the gene MaMyb2. However, by studying the gene flow within the species complex, they found that the *cis*-regulatory mutation of the gene MaMyb2 evolved *de novo* in the common ancestor of *M. aurantiacus* ssp. *flemingii* and was transferred by introgressive hybridisation to *M. aurantiacus* ssp. *puniceus*.

Naturally-occurring hybrid species from parental CAM and C₃ species are not common, perhaps due to distinctive climatic niches. However, there is a natural hybrid species in the genus *Yucca*, *Y. gloriosa*, whose parental species are constitutive CAM (*Y. aloifolia*) and constitutive C₃ (*Y. filamentosa*) (Rentsch and Leebens-Mack, 2012; Heyduk *et al.*, 2016b). *Y. gloriosa* is a weak CAM species (−22.0‰, Martin *et al.* 1982) which can fix CO₂ at night time. *Y. gloriosa* exhibits intermediate characteristics in comparison with its parental species, and during drought stress 100% of the CO₂ uptake occurs at night time (Heyduk *et al.*, 2016b). It is possible that the common ancestor of *P. chilensis* after the hybridisation step would be similar to *Y. gloriosa* and show a weak CAM photosynthesis.

Yucca gloriosa is a stable hybrid species and therefore it is a useful model to investigate CAM evolution but not to investigate the initial steps of hybridization between a CAM and a C₃ species. Unfortunately, there is no such model in CAM plants. However, hybrids of C₄ and C₃ plants have been developed in laboratory conditions. This is the case for artificial hybrids of *Atriplex prostrata* (C₃) and *A. rosea* (C₄). After hybridisation, the great majority of F₁ and F₂ exhibited physiological characteristics more similar to the parental C₃ species than the parental C₄ species (Oakley *et al.*, 2014). The majority of F₁ and F₂ individuals exhibited an intermediate anatomy. Only 1 in 23 successful F₂ lines showed similar physiological characteristics with their parental C₄ species. In a

hypothetical scenario where the successful F₂ hybrids naturally occur and the conditions favour only C₄ plants, only a small fraction of the F₂ hybrids will have a higher fitness and will therefore survive and reproduce. After generations, if the hybrid backcross with the paternal C₃ species, C₄ isoforms could be fixed into the population gene pool. In the case of *P. chilensis* and *P. alpestris* ssp. *zoellneri*, it is possible that the early hybrids that exhibited CAM characteristics were favoured by the environmental conditions in the northern part of central Chile. It is likely that the introgressed genes were finally incorporated into the gene pool by being crossed back to the parental species as suggested by the *ppc-1* phylogeny (Fig. 4).

Two additional points must be considered: first, whether the common ancestor of *P. chilensis* had the Arg679 residue; and secondly, whether *P. chilensis* lost its CAM capacity when it colonised the more southern areas of Chile. Despite the fact that early divergent species of the yellow clade occurred in mesic areas, the phylogenetic position of the clade indicates that the common ancestor was more likely to have the Arg679 residue in the PEPC isoform (Fig. 2). However, acquisition of Arg679 in *P. chilensis* is likely to have been through hybridisation. Future studies should test whether the CAM polymorphism is present in *P. boliviensis*, *P. gilmartiniae* and *P. hoxeyi* and if the CAM polymorphism is present species from the northern Andes clade.

The second issue is more problematic because the present study only analysed the PEPC sequences of *P. chilensis* CAM populations, and not C₃ populations. When the common ancestor of *P. chilensis* (C₃ populations) was migrating further south in central Chile to more humid environments, it is possible that the regulatory genes that trigger CAM were under relatively less pressure than in the northern part of central Chile. This resulted in a change between constitutive CAM niche to non-constitutive CAM niche (Chapter 2, Supplementary Fig. 9). As a result, constitutive CAM was lost. Future studies

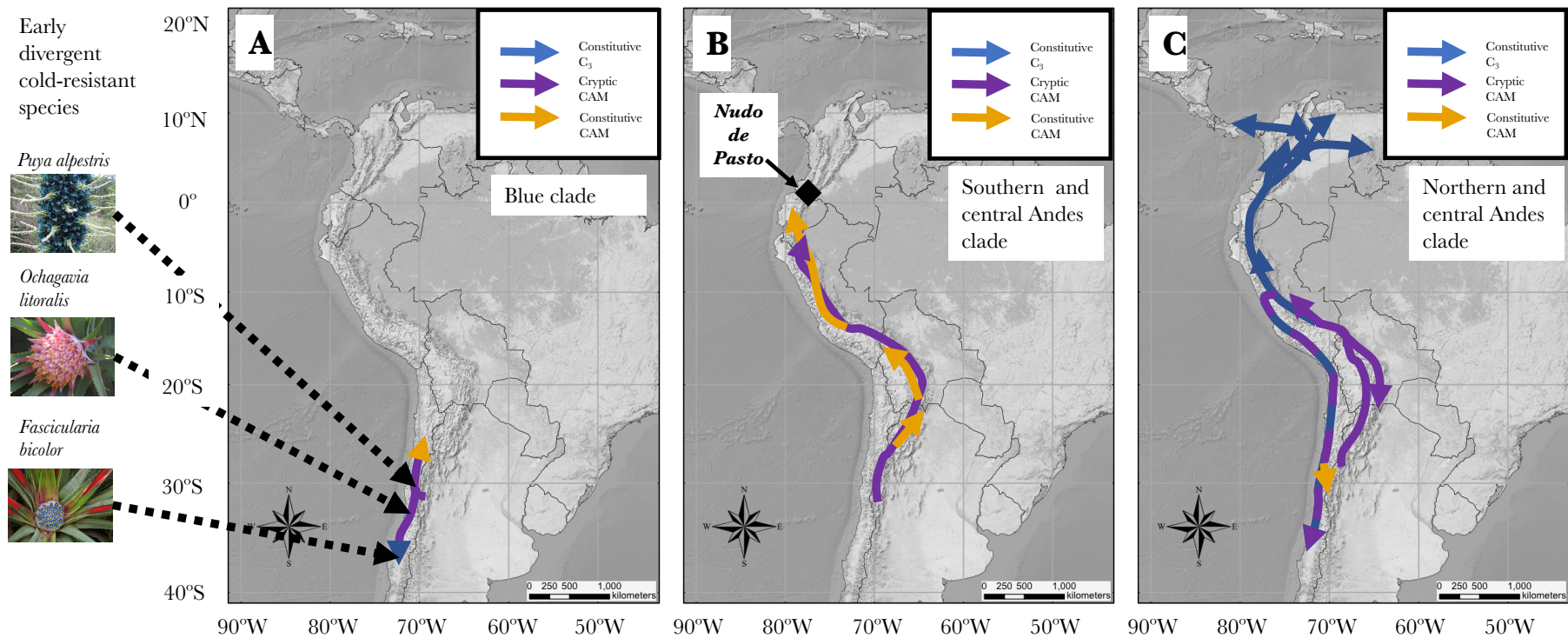
should test this hypothesis by studying (in the field and in common garden conditions) the physiology of *P. chilensis* (C₃ and CAM populations) and *P. alpestris* (both subspecies).

The colonisation of the Andes by the genus *Puya*

The rise of the Andes has increased the diversification of species in multiple ways including: allopatric speciation, especially in montane taxa; generation of new environments and therefore increased ecological speciation; and creation of biotic corridors for taxa exapted to montane conditions increasing their range (Antonelli and Sanmartin, 2011).

All evaluated *Puya* species have a genetic component to cold resistance (Fig. 5; Chapter 3). The present study evaluated *P. alpestris*, an early divergent species of *Puya* and also an early divergent species of the subfamily Bromelioideae. Given the phylogenetic position of these species, it is likely that they reflect the state of the common ancestor of Bromelioideae and Puyoideae. The crown age of the core *Puya* clade is 4.56 Ma while the stem age is 9.09 Ma (based on Givnish *et al.*, 2014). The stem age corresponds to the final rise of the central Andes, while the crown age corresponds to the point where the central Andes plateau was stable (Garzzone *et al.*, 2008). The present study suggests that the rise of the central Andes *per se* did not increase the diversification of core *Puya*. However, once the central Andes was stable, diversification increased (Fig. 5D–E). This may be a result of Pleistocene glaciation-interglaciation cycles that expanded and contracted vertically the distribution of puna and the paramo. The expansions formed corridors (increasing the range of *Puya* species) and the contractions formed isolated paramos and punas influencing allopatric speciation (Särkinen *et al.*, 2011, 2012, 2014). *Puya* colonised the Andes from the south to the north, and the colonisation of dry environments that have a high cloud cover seasonality triggered constitutive CAM. The fact that the rise of central Andes did not increase diversification in *Puya* is consistent with the apparent exaptation

of cold resistance, as showed by the common ancestor of *Puya*. In these new high-elevation environments, cold resistant species had an advantage (Fig. 5A–C, F).



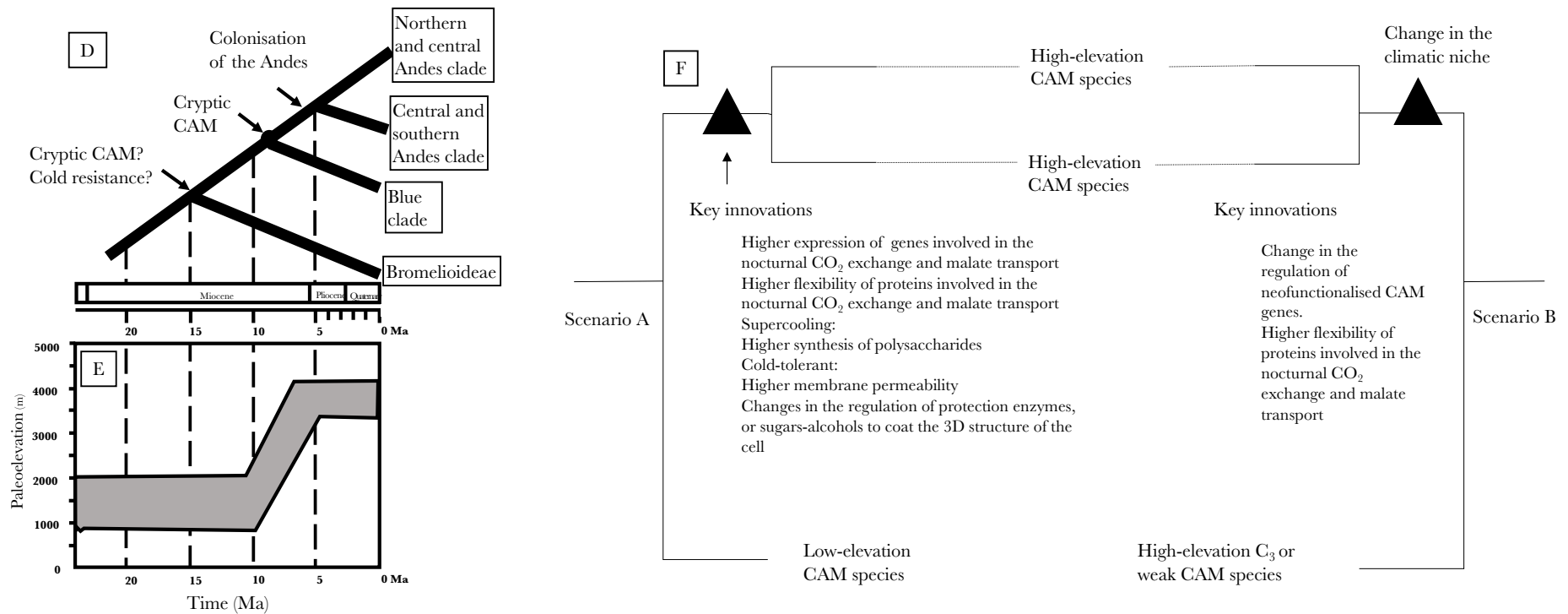


Figure 5. Synthesis of the thesis. Panels (A–C) represent the most likely biogeographic paths correlated with constitutive CAM evolution in *Puya*. Panel (A) represents the blue clade. Panel (B) represents the central and southern Andes clade. Panel (C) represents the northern and central Andes clade. The arrows indicate the most likely direction of colonisation by *Puya*. The distribution of early divergent cold resistant species is shown in Panel (A). This circumstantial evidence suggests that the common ancestor of Puyoideae and Bromelioideae was a cold-resistant plant. In Panel (B), the biogeographical barrier *Nudo de Pasto* is highlighted. In Panel (C), the path of colonisation of Chile by the yellow clade is shown in purple and blue colours because the Arg679 in *P. Chilensis* has a hybrid origin (Fig. 3). The current data suggests that this barrier does not allow weak or

constitutive CAM species to pass through the barrier. The evolution of constitutive CAM is represented in orange when there was a change in the climatic niche. The reversal to constitutive C₃ is represented in blue and weak CAM is represented in purple. Panel **(D)** is a schematic representation of the phylogeny of *Puya* with the evolution of weak CAM and the colonisation of the Andes. Panel (D) is calibrated based on Panel **(E)**, which shows the paleoelevation of central Andes during the last 20 Ma based on Fig. 4 of Garzzone *et al.* (2008). Panel **(F)** presents the schematic hypothesis of evolution of CAM in high elevation environments considering two possible scenarios. Scenario A: the common ancestor was a low-elevation CAM species. The key innovations to colonise high elevations are related with an increase in the cold resistance and the flexibility of CAM-related enzymes, which are active at night time. Scenario B: the common ancestor was a high-elevation C₃ species or a weak CAM species. In this case, the evolution of CAM is correlated with changes in the climatic niche. The key innovations to colonise the high-elevations are related to neofunctionalisation of CAM genes and their regulation. Also an increase in the flexibility of the enzymes active at night times such as PEPC. These hypothetical scenarios can be tested in taxa containing CAM species at high elevation, by studying the transcriptome and the ecophysiology.

The northern Andes was colonised by the northern Andes clade around ~ 1.58 Ma (Madriñan *et al.*, 2013). As discussed in Chapter 2, it is not yet clear whether the northern Andes was colonised twice; this cannot be determined until the phylogenetic position of *Puya floccosa* has been studied. Chapter 2 also demonstrated that the low dispersal of *Puya* and the extremely humid region of the *Nudo de Pasto* would have acted as a biogeographic barrier that prevented the dispersal of constitutive CAM species but allowed the dispersal of C₃ species. The northern-most distributed constitutive CAM species is *P. lananta*. This species belongs to the central and southern Andes clade.

The evidence provided in Chapter 4 indicates that the most likely common ancestor of *Puya* was a weak CAM plant. It is not clear whether the CAM polymorphism was present in the common ancestor of the northern and central Andes clade. Based on parsimonious reconstruction, it is likely that the CAM polymorphism was lost in the clade containing *P. raimondii* and the northern Andes clade (Fig. 2). However, this does not consider whether the CAM polymorphism could be gained in *P. chilensis* by introgressive hybridisation with *P. alpestris* ssp. *zoellneri*.

Assuming that the CAM polymorphism was lost in the clade containing *P. raimondii* and the northern Andes clade, it is possible that weak CAM species could have dispersed across the *Nudo de Pasto*. *P. weberbaueri* is the northern-most species of the early divergent species of the northern and central Andes clade. Its northern-most population occurs at 7° 25' S. This circumstantial evidence suggests that the *Nudo de Pasto* was not only an effective barrier for constitutive CAM species but also for weak CAM species. This is based on the AFLP phylogeny, which only includes 30% of the species in the genus. Future research should incorporate more species to test whether weak CAM species could cross the north of the *Nudo de Pasto*.

Given all the evidence exposed by the present study, two hypotheses are proposed to explain the colonisation of the Andes by the northern and central Andes clade, and by the central and southern Andes clade (Fig. 5 B–C). The colonisation of the Andes proceeded from the south to the north via the east cordillera in the central Andes. In the case of the northern and central Andes clade a recolonization of the south occurred by the yellow clade. This recolonization used the Lomas formation of the Atacama Desert as stepping stones to colonise central Chile. In the central and southern Andes clade, constitutive CAM evolved as a consequence of the colonisation of environments that were drier and had higher cloud cover seasonality. The colonisation of CAM species to higher elevation environments imposed a pressure to increase the cold resistance (either cold tolerance or supercooling) and to increase the flexibility of the enzymes active at night time during the nocturnal CO₂ fixation and malate transportation (Fig. 5F). The present study proposes a hypothesis which explains the evolution of CAM species in montane environments. First, the case that CAM evolved before the colonization of high-elevation environments. In this case cold resistance should increase in comparison with the low-elevation common ancestor by changes in the regulation of protective enzymes and changes in the permeability of the membrane (for cold-tolerant species) or by the maintaining of a gel-like state cytoplasm by increased the synthesis of polysaccharides (supercooling species). In either of the two cold resistance mechanism enzymes involved in nocturnal CO₂ exchange and malate transportation have to be more flexible and/or be expressed relatively higher to compensate the decreased in entropy by low temperatures.

Secondly, the case where CAM evolved at high-elevations. In this case presumably the common ancestor was already a cold resistant plant that colonised drier areas with high cloud cover seasonality. That change in the niche could trigger the

neofunctionalisation of CAM genes. In that case the neofunctionalised enzymes involved in nocturnal CO₂ exchange and malate transportation would be more flexible. Future research should focus on investigating the evolution of CAM in high elevation environments. Moreover, future research can test for positive selection in genes from high-elevation environments and study the molecular structure of PEPC from high-elevation species to test if it is more flexible than the structure of PEPC from low-elevation species.

Varadarajan's hypotheses

Varadarajan (1990) in his excellent study about geographical patterns of *Puya*, established the most likely geographic origin of *Puya* and the direction of colonisation of the Andes by *Puya*. At that time, it was believed that *Puya* was one of the earliest divergent genera in Bromeliaceae, among other genera distributed in the Guiana Shield such as *Brocchinia* (Smith and Downs, 1974). Based on that, he suggested that the most likely direction of colonisation of the Andes was from the north of south America to Chile. The assumption of that hypothesis was falsified by the most recent molecular phylogenies (Crayn *et al.*, 2004; Givnish *et al.*, 2014; Jabaily and Sytsma; 2010, 2013). According to the current understanding, *Puya* is the most derived genera of the polyphyletic subfamily Pitcairnioideae.

Varadarajan (1990) suggested that the evolution of CAM was the result of changing from the semi-xeric environments in the Guiana Shield to more xeric environments in the central Andes. This thesis demonstrates that changes in the climatic niche from mesic to dryer environments is correlated with the evolution of CAM in the genus *Puya*. This process took place mainly in the central Andes. However, the lack of CAM species in the northern Andes (including the Guiana Shield) was an evolutionary loss rather than an ancestral characteristic of *Puya*.

Varadarajan (1990) also discussed whether *Puya* colonised the Andes either horizontally (in the case of the common ancestor that was a high-elevation species) or vertically (in the case of the common ancestor that was a low-elevation species). He assumed that the common ancestor was a high-elevation species and therefore early in the diversification of *Puya* the colonisation of the Andes was horizontal rather than vertical. The present study suggests that the ancestral cold resistance facilitated the colonisation of the Andes. The earliest colonisation of the Andes was initially vertical from the central Chile and the subsequent diversification was in a horizontal direction. As a consequence, the highest abundance of *Puya* species occur at 3000 m (Jabaily and Sytsma, 2013).

A comparison of the hypotheses presented by Varadajan (1990) with the hypotheses presented in this thesis shows the importance of the phylogenetic hypothesis in biogeographical analyses. It is very important to complete a more detailed phylogeny to test the hypotheses that the present study proposes.

The ecological advantage of weak CAM

Given that the common ancestor of *Puya* was a weak CAM species, the ecological relevance of weak CAM must be explored. Weak CAM species in tropical orchids in Panama could increase their survival rates during drought conditions when diurnal CO₂ uptake is reduced (Silvera *et al.* 2005). It has also been suggested that the ecological importance of weak CAM in the juvenile stages of the epiphytes *Guzmania monostachia*, *G. lingulata* and *Werauhia sanguinolenta* has contributed to their survival during drought periods (Pierce *et al.*, 2002a; Beltrán *et al.*, 2013). Winter and Holtum (2015) suggested that cryptic CAM (the weakest version of weak CAM) in *Jatropha curcas* was more closely related with carbon conservation than water conservation. Future studies should investigate in greater

depth the ecological importance of weak CAM species, especially when the weak CAM expression is not modulated by drought (Silvera *et al.*, 2005).

There is evidence of weak CAM in the early divergent species *Ochagavia carnea* (pers. comm., Jamie Males, University of Cambridge) from the subfamily Bromelioideae. According to the present study, the ancestral character state of *Puya* was a weak CAM species. It is therefore very likely that the common ancestor of the subfamily Bromelioideae and the subfamily Puyoideae was a weak CAM species. Future studies should test this hypothesis by evaluating the CAM polymorphism of more basal species in the subfamily Bromelioideae. The common ancestor of the subfamily Bromelioideae and Puyoideae was likely to have originated in central Chile around ~14.1 Ma. At that time, there was a global reduction of CO₂ (Forster *et al.*, 2012). Perhaps the evolution of weak CAM in Puyoideae and Bromelioideae is more related with carbon conservation than water conservation. During the early evolution of *Puya* (Mid Miocene, ~9.4 Ma) the planet experienced the beginnings of another depletion of CO₂ that had its most dramatic peak around ~8–6 Ma (Bolton and Stoll, 2013; Bolton *et al.*, 2016; Amidon *et al.*, 2017). It is possible that weak CAM evolved as a mechanism to conserve CO₂ rather than water. This characteristic may have contributed to the core *Puya* clade having colonised the Andes where the ratio of internal to external partial pressure of CO₂ is low.

Physiological and biochemical adaptations to high-elevation environments in other organisms

Recent studies in other organisms could provide insights for future research avenues for molecular adaptations of CAM plants in high-elevation environments. Physiological and biochemical adaptations to high elevation environments have previously been studied in-depth using animals as models. The deer mouse, *Peromyscus maniculatus*, a model organism to study adaptation in high-elevation environments, occurs

in the western parts of North America from sea level to 4800 m. Cheviron *et al.* (2014) studied the molecular adaptations of populations from high and low elevations. These populations were studied before and after acclimation at low-elevations in a common garden experiment. They measured the apparent $V_{\max}$ of key enzymes related to cellular lipid oxidation and glycolysis, and the relative expression using RNA-seq data. They found that the populations from high elevations always showed higher cellular aerobic capacities and enhanced capacity of cellular lipid oxidation (Cheviron *et al.*, 2012). Moreover, they found a large transcriptomic plasticity that could explain the broad climatic niche of the deer mouse.

Scott *et al.* (2015) studied the changes in gene expression and morphological changes in the gastrocnemius muscle of the deer mouse, using a similar experimental design to Cheviron *et al.* (2014). Scott *et al.* (2015) found higher oxidative fibre density and higher capillary density in the muscles for high-elevation populations, *in situ*, after acclimation at low elevations and in the F₁ under acclimation at low elevations. According to Mahalingan *et al.* (2017) there are morphological differences in the mitochondria of the deer mouse, for example in the distribution of mitochondria per cell of the gastrocnemius muscle and total mitochondrial volume density. The present study found significant physiological differences between high- and low-elevation species in *Puya*. This is similar to the results found by Cheviron *et al.* (2014) and Scott *et al.* (2015) under common garden conditions for the deer mouse. It is possible that if the same physiological variables were measured *in situ*, then the differences between high- and low-elevations would be maintained but perhaps the absolute difference would be higher.

Future studies should investigate the proteome and transcriptome of *Puya castellanosi* and *P. harmsii* and other high-elevation CAM species that are distributed even higher and more northerly than *Puya castellanosi* and *P. harmsii*, for example *P. cerrateana*,

P. ctenorhyncha and *P. longistyla*. These species could also be compared with other low-elevation CAM species such as *P. laxa*, which is distributed more northerly and is from low elevations. The molecular adaptations of high-elevation plant species can be revealed by comparing the morphology, physiology, proteome of different populations at different elevations such as in *Kobresia pygmaea* (a high-elevation species from the central Tibetan Plateau; Li *et al.*, 2014); *Lamiophlomis rotata* (an alpine plant with a wide elevation range from the northern Tibetan Plateau; Ma *et al.*, 2015a); and *Potentilla saundersiana* (an alpine plant with a wide elevation range from the Northwestern Tibetan Plateau; Ma *et al.* 2015b). Furthermore, the molecular adaptations could be studied by comparing the transcriptomes of high-elevation species such as *Lobelia aberdarica* and *L. telekii* (afroalpine species; Zhao *et al.*, 2016). Zhao *et al.* (2016) demonstrated that genes that evolved under positive selection are related to cold stimulus, DNA repair and regulation of gene expression which are consistent with characteristics of high-elevation environments. None of these species show CAM photosynthesis.

Summary of the key results

Hypothesis 1: Constitutive CAM species and non-constitutive CAM species of *Puya* have different climatic niches.

The present study demonstrates that the differences in the climatic niches of constitutive and non-constitutive CAM species of *Puya* are driven by aridity index, cloud cover seasonality, mean annual cloud frequency and precipitation seasonality. Despite the high degree of niche conservatism in *Puya* during the evolution of CAM, there has been a niche change between constitutive CAM clades and non-constitutive CAM clades.

Hypothesis 2: CAM evolved multiple times in *Puya*

The present study demonstrates that constitutive CAM evolved multiple times in *Puya* but neofunctionalised CAM genes evolved only once. The evolution of constitutive CAM was complex. Using only carbon isotope values, it was not possible to reconstruct the ancestral photosynthetic pathway of the common ancestor. By evaluating the evolution of the PEPC genes, it was possible to determine that the most likely photosynthetic pathway of the common ancestor of *Puya* was weak CAM. Weak CAM was lost in the northern and central Andes clade. Therefore, while the present study concludes that constitutive CAM evolved multiple times, the neofunctionalised enzymatic machinery was present in the common ancestor of *Puya*.

Hypothesis 3: There are physiological and biochemical differences between low-elevation and high-elevation species of *Puya*.

The present study demonstrates that there are significant differences in the cold resistance and the relative activity of PEPC at low and high temperature between high- and low-elevation CAM species. It is likely that the common ancestor of *Puya* was a cold-resistant plant. The present study suggests that, during the colonisation of the Andes, the cold-resistance increased and that this increase was reflected in the temperature boundaries of PEPC activity.

Hypothesis 4: There are differences in the amino acid residues of PEPC between CAM and non-constitutive CAM species of *Puya*

CAM polymorphism

The present study demonstrates that the most expressed isoform of PEPC in *Puya* species is an orthologue of the PEPC isoform involved in CAM in *Ananas comosus*. By using five independent origins of CAM in monocots, the present study demonstrates that the

residue Arg6779 or Asn679 (in orchids) is a molecular signature of CAM plants in the CAM PEPC isoform.

CAM polymorphism in *Puya*

By evaluating this signature in *Puya* species, the present study demonstrates that the non-constitutive CAM species from the central and southern Andes clade, the blue clade and *P. ferruginea* contained the molecular CAM polymorphism. The relative expression of the two plant PEPC isoforms confirms that the non-constitutive CAM plants from the central and southern Andes clade, the blue clade and *P. ferruginea* are part of the C₃–CAM continuum described in Silvera *et al.* (2010b). *P. raimondii* was the only species that did not show the CAM polymorphism. Given the ratio of PEPC expression, it is likely that it belongs to the constitutive C₃ part of the C₃–CAM continuum. The use of carbon isotope values and rapid screening of the CAM polymorphism can improve ancestral photosynthetic reconstructions. Future research should focus on identify similar polymorphisms in the CAM species in eudicots.

Conclusions

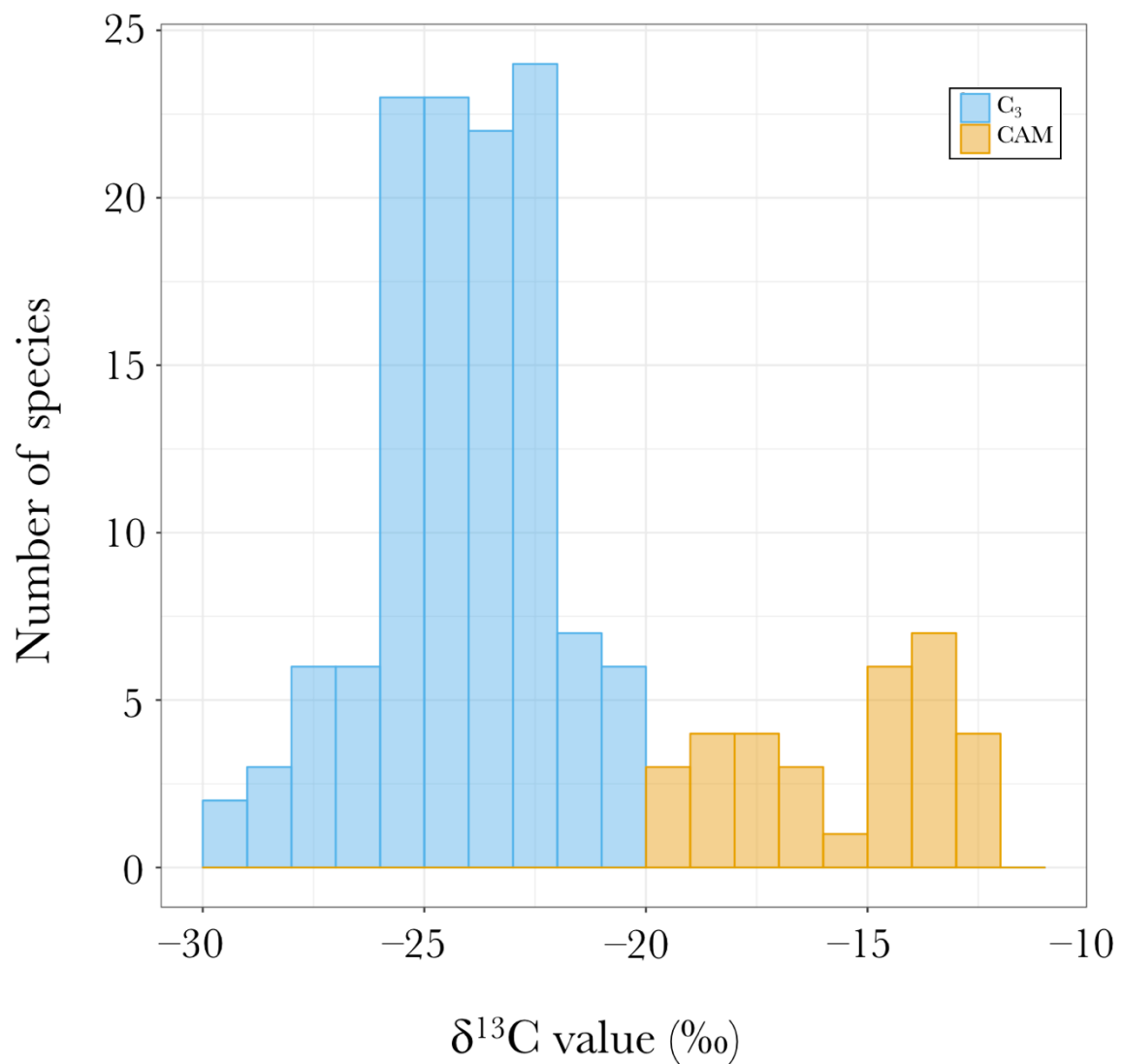
The aim of this thesis was to uncover new evidence to understand the ecophysiological and evolutionary significance of CAM in the montane genus *Puya*. Given that *Puya* is a montane genus, and that the great majority of species occur in the Andes, this thesis proposed a general hypothesis for the evolution of CAM in mountain environments. During the colonisation of the Andes, high elevation species were pressured by new conditions to increase their cold resistance and flexibility in the PEPC CAM isoform. The proposed hypothesis is that species from high elevations have higher cold resistance than their low-elevation counterparts by maintaining higher levels of protective enzymes and higher permeability of the membranes or higher synthesis of polysaccharides, depending on the cold resistance mechanism. In addition, the enzymes

involved in nocturnal CO₂ fixation and malate transportation have a relatively higher flexibility and/or are relatively highly expressed in comparison with their low-elevation relatives, to compensate for the low entropy at low temperatures.

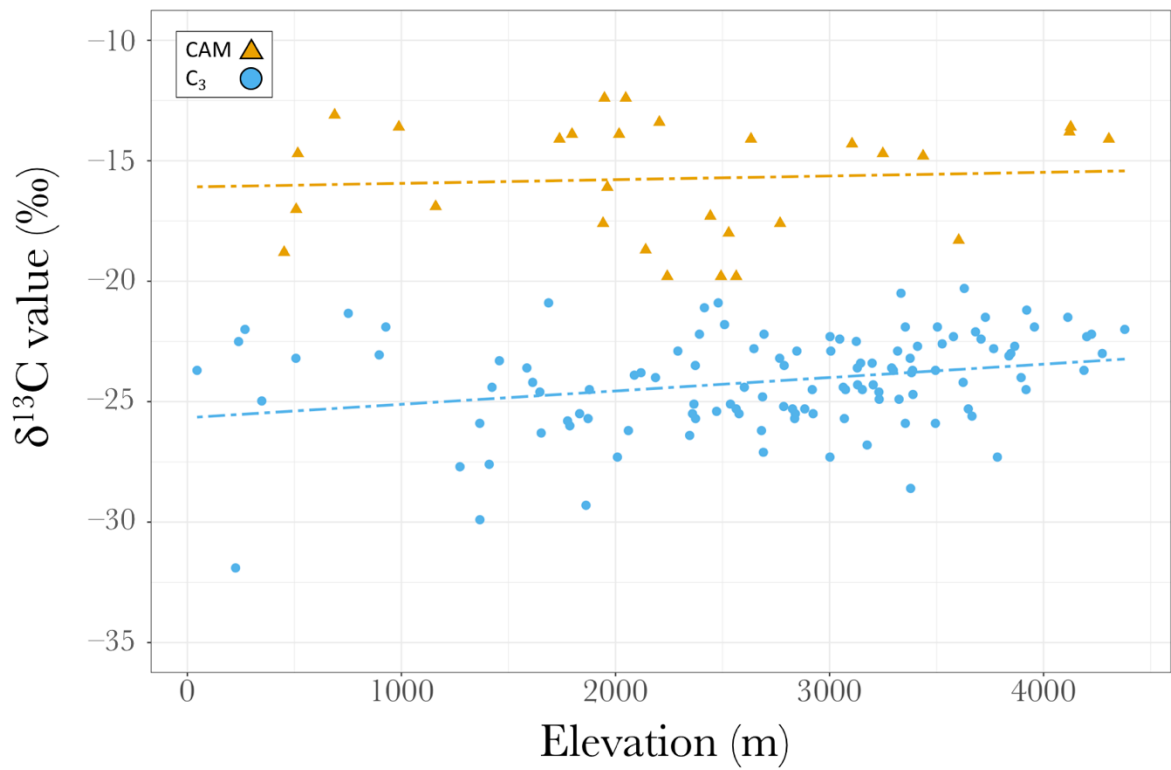
The colonisation of the Andes was facilitated because the common ancestor of *Puya* was likely to be a cold resistant plant. Once the Andean *Puya* diversified across the Andes (from south to north) constitutive CAM evolved multiple times associated with changes in the climatic niche. These multiple events cannot be considered to be independent because it is very likely that the common ancestor of *Puya* was a weak CAM species. The species from the northern Andes clade lost their CAM capacity and this contributed to the colonisation of the northern Andes by *Puya*. The present study found suitable areas for CAM in the northern Andes. However, there are no indications of constitutive CAM species in this region of the Andes. By studying *Puya*, this thesis has uncovered relevant new evidence about the ecophysiology and evolution of CAM in the Andes.

Supplementary Figures

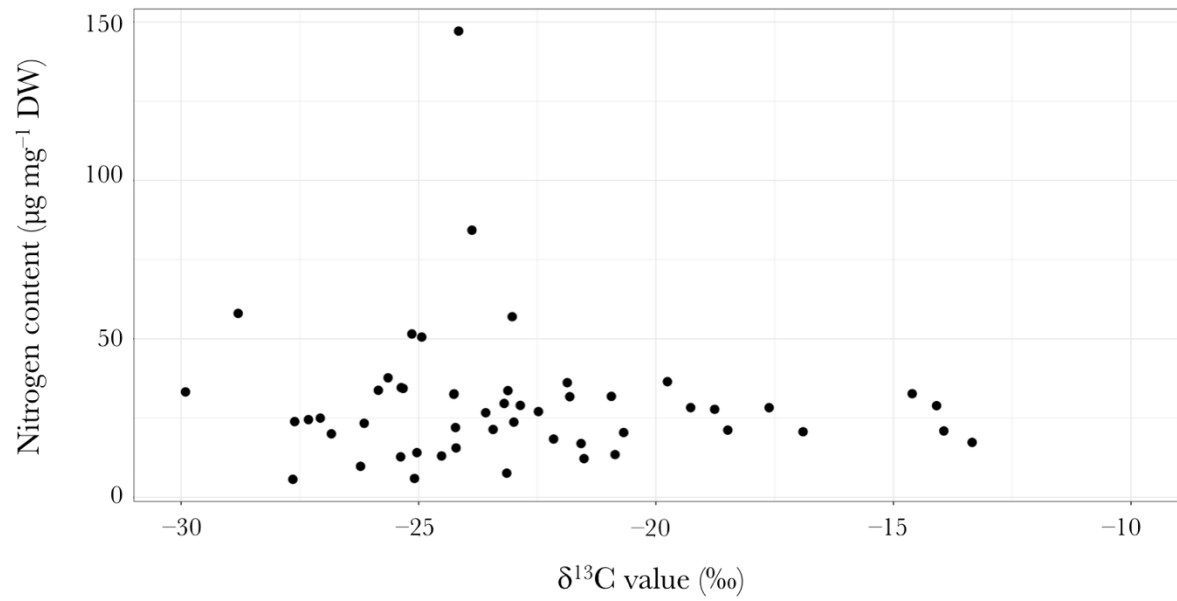
Supplementary Figures – Chapter 2



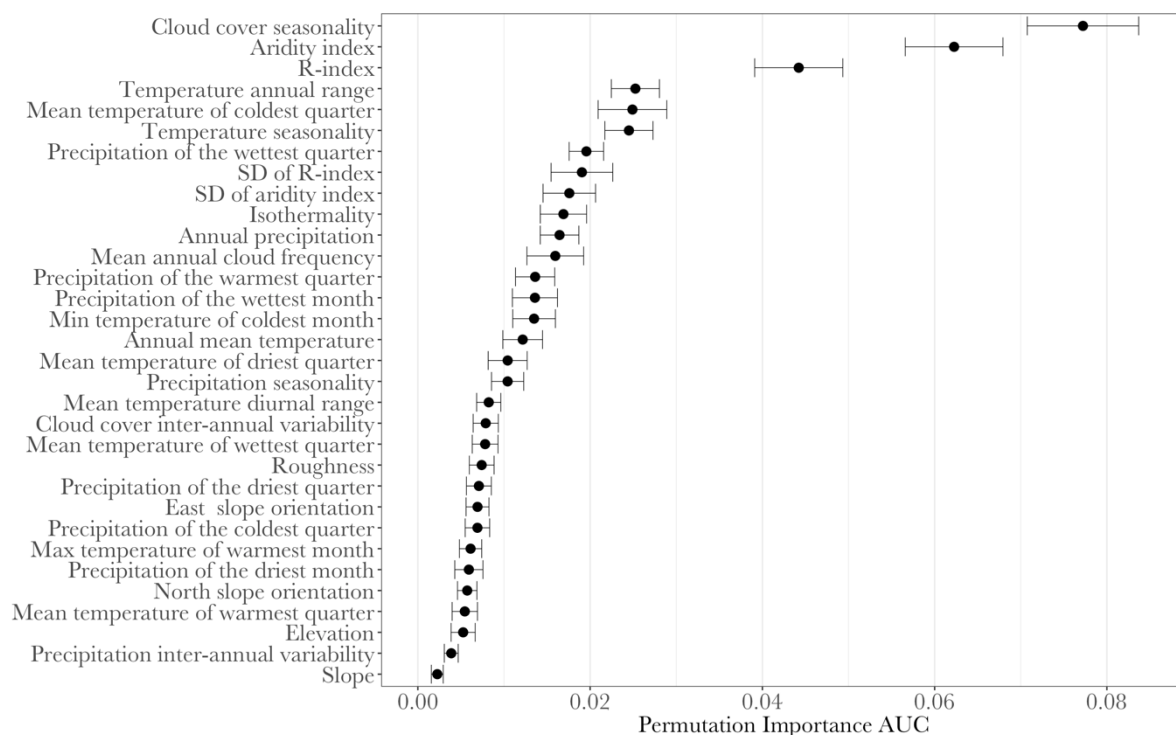
Supplementary Figure 1. Frequency histogram of $\delta^{13}\text{C}$ values of 161 species of *Pitya* plotted in class intervals of 1.0‰.



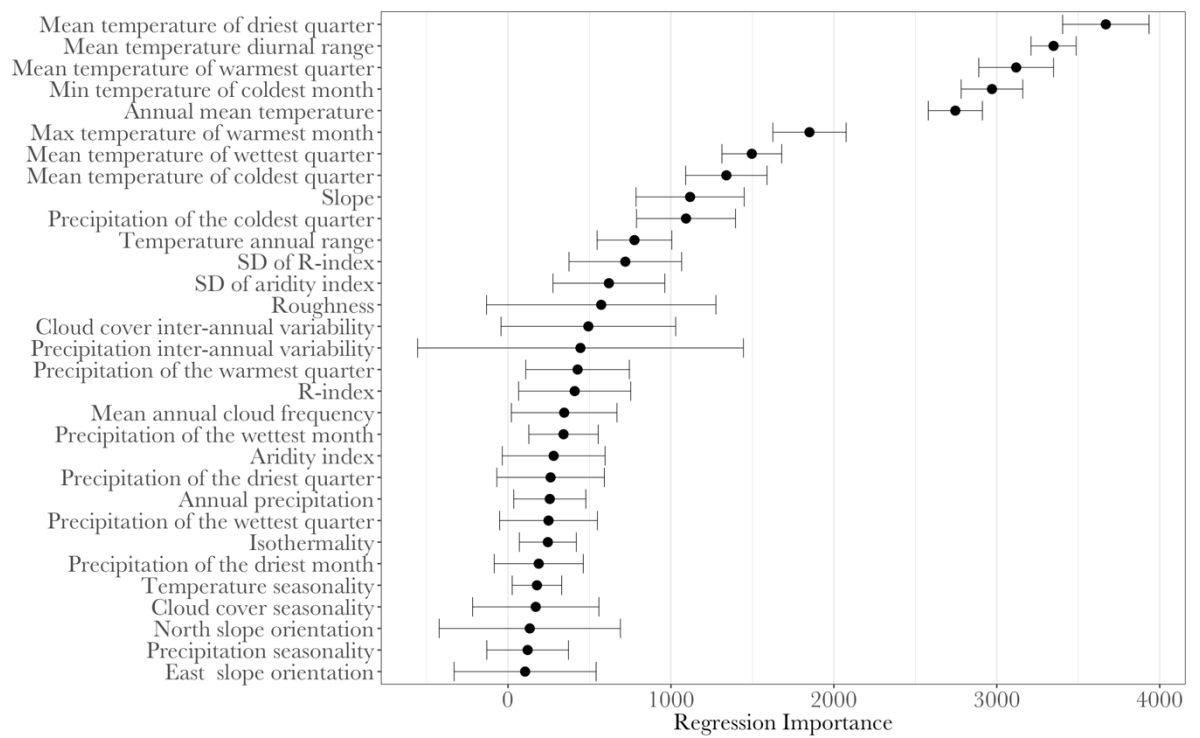
Supplementary Figure 2. Correlation between $\delta^{13}\text{C}$ values and elevation in 147 species of *Puya*. C₃ species (116 species; blue circles) show a positive and significant correlation (slope = 0.0006, $p = 0.0026$, $r^2 = 0.0674$, $y = -25.66 + 0.0006 x$), while CAM species (31 species; orange triangles) do not show a significant correlation (slope = 0.0001, $p = 0.7237$, $r^2 = 0.005$, $y = -16.09 + 0.0001 x$).



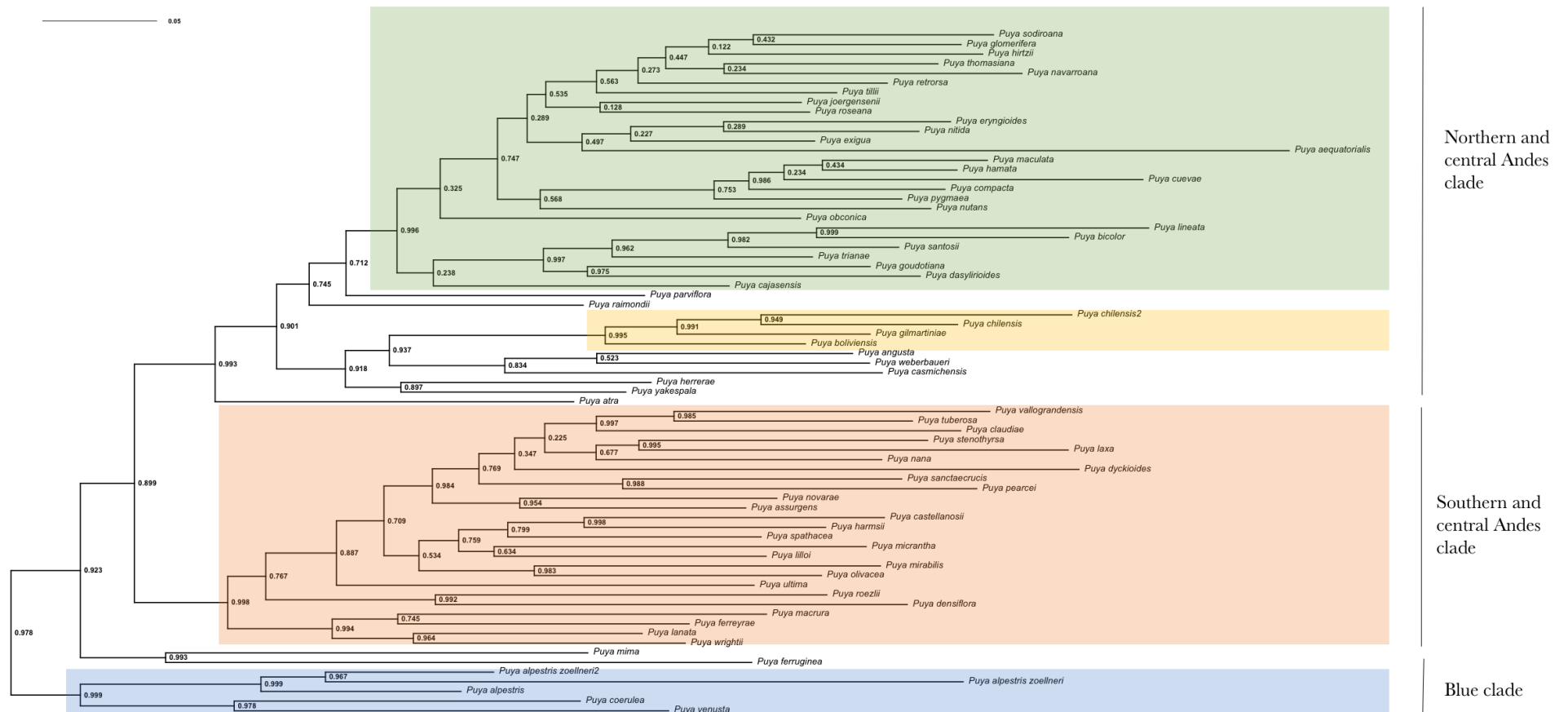
Supplementary Figure 3. Correlation between $\delta^{13}\text{C}$ value and nitrogen content in 52 species of *Puya*. There is not a significant linear correlation between $\delta^{13}\text{C}$ values and nitrogen content ($p = 0.4916$).



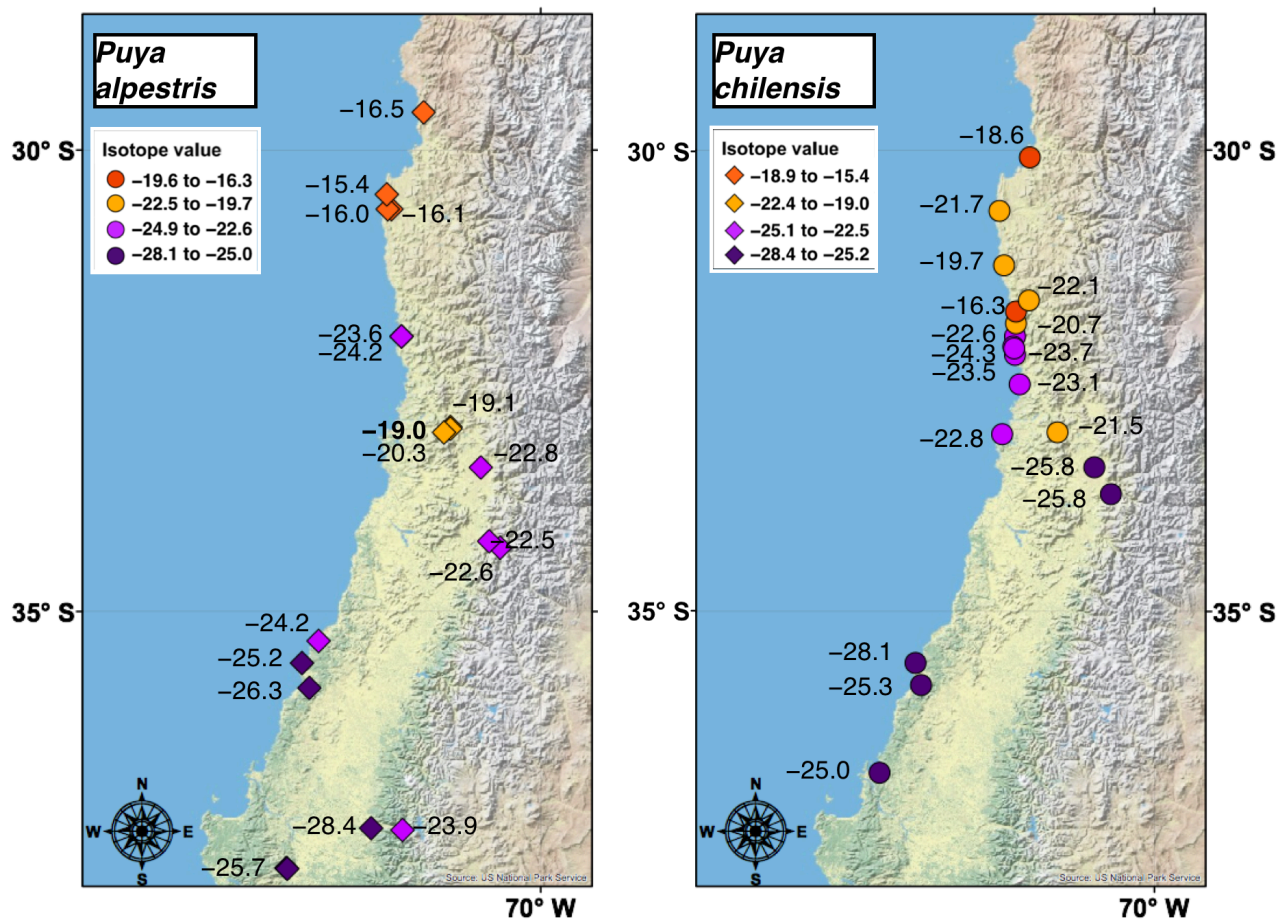
Supplementary Figure 4. Mean permutation importance of 32 climatic variables in 50 independent Random Forest ordinations to classify the photosynthetic pathway of 110 species of *Puya* from central and southern Andes and Chile. The importance is relative to the bars represent one standard deviation above and below the mean value. The permutation importance is calculated based on the change of accuracy of the area under the receiver operating characteristic curve (ROC) after permutation each variable.



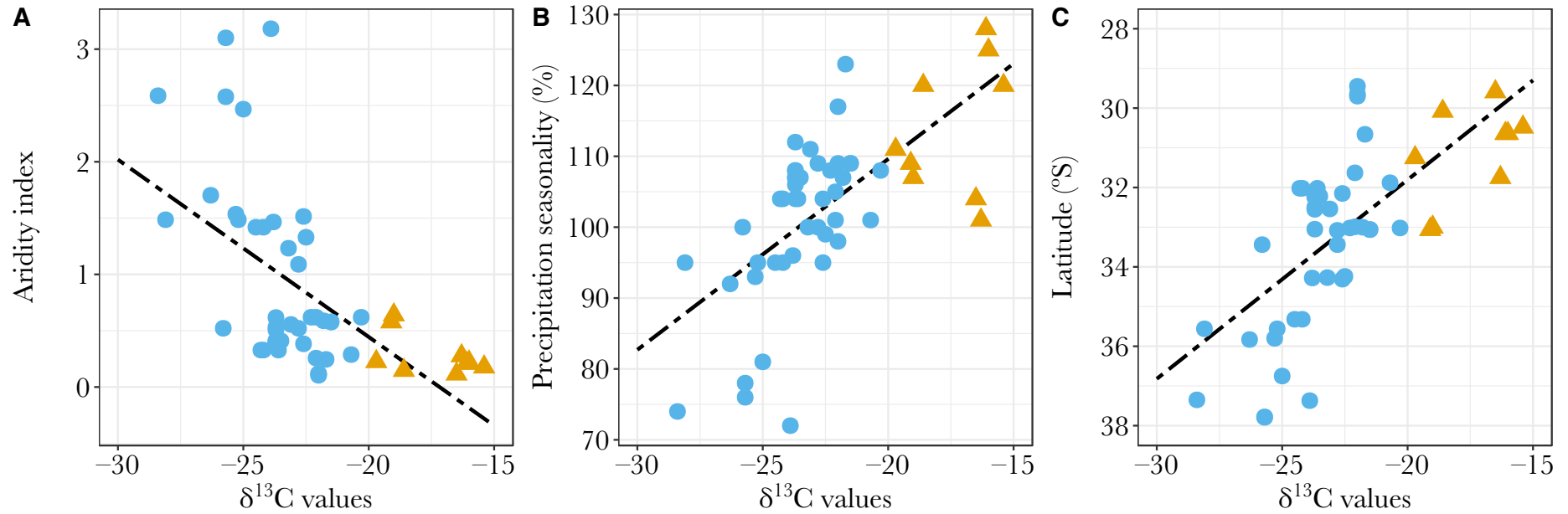
Supplementary Figure 5. Mean regression importance of 32 climatic variables to model elevation in 50 independent Random Forest regressions of 31 CAM species of *Puya*. The bars represent one standard deviation above and below the mean value.



Supplementary Figure 6. Bayesian inference phylogeny of 69 species of *Puya* based on AFLP information from Jabaily and Sytsma (2013). Posterior probability is shown at each node. The green, yellow, orange and blue highlighting indicates the northern Andes, yellow, southern and central Andes, and blue clades, respectively.

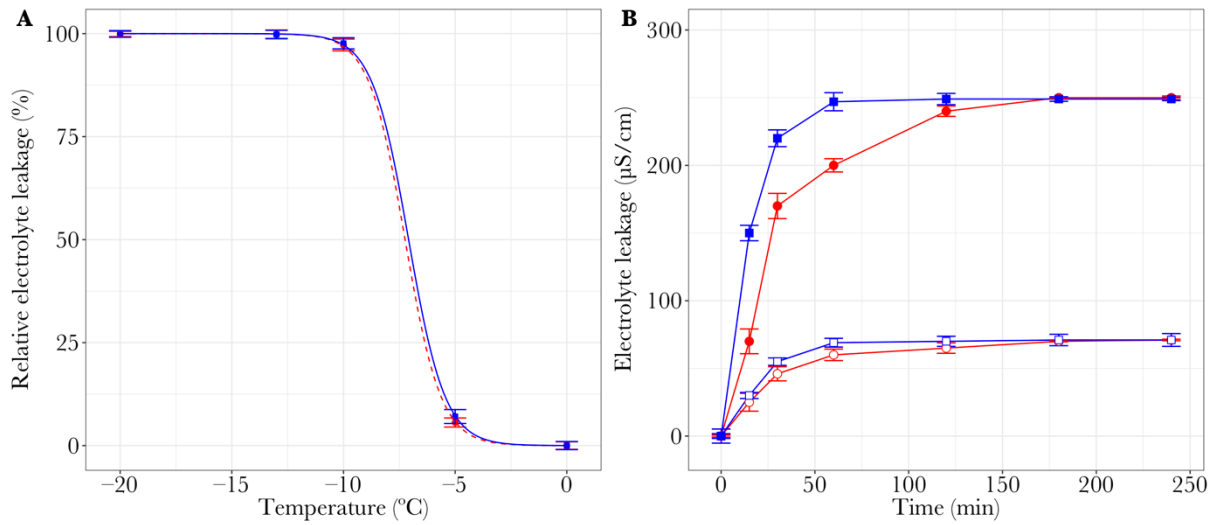


Supplementary Figure 7. Distribution of *Puya alpestris* and *Puya chilensis* by their carbon isotope values. Diamonds represent populations of *P. alpestris*, and circles represent populations of *P. chilensis*. Dark orange colours show constitutive CAM populations.

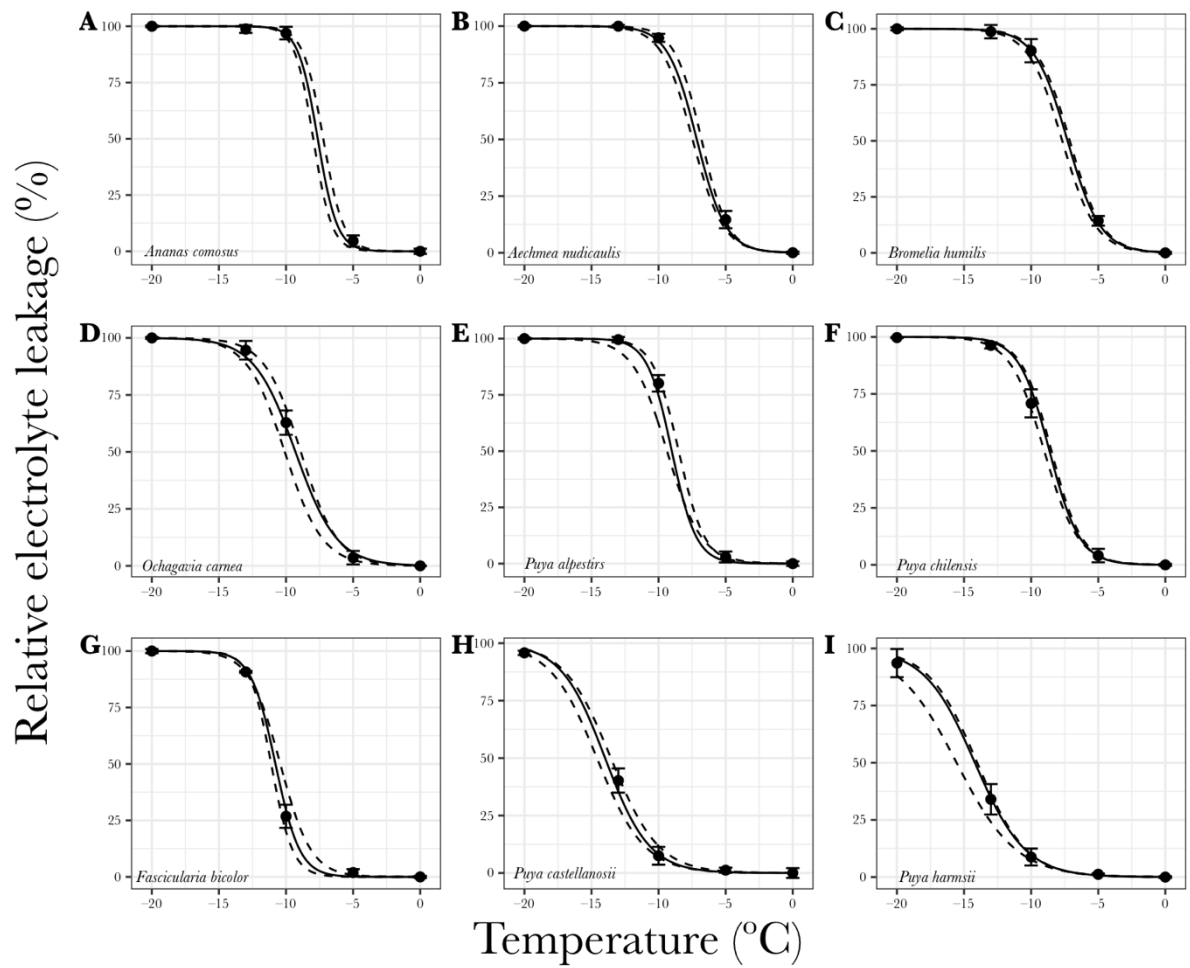


Supplementary Figure 8. Linear correlation between aridity index, precipitation seasonality latitude and $\delta^{13}\text{C}$. the carbon isotope values and climatic variables of 5 species (51 individuals) of *Puya* from Chile. **(A)** Correlation between aridity index and $\delta^{13}\text{C}$ values (slope = -0.1575 , $p < 0.001$, $r^2 = 0.3251$, $y = -2.8294 - 0.1575x$). **(B)** Correlation between precipitation seasonality and $\delta^{13}\text{C}$ values (slope = 2.6827 , $p < 0.001$, $r^2 = 0.4215$, $y = 163.3745 + 2.6827x$). **(C)** Correlation between latitude and $\delta^{13}\text{C}$ values (slope = 0.5015 , $p < 0.001$, $r^2 = 0.4453$, $y = -21.7767 + 0.5015x$). The species included are *P. alpestris*, *P. chilensis*, *P. coerulea*, *P. gilmartiniae*. Orange triangles represent constitutive CAM populations. Blue circles represent non-constitutive CAM populations.

Supplementary Figures – Chapter 3



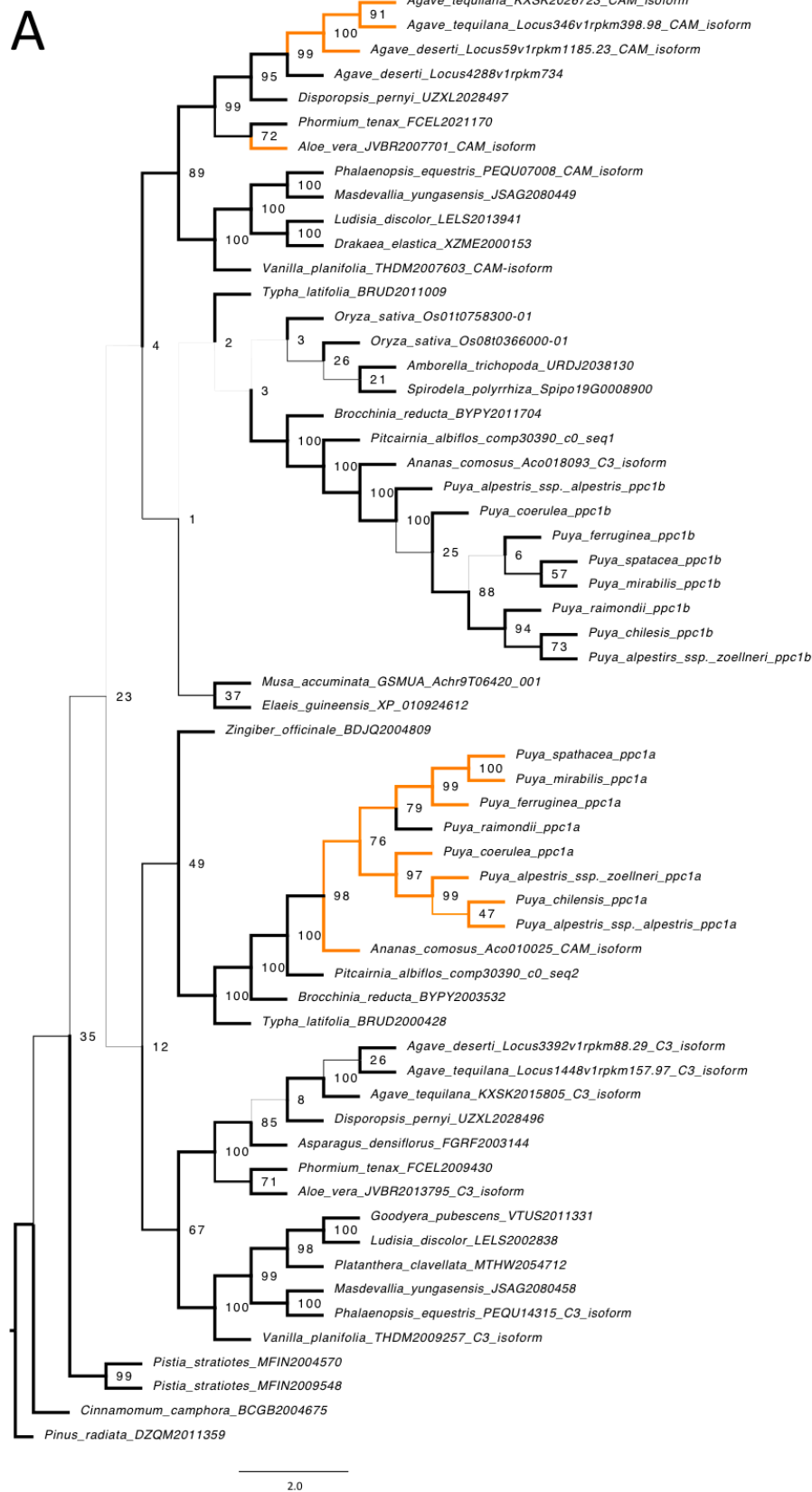
Supplementary Figure 1. Optimisation of the electrolyte leakage protocol followed in the present study. **(A)** Relative electrolyte leakage of rectangular-shaped (blue squares) or circular-shaped (red circles) excised tissue from leaves of *Ananas comosus*. Five excised tissues were incubated for 3 h at -80 , -20 , -13 , -10 , -5 , 0 and 4 °C (data for -80 °C and 4 °C are not shown), then submerged in distilled water (5 ml) for 1 h. The first measured of electrolyte leakage was obtained. The water and the tissue in a sealed vial where incubating at 100 °C for 1 h. After reach room temperature, the electrolyte leakage was measured again. The relative electrolyte leakage are the ratios of initial over final electrolyte leakage. The data was transform to so the total damage was produced at -80 °C. The logistic regression of the relative electrolyte leakage of rectangular-shaped and circular-shaped tissues are shown by the blue continuous line and red dashed line, respectively. There were no obvious differences between circular- and rectangular-shaped tissues in electrolyte leakage. **(B)** Electrolyte leakage over time using rectangular (blue squares) or circular-shaped (red circles) tissue from leaves of *Ananas comosus*. The tissues were exposed to liquid nitrogen (filled points) or kept at 4 °C (open points) for 3 hours. After that time, the tissue was submerged in distilled water (5 ml) and electrolyte leakage measured after 15, 30, 60, 120, 180 and 240 minutes. Data are means $\pm$ standard deviation, $n = 3$ experiments.



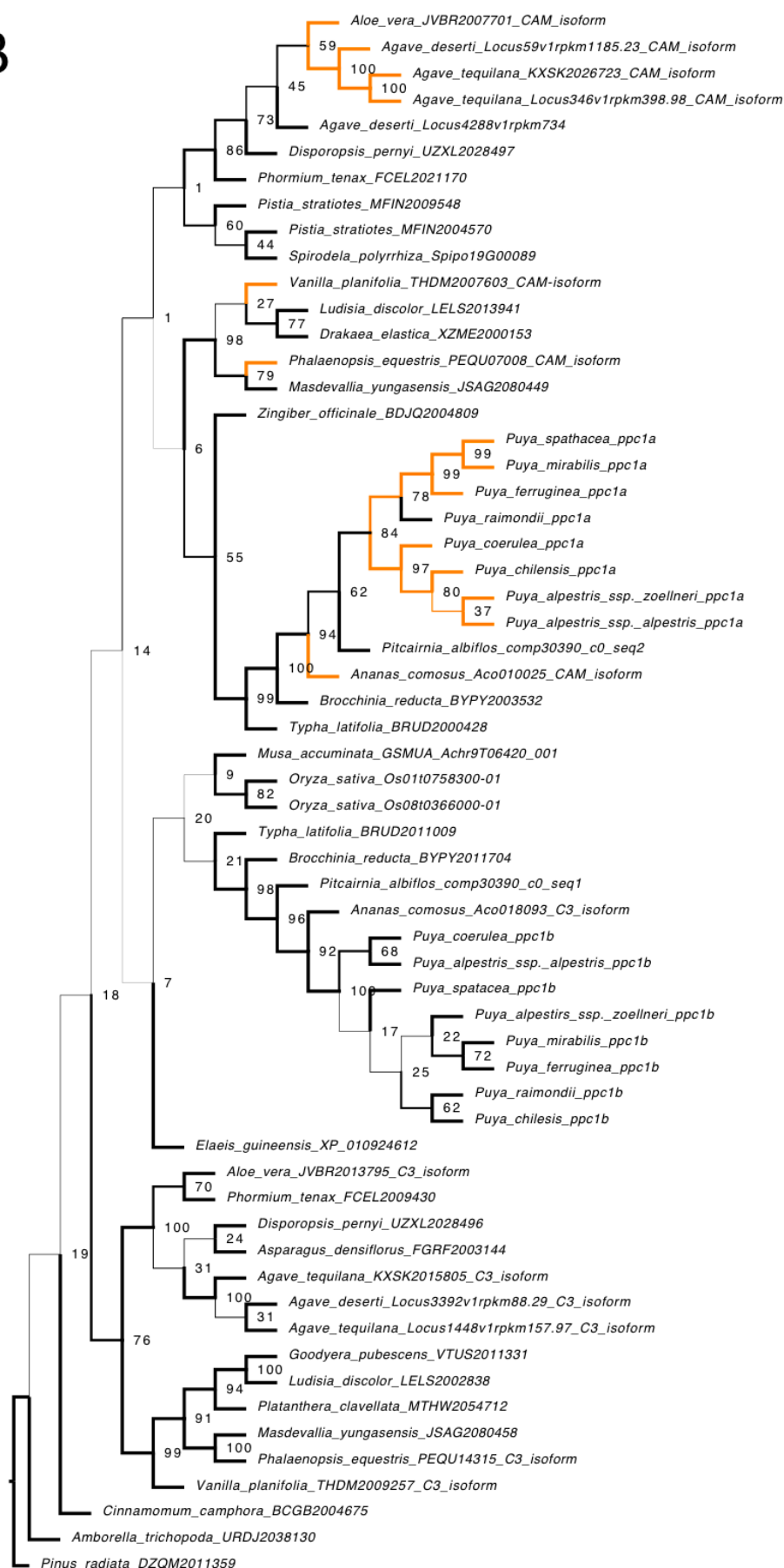
Supplementary Figure 2. Quasi-binomial logistic regression of relative electrolyte leakage and temperature. Five rectangular-shaped excised tissues were incubated for 3 h at -80 , -20 , -13 , -10 , -5 , 0 and 4 °C (data for -80 °C and 4 °C are not shown), then submerged in distilled water (5 ml) for 1 h. The first measured of electrolyte leakage was obtained. The water and the tissue in a sealed vial where incubating at 100 °C for 1 h. After reach room temperature, the electrolyte leakage was measured again. The relative electrolyte leakage are the ratios of initial over final electrolyte leakage. The data was transform to so the total damage was produced at -80 °C. The leaf tissue came from the following species: **(A)** *Ananas comosus*, **(B)** *Aechmea nudicaulis*, **(C)** *Bromelia humilis*, **(D)** *Ochagavia carnea*, **(E)** *Puya alpestris*, **(F)** *Puya chilensis*, **(G)** *Fascicularia bicolor*, **(H)** *Puya castellanosi*, and **(I)** *Puya harmsii*. The continuous lines represent the mean logistic regression of all independent experiments. The dashed lines represent the standard error of the mean of the regression. For the Table. 1. each independent experiment derived a single T_{50} value. n = independent experiments; $n = 3$ for *Aechmea nudicaulis* and *Puya*

chilensis; n = 6 for *Bromelia humilis*, *Ochagavia carnea*, and *P. castellanosi*; n = 9 for *Ananas comosus*, *Fascicularia bicolor*, *P. alpestris* ssp. *zoellneri* and *P. harmsii*.

Supplementary Figures – Chapter 4



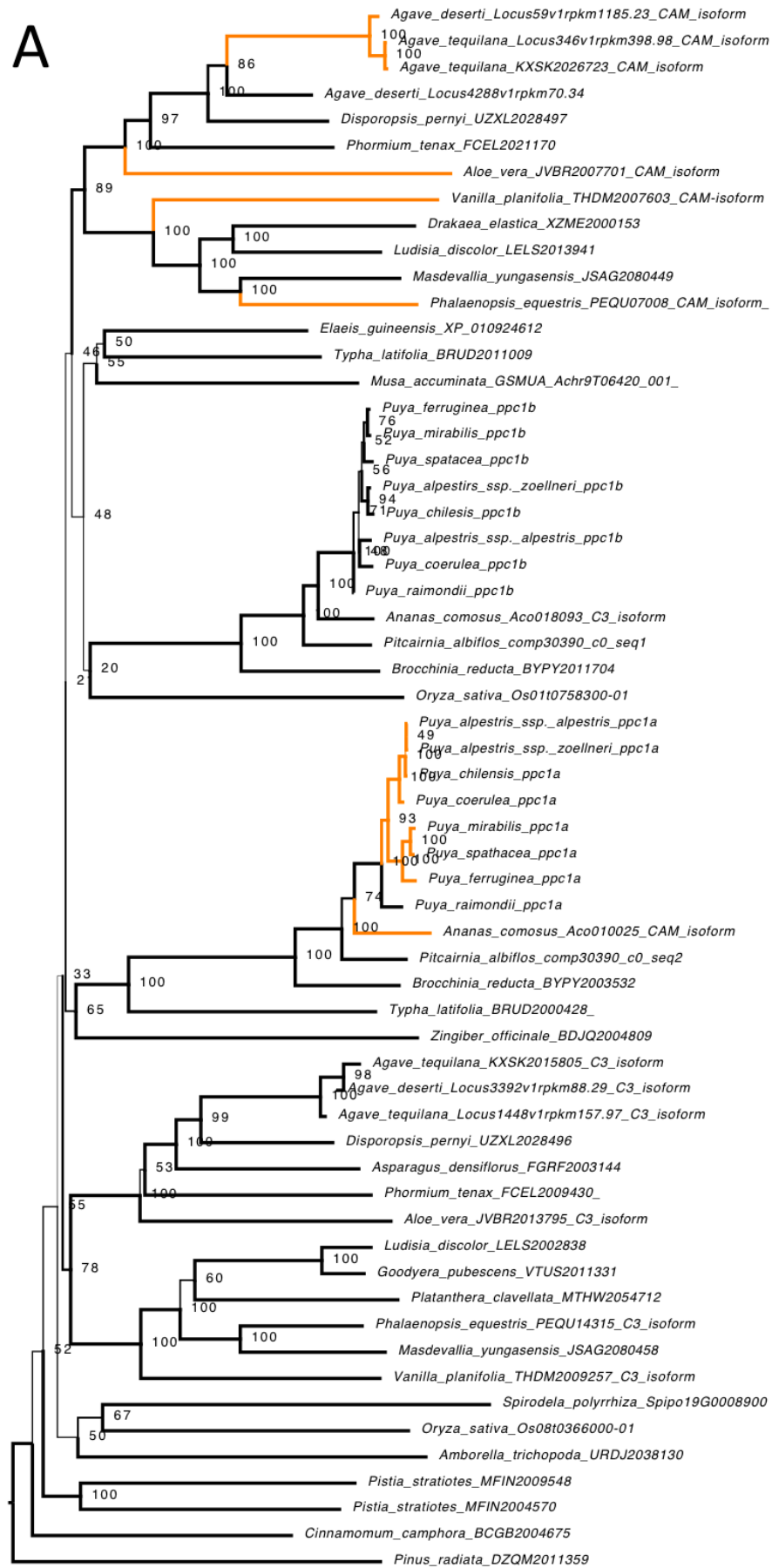
B



2.0

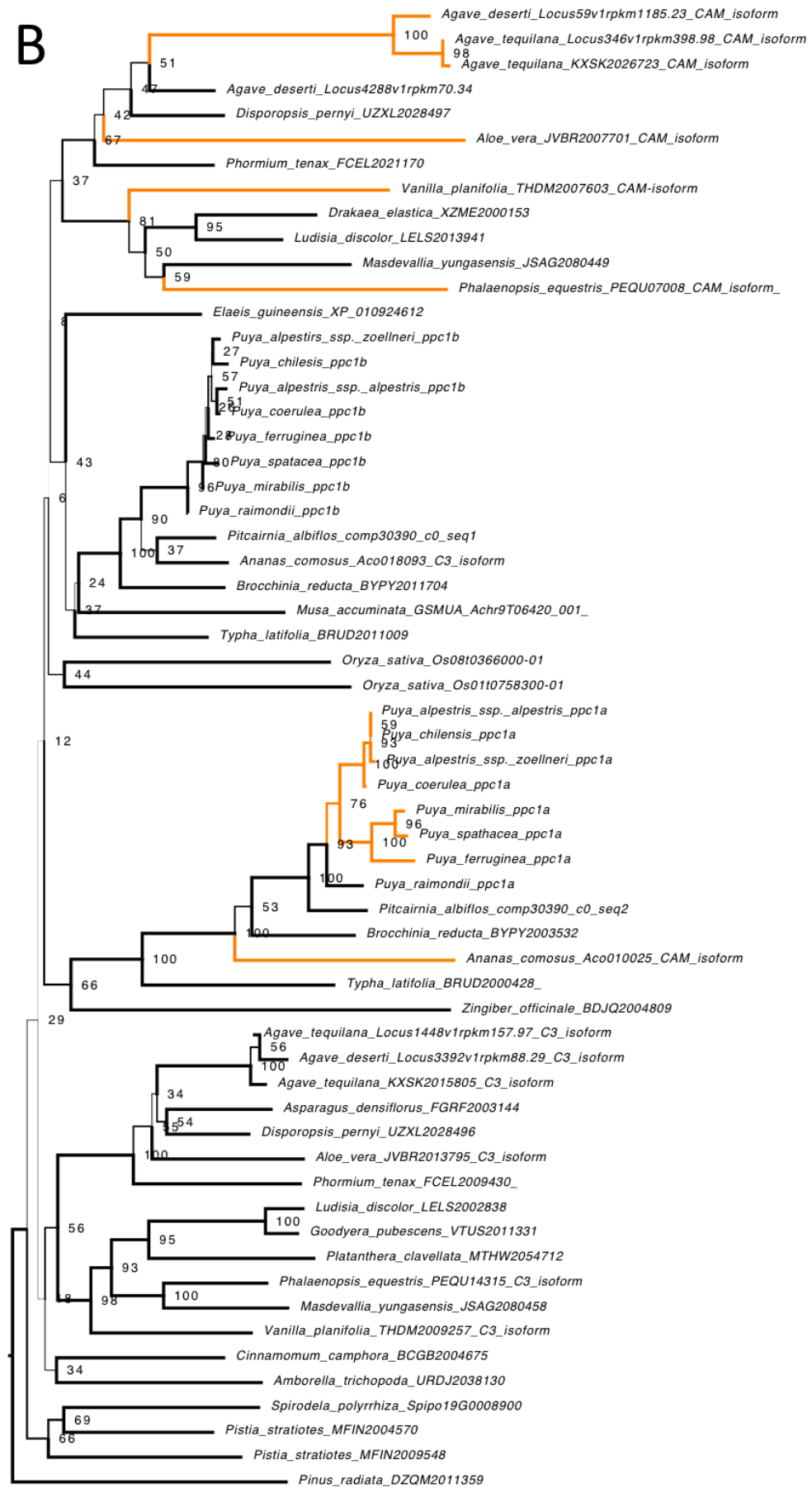
Supplementary Figure 1. Maximum parsimony (MP) trees showing the evolution of *ppc-1* genes in monocots. Branch support was based on 1000 bootstrap replicates. The branch thickness is directly proportional with the branch support. **(A)** Phylogenetic relationships among *ppc-1* using the nucleotide alignment. **(B)** Phylogenetic relationships among *ppc-1* using the amino acid alignment. Orange branches are taxa showing the residue Arg679 or Asn679 (*A. comosus* CAM isoform numbering) in the PEPC sequence, black branches are taxa exhibiting the residue His679 (*A. comosus* CAM isoform numbering) in the PEPC sequence.

A



0.02

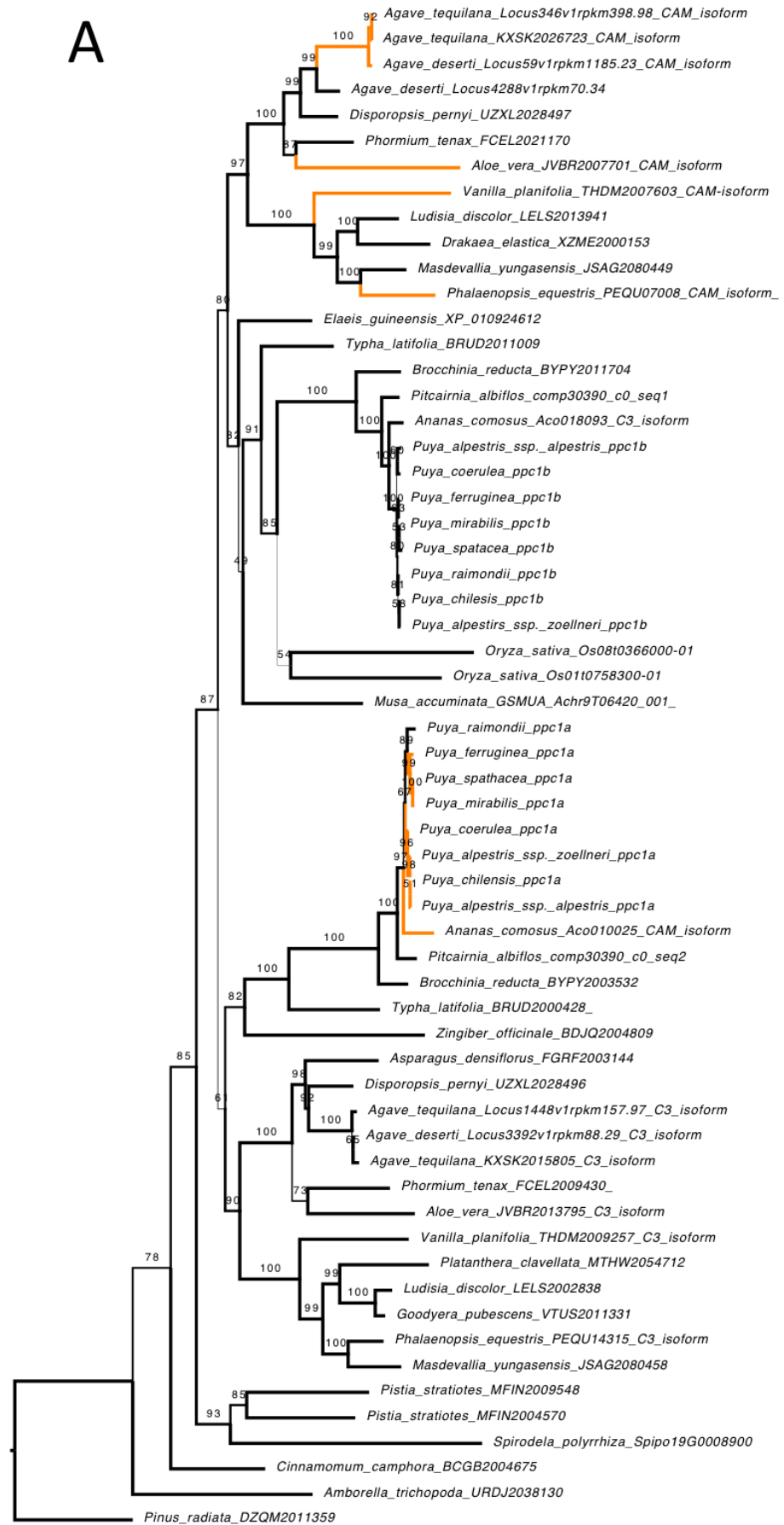
B



0.02

Supplementary Figure 2. Neighbour-joining (NJ) trees showing the evolution of *ppc-1* genes in monocots. Branch support was based on 1000 bootstrap replicates, the branch thickness is directly proportional with the branch support. **(A)** Phylogenetic relationships among *ppc-1* using the nucleotide alignment. **(B)** Phylogenetic relationships among *ppc-1* using the amino acid alignment. Orange branches are taxa showing the residue Arg679 or Asn679 (*A. comosus* CAM isoform numbering) in the PEPC sequence, black branches are taxa exhibiting the residue His679 (*A. comosus* CAM isoform numbering) in the PEPC sequence.

A



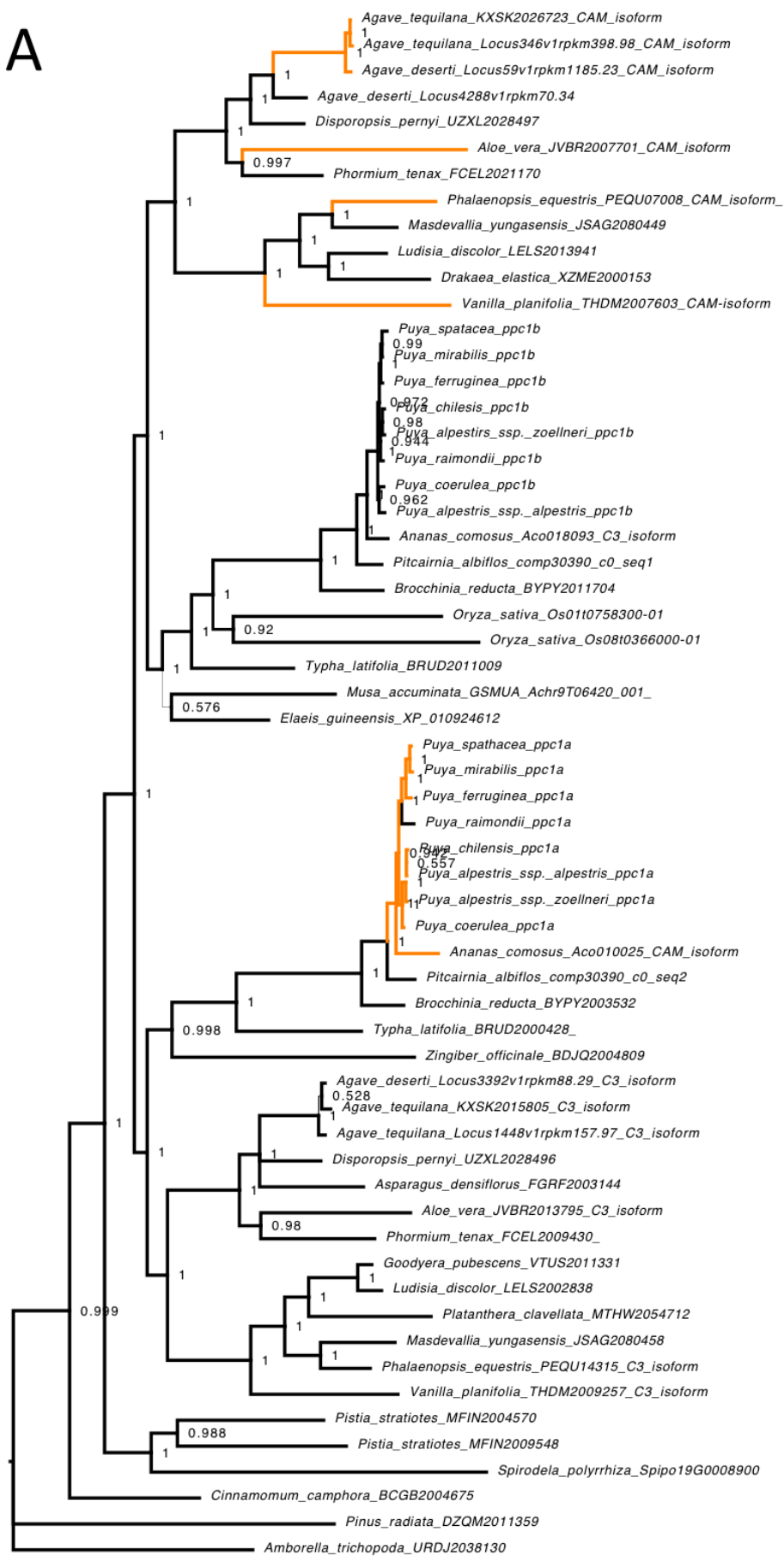
0.06

B



Supplementary Figure 3. Maximum likelihood (ML) trees showing the evolution of *ppc-1* genes in monocots. Branch support was based on 1000 bootstrap replicates. The branch thickness is directly proportional with the branch support. **(A)** Phylogenetic relationships among *ppc-1* using the nucleotide alignment. **(B)** Phylogenetic relationships among *ppc-1* using the amino acid alignment. Orange branches are taxa showing the residue Arg679 or Asn679 (*A. comosus* CAM isoform numbering) in the PEPC sequence, black branches are taxa exhibiting the residue His679 (*A. comosus* CAM isoform numbering) in the PEPC sequence.

A



0.04

B



0.02

Supplementary Figure 4. Bayesian inference (BI) trees showing the evolution of *ppc-1* genes in monocots. Branch support was based on posterior probability. The branch thickness is directly proportional with the branch support. **(A)** Phylogenetic relationships among *ppc-1* using the nucleotide alignment. **(B)** Phylogenetic relationships among *ppc-1* using the amino acid alignment. Orange branches are taxa showing the residue Arg679 or Asn679 (*A. comosus* CAM isoform numbering) in the PEPC sequence, black branches are taxa exhibiting the residue His679 (*A. comosus* CAM isoform numbering) in the PEPC sequence.

Supplementary Tables

Supplementary Tables – Chapter 1

Supplementary Table 1. List of the taxonomic names mentioned in this thesis and the accepted taxonomic names of *Puya* species.

Accepted species names of *Puya*

Puya adscendens L.B. Smith

Puya aequatorialis André

Puya alata L.B. Smith

Puya alba L.B. Smith

Puya alpestris (Poeppig) Gay

Puya alpestris ssp. *alpestris* (Poeppig) Zizka, J.V. Schneid. & Novoa

Puya alpestris ssp. *zoellneri* (Mez) Zizka, J.V. Schneid. & Novoa

Puya alpicola L.B. Smith

Puya angelensis E. Gross & Rauh

Puya angulonis L.B. Smith

Puya angusta L.B. Smith

Puya antioquiensis L.B. Smith & R.W. Read

Puya araneosa L.B. Smith

Puya argentea L.B. Smith

Puya aristeguietae L.B. Smith

Puya asplundii L.B. Smith

Puya assurgens L.B. Smith

Puya atra L.B. Smith

Puya barkleyana L.B. Smith

Puya bermejiana Gómez, Slanis & Grau
Puya bicolor Mez
Puya boliviensis Baker
Puya boopiensis R. Vásquez, Ibisch & R. Lara
Puya boyacana Cuatrecasas
Puya brachystachya (Baker) Mez
Puya brackeana Manzanares & W. Till
Puya bravoii Aráoz & A. Grau
Puya brittoniana Baker
Puya cahuachensis A. Galán, J. Montoya, J.A. Vicente & E. Linares
Puya cajasensis Manzanares & W. Till
Puya cardenasii L.B. Smith
Puya cardonae L.B. Smith
Puya casmichensis L.B. Smith
Puya castellanosii L.B. Smith
Puya cerrateana L.B. Smith
Puya chilensis Molina
Puya claudiae Ibisch, R. Vásquez & E. Gross
Puya clava-herculis Mez & Sodiro
Puya cleefii L.B. Smith & R.W. Read
Puya cochabambensis R. Vásquez & Ibisch
Puya coerulea Lindley
Puya commixta L.B. Smith
Puya compacta L.B. Smith
Puya coriacea L.B. Smith

Puya cristata L.B. Smith
Puya cryptantha Cuatrecasas
Puya ctenorhyncha L.B. Smith
Puya cuatrecasasii L.B. Smith
Puya cuevae Manzanares & W. Till
Puya cylindrica Mez
Puya dasylirioides Standley
Puya densiflora Harms
Puya depauperata L.B. Smith
Puya dichroa L.B. Smith & R.W. Read
Puya dodsonii Manzanares & W. Till
Puya dolichostrobila Harms
Puya donneriana R. Vasquez, S. Altamirano & Ibisch
Puya dyckiioides (Baker) Mez
Puya elviragrossiae R. Vásquez & P.L. Ibisch
Puya entre-riosensis Ibisch & E. Gross
Puya erlenbachiana Ibisch & R. Vásquez
Puya eryngioides André
Puya exigua Mez
Puya exuta L.B. Smith & R.W. Read
Puya fastuosa Mez
Puya ferox Mez
Puya ferreyrae L.B. Smith
Puya ferruginea (Ruiz & Pavón) L.B. Smith
Puya fiebrigii Mez

Puya floccosa (Linden) E. Morren
Puya fosteriana L.B. Smith
Puya fulgens L.B. Smith
Puya furfuracea (Willdenow) L.B. Smith
Puya gargantae L.B. Smith
Puya gerd-muelleri W. Weber
Puya gerdae W. Weber
Puya gigas André
Puya gilmartiniae G.S. Varadarajan & Flores
Puya glabrescens L.B. Smith
Puya glandulosa L.B. Smith
Puya glareosa L.B. Smith
Puya glaucovirens Mez
Puya glomerifera Mez & Sodiro
Puya goudotiana Mez
Puya gracilis L.B. Smith
Puya grafii Rauh
Puya grandidens Mez
Puya grantii L.B. Smith
Puya grubbii L.B. Smith
Puya gutteana W. Weber
Puya hamata L.B. Smith
Puya harmsii (Castellanos) Castellanos
Puya herrerae Harms
Puya herzogii Wittmack

Puya hirtzii Manzanares & W. Till
Puya hofstenii Mez
Puya horrida L.B. Smith & R.W. Read
Puya hortensis L.B. Smith
Puya hoxeyi Janeba
Puya huancavelicae L.B. Smith
Puya humilis Mez
Puya hutchisonii L.B. Smith
Puya ibischii R. Vásquez
Puya iltisiana L.B. Smith
Puya isabellina Mez
Puya joergensenii H. Luther
Puya killipii Cuatrecasas
Puya kuntzeana Mez
Puya laccata Mez
Puya lanata Kunth
Puya lanuginosa (Ruiz & Pavón) Schultes f.
Puya larae R. Vásquez & Ibisch
Puya lasiopoda L.B. Smith
Puya laxa L.B. Smith
Puya lehmanniana L.B. Smith
Puya leptostachya L.B. Smith
Puya lilloi Castellanos
Puya lineata Mez
Puya llatensis L.B. Smith

Puya loca Madriñán
Puya lokischmidtiae R. Vásquez & Ibisch
Puya longisepala Mez
Puya longispina Manzanares & W. Till
Puya longistyla Mez
Puya lopezii L.B. Smith
Puya lutheri W. Till
Puya macbridei L.B. Smith
Puya macropoda L.B. Smith
Puya macrura Mez
Puya maculata L.B. Smith
Puya mariae L.B. Smith
Puya medica L.B. Smith
Puya membranacea L.B. Smith
Puya meziana Wittmack
Puya micrantha Mez
Puya mima L.B. Smith & R.W. Read
Puya minima L.B. Smith
Puya mirabilis (Mez) L.B. Smith
Puya mitis Mez
Puya mollis Baker
Puya mucronata Manzanares
Puya nana Wittmack
Puya navarroana Manzanares & W. Till
Puya nigrescens L.B. Smith

Puya nitida Mez

Puya nivalis Baker

Puya novarae Varadarajan ex Gomez Romero & A. Grau

Puya nutans L.B. Smith

Puya obconica L.B. Smith

Puya occidentalis L.B. Smith

Puya ochroleuca Betancur & Callejas

Puya olivacea Wittmack

Puya oxyantha Mez

Puya pachyphylla R. Vásquez & Ibisch

Puya parviflora L.B. Smith

Puya pattersoniae Manzanares & W. Till

Puya paupera Mez

Puya pearcei (Baker) Mez

Puya penduliflora L.B. Smith

Puya pitcairnioides L.B. Smith

Puya pizarroana R. Vásquez, Ibisch & St. Beck

Puya ponderosa L.B. Smith

Puya potosina L.B. Smith

Puya pratensis L.B. Smith

Puya prosanae Ibisch & E. Gross

Puya pseudoeryngioides H. Luther

Puya pygmaea L.B. Smith

Puya pyramidata (Ruiz & Pavón) Schultes f.

Puya raimondii Harms

Puya ramonii L.B. Smith
Puya ramosa L.B. Smith
Puya rauhii L.B. Smith
Puya reducta L.B. Smith
Puya reflexiflora Mez
Puya retrorsa Gilmartin
Puya riparia L.B. Smith
Puya robin-fosteri H. Luther
Puya roezlii E. Morren
Puya roldanii Betancur & Callejas
Puya roseana L.B. Smith
Puya rusbyi (Baker) Mez
Puya sagasteguii L.B. Smith
Puya sanctae-crucis (Baker) L.B. Smith
Puya sanctae-martae L.B. Smith
Puya santanderensis Cuatrecasas
Puya santosii Cuatrecasas
Puya secunda L.B. Smith
Puya sehuencasensis R. Vásquez, Ibisch & R. Lara
Puya serranoensis Rauh
Puya silvae-baccae L.B. Smith & R.W. Read
Puya simulans L.B. Smith
Puya smithii Castellanos
Puya sodiroana Mez
Puya solomonii G.S. Varadarajan

Puya spathacea (Grisebach) Mez
Puya stenothyrsa (Baker) Mez
Puya stipitata L.B. Smith
Puya strobilantha Mez
Puya textoragicolae W. Weber
Puya thomasiona Andre
Puya tillii Manzanares
Puya tovariana L.B. Smith
Puya trianae Baker
Puya tristis L.B. Smith
Puya trollii L.B. Smith
Puya tuberosa Mez
Puya tunarensis Mez
Puya tyleriana Sagástegui, Zapata & M.O. Dillon
Puya ugentiana L.B. Smith
Puya ultima L.B. Smith
Puya valida L.B. Smith
Puya vallo-grandensis Rauh
Puya vargasiana L.B. Smith
Puya vasquezii Ibisich & E. Gross
Puya venezuelana L.B. Smith
Puya venusta Philippi
Puya vestita André
Puya volcanensis Castillon
Puya weberbaueri Mez

Puya weberiana E. Morren ex Mez
Puya weddelliana (Baker) Mez
Puya werneriana R.W. Read & L.B. Smith
Puya westii L.B. Smith
Puya wrightii L.B. Smith
Puya wurdackii L.B. Smith
Puya x *loxensis* Manzanares & W. Till
Puya x *pichinchae* Mez & Sodiro
Puya yakespala Castellanos

Species mentioned in the present study different from *Puya* spp.

Aechmea blanchetiana (Baker) L.B.Sm.
Aechmea nudicaulis (L.) Griseb.
Agave parryi Engelm.
Agave tequilana F.A.C. Weber
Agave utahensis Engelm.
Allium x *proliferum* (Moench) Schrad. ex Willd.
Aloe vera (L.) Burm.f.
Amborella trichopoda Baill.
Ananas comosus (L.) Merr.
Atriplex prostrata Boucher ex DC.
Atriplex rosea L.
Brocchinia reducta Baker
Cinnamomum camphora (L.) J. Presl
Crassula argentea Thunb. (syn. *Crassula ovata* (Mill.) Druce.)

Denmoza rhodacantha (Salm-Dyck) Britton & Rose
Disporopsis pernyi (Hua) Diels
Echeveria quitensis (Kunth) Lindl. (syn. *Echeveria columbiana* Poelln)
Eriosyce ceratistes (Otto ex Pfeiff.) Britton & Rose (syn. *Eriosyce aurata* (Pfeiff.) Backeb)
Escobaria vivipara (Nutt.) Buxb
Espelitia shultzii Wedd.
Fascicularia bicolor (Ruiz & Pav.) Mez
Flaveria trinervia (Spreng.) C.Mohr
Guzmania lingulata (L.) Mez
Guzmania monostachia (L.) Rusby ex Mez
Kalanchoe daigremontiana Raym. Hamet & H. Perrier
Kobresia pygmaea (C.B. Clarke) C.B. Clarke
Lamiophlomis rotata (Benth. ex Hook.f.) Kudô
(syn. *Phlomoides rotata* (Benth. ex Hook.f.) Mathiesen)
Lobelia aberdarica R.E.Fr. & T.C.E.Fr.
Lobelia keniensis R.E.Fr. & T.C.E.Fr. (syn. *Lobelia gregoriana* Baker f.)
Lobelia telekii Schweinf.
Ludisia discolor (Ker Gawl.) A. Rich.
Masdevallia yungasensis T. Hashim.
Mesembryanthemum crystallinum L.
Mimulus aurantiacus Curtis
Nicotiana tabacum L.
Ochagavia carnea (Beer) L.B.Sm. & Looser
Opuntia erinacea Engelm. & J.M. Bigelow
Opuntia ficus-indica (L.) Mill.

Opuntia fragilis (Nutt.) Haw.
Opuntia humifusa (Raf.) Raf.
Oroya peruviana (K. Schum.) Britton & Rose
Peperomia camptotricha Miq.
Phalaenopsis equestris (Schauer) Rchb.f.
Phormium tenax J.R. Forst. & G. Forst.
Pinus radiata D. Don
Pitcairnia albiflos Herb.
Polylepis tarcapana Phil. (syn. *Polylepis tomentella* Wedd.)
Portulaca oleraceae L.
Potentilla saundersiana Royle
Sempervivum montanum L.
Tephrocactus floccosus (Salm-Dyck) Backeb.
 (syn. *Austrocylindropuntia floccosa* (Salm-Dyck ex Winterfeld) F.Ritter)
Vanilla planifolia Jacks. ex Andrews
Werauhia sanguinolenta (Cogn. & Marchal)
Yucca aloifolia L.
Yucca filamentosa L.
Yucca gloriosa L.
Zea mays L.

Supplementary Tables – Chapter 2

Supplementary Table 1. List of 32 climatic and topographic variables used in the present study and their source.

Climatic and topographical variables	Shapiro test <i>p</i>-value	Bartlett test <i>p</i>-value
Precipitation inter-annual variability	< 0.0001	< 0.0001
SD of aridity index	< 0.0001	< 0.0001
Aridity index	< 0.0001	< 0.0001
Precipitation of the driest month	< 0.0001	0.0003
Roughness	0.0001	0.0009
Precipitation of the coldest quarter	< 0.0001	0.0010
Cloud cover seasonality	0.0157	0.0013
Precipitation of the driest quarter	< 0.0001	0.0015
Slope	< 0.0001	0.0025
North slope orientation	0.0058	0.0082
Temperature annual range	0.0001	0.0144
Max temperature of warmest month	0.0037	0.0154
R-index	0.1515	0.0155
Annual precipitation	0.0004	0.0364
Mean temperature of warmest quarter	0.0094	0.0449
Precipitation of the wettest month	< 0.0001	0.0583
Mean temperature diurnal range	0.0709	0.0724
Precipitation of the wettest quarter	< 0.0001	0.0803
SD of R-index	0.2628	0.0997
Mean annual cloud frequency	0.0001	0.1578
Mean temperature of wettest quarter	0.0514	0.3409
Temperature seasonality	< 0.0001	0.4068
Annual mean temperature	0.4754	0.4324
Precipitation seasonality	0.0047	0.5835
Precipitation of the warmest quarter	0.0001	0.5849
Elevation	0.0665	0.6503
Mean temperature of coldest quarter	0.8594	0.7146
Min temperature of coldest month	0.0848	0.8134
East slope orientation	0.1557	0.9735
Cloud cover inter-annual variability	0.0012	0.9799
Isothermality	0.0001	0.9810
Mean temperature of driest quarter	0.3701	0.9867

Supplementary Table 2. Differences between CAM and C₃ species in their mean, median, minimum and maximum values based on bootstrap test of 100 000 independent iterations. *p*-values indicates significant differences.

Variable	Mean			Median			Maximum			Minimum		
	CAM	C ₃	<i>p</i> -value	CAM	C ₃	<i>p</i> -value	CAM	C ₃	<i>p</i> -value	CAM	C ₃	<i>p</i> -value
Cloud cover seasonality (SD %)	10.2595	7.8341	0.0005	10.165	7.875	0.0152	12.81	12.97	0.2796	8.28	2.73	< 0.0001
Aridity index (mm mm ⁻¹)	0.3902	0.8332	0.0023	0.4157	0.7469	0.03	0.5924	3.0543	< 0.0001	0.0965	0.0113	0.6479
R-index (mm mm ⁻¹)	0.3551	0.6005	0.0025	0.3766	0.5978	0.06	0.6187	0.9991	< 0.0001	0.1021	0.0079	0.6392
Precipitation seasonality (%)	8.5744	6.2694	0.025	7.9667	6	0.1388	13.1	10.9333	0.0016	5.5	1.9667	0.6199
Precipitation of the driest quarter (mm)	3.6444	8.741	0.0369	2.8	6.9	0.1593	20.2	34	0.2472	0	0	0.2877
Cloud cover inter-annual variability (SD %)	16.2656	10.0004	0.037	16.65	7.665	0.0409	30.235	33.33	0.8558	3.92	1.19	0.03
Mean annual cloud frequency (%)	60.9664	76.9255	0.0374	60.67	84.69	0.025	90.79	97.92	0.0363	24.7117	21.4633	0.5281
Precipitation of the coldest quarter (mm)	5.2153	14.5597	0.0391	3.3667	9.1	0.2067	23.8	99.6	0.0529	0	0.4	0.2051
Temperature seasonality (SD °C)	1.585	1.3014	0.0455	1.353	0.722	0.1796	4.4433	4.3378	0.2709	0.385	0.139	0.3542
SD of aridity index	0.3172	0.4896	0.0562	0.2983	0.3936	0.1937	0.5195	2.3464	0.0023	0.117	0.0167	0.9237
Max. temperature of warmest month (°C)	24.1011	20.3067	0.0644	24.5	20.5	0.137	31.2	33.1	0.4343	15.75	8.5	0.0242
Precipitation of the wettest quarter (mm)	32.1725	42.3251	0.0814	34.2	38.3667	0.2671	53.7	113.0444	0.0245	9.9667	0.7333	0.735
Annual precipitation (mm)	744.6781	1159.181	0.0933	798.8	1076.4	0.1858	1336.8	3145.0667	0.0225	178.4	12.4	0.365
Precipitation of the wettest month (mm)	12.3292	15.8287	0.1023	13.2	14.57	0.3096	19.1	42	0.0126	3.8333	0.3	0.7468
Mean temperature diurnal range (°C)	13.6575	12.1599	0.1026	13.95	12	0.1449	16.3	19.5667	0.5118	9.5667	8.0556	0.0229
Mean temperature of warmest quarter (°C)	17.1058	13.7844	0.1267	17.7	13.7429	0.1378	25.2	26.4	0.4255	8.3	3.45	0.1321
Precipitation of the warmest quarter (mm)	24.4394	31.1444	0.1275	24.7	30	0.1934	39.7	77.9	0.0778	0	0	0.5906
Mean temperature of wettest quarter (°C)	16.4222	13.0836	0.1382	16.8	12.7	0.1421	24.9	26.4	0.4542	8	2.75	0.1346
Annual mean temperature (°C)	15.3408	12.3753	0.1642	15.3	12.56	0.2685	24.8	22.1	0.1066	7.0333	3.05	0.1596
Temperature annual range (°C)	18.974	16.6896	0.2071	18.8667	16.4	0.2231	26.4	31.3	0.4012	12.3	9.475	0.0559
Mean temperature of coldest quarter (°C)	13.1198	10.5182	0.2572	13.1	10.35	0.3752	24.1	19.625	0.1067	4.7333	1.8	0.2078
Mean temperature of driest quarter (°C)	13.6694	11.1947	0.299	13.9	10.75	0.3994	24.3	20.1083	0.174	4.7667	2.9333	0.2109
Elevation (m)	2255.2694	2747.319	0.3229	2238	2969	0.2191	4049	4374	0.4901	126	63	0.6988
Precipitation of the driest month (mm)	0.9266	2.2579	0.5932	0.6	1.7	0.2304	5.6	9.9	0.9641	0	0	0.0586
Min. temperature of coldest month (°C)	5.1271	3.6171	0.6043	5.65	4.3	0.8077	18.9	14.2583	0.197	-5.5	-8.7333	0.5975
Isothermality	7.2667	7.5231	0.6215	7.2	7.7333	0.5925	8.8	9.4	0.4702	5.1333	5.2	0.7584
Precipitation inter-annual variability (mm)	3.7861	4.6113	0.6571	2.2595	2.3599	0.9682	21.3635	155	0.7469	1.8672	1.6669	0.2011
Roughness	539.0919	484.557	0.6631	467	451.25	0.7858	1423	1200	0.3137	49	141	0.8183
Slope (degrees)	9.6474	9.0831	0.8217	8.0246	8.561	0.9443	31.6212	29.6604	0.4846	0.7067	1.5537	0.5144
North slope orientation	0.0477	0.0829	0.832	-0.0183	0.0957	0.8969	0.9975	1	0.5574	-0.9914	-1	0.5833
East slope orientation	0.0072	-0.012	0.964	0.0304	-0.0128	0.9642	0.9732	0.9981	0.5482	-0.965	-0.9979	0.5661
SD of R-index	0.1262	0.1264	0.9696	0.1222	0.1316	0.8562	0.2008	0.2672	0.4283	0.0504	0.0032	0.1385

Supplementary Table 3. PCA eigenvalues of the climatic and topographic variables for PC1, PC2 and PC3.

Variable	PC1	PC2	PC3
R-index (mm mm ⁻¹)	-0.2715	0.0328	0.0609
Precipitation of the driest quarter (mm)	-0.2502	0.1111	-0.0911
Aridity index (mm mm ⁻¹)	-0.2477	0.0334	0.1721
Annual precipitation (mm)	-0.2401	0.1264	0.2133
Mean annual cloud frequency (%)	-0.2395	-0.0139	-0.0312
Precipitation of the driest month (mm)	-0.239	0.1103	-0.1245
Isothermality (mean temperature diurnal range/temperature annual range)	-0.2082	0.0472	-0.1576
Precipitation of the coldest quarter (mm)	-0.1901	0.0995	0.0817
Precipitation of the wettest quarter (mm)	-0.1791	0.1162	0.3739
Precipitation of the wettest month (mm)	-0.1776	0.1052	0.3748
Elevation (m)	-0.1431	-0.2978	0.003
Precipitation of the warmest quarter (mm)	-0.1375	0.0699	0.2809
SD of aridity index	-0.1157	0.0054	0.3437
Roughness	-0.0287	-0.0427	0.1543
Slope (Radians)	-0.0254	-0.0475	0.1441
Min. temperature of coldest month (°C)	-0.0195	0.381	-0.1063
North slope orientation	-0.0094	0.0229	-0.0225
East slope orientation	0.0101	-0.0346	0.1875
Mean temperature of coldest quarter (°C)	0.0537	0.3694	-0.0312
Precipitation inter-annual variability (mm mm ⁻¹)	0.064	0.0153	-0.1688
SD of R-index	0.0814	0.0107	0.308
Mean temperature of driest quarter (°C)	0.0837	0.3606	-0.0335
Annual mean temperature (°C)	0.1334	0.3449	0.0202
Mean temperature of wettest quarter (°C)	0.1565	0.2845	0.0552
Mean temperature of warmest quarter (°C)	0.1825	0.2976	0.0523
Mean temperature diurnal range (°C)	0.191	-0.1561	0.1813
Cloud cover seasonality (SD)	0.2056	-0.1021	0.1953
Max. temperature of warmest month (°C)	0.2108	0.2488	0.0875
Temperature seasonality (SD °C)	0.2192	-0.0546	0.1336
Temperature annual range (°C)	0.232	-0.1288	0.1941
Cloud cover inter-annual variability (SD)	0.2328	-0.0046	0.049
Precipitation seasonality (%)	0.2383	-0.0656	0.1668

Supplementary Table 4. Accuracy of five different ordinations to classify CAM and C₃ species using climatic data and random forest ordinations. The columns called “CAM” and “C₃” are the accuracy of the ordination classifying correctly CAM and C₃ species, respectively. The “Overall” column is the accuracy of the whole species. The ordination including all the variables include the variables listed in Table 2. The rest of ordinations are based on the most important variables found in Fig. 4. Ordination 2 variables: aridity index and cloud cover seasonality. Ordination 3 variables: aridity index, cloud cover seasonality, and mean annual cloud frequency. Ordination 4 variables: aridity index, cloud cover seasonality, mean annual cloud frequency and precipitation seasonality. Ordination 5 variables: aridity index, cloud cover seasonality, mean annual cloud frequency, precipitation seasonality and temperature seasonality.

Ordination	Accuracy		
	CAM	C₃	Overall
Ordination 4 variables	1.000	0.922	0.9388
Ordination 5 variables	0.968	0.905	0.9184
Ordination all variables	0.980	0.905	0.9110
Ordination 3 variables	1.000	0.879	0.9048
Ordination 2 variables	0.967	0.871	0.8912

Supplementary Tables – Chapter 3

Supplementary Table 1. Summary of the PEPC temperature optimum at saturated conditions based on the smooth function fitting a polynomial surface for each species. Data are means $\pm$ standard error of the fitting function.

Species	T_{opt} (°C)	V_{max} (mmol h⁻¹ g⁻¹ DW)
<i>Puya alpestris</i> ssp. <i>zoellneri</i>	39.80 $\pm$ 0.38	0.6525
<i>Puya castellanosi</i>	44.00 $\pm$ 0.42	0.5385
<i>Puya chilensis</i>	43.15 $\pm$ 0.48	0.4879
<i>Puya harmsii</i>	43.00 $\pm$ 0.46	0.4593

Supplementary Table 2. ANOVA results from the linear mixed-effect model based on residual PEPC activity after incubation at 50 °C, time and species. The random effect was calculated based on each individual plant.

Factor	Chi-square	Degrees of freedom	p-value
Species	18.6447	6	0.0048
Minutes	145.2795	1	< 0.0001
Species* Minutes	5.3398	6	0.501

Supplementary Table 3. ANOVA results from the linear mixed-effect model based on residual PEPC activity after incubation at 50 °C, time and species groups. Groups are defined as equatorial bromelioid species, low-elevation *Puya* and high-elevation *Puya* species. The random effect was calculated based on each individual plant.

Factor	Chi-square	Degrees of freedom	p-value
Groups	15.9271	2	0.0003
Minutes	147.9593	1	< 0.001
Groups *Minutes	4.8900	2	0.0867

Supplementary Table 4. Summary of the residual PEPC $V_{\max}$ activity (%) at 5 °C. Data are means $\pm$ standard deviation.

Species	Residual PEPC $V_{\max}$ activity (%)
<i>Ananas comosus</i>	10.88 $\pm$ 2.28
<i>Aechmea nudicaulis</i>	10.67 $\pm$ 1.48
<i>Bromelia humilis</i>	10.64 $\pm$ 0.46
<i>Puya alpestris</i>	15.01 $\pm$ 0.66
<i>Puya castellanosi</i>	17.56 $\pm$ 0.32
<i>Puya chilensis</i>	15.06 $\pm$ 1.07
<i>Puya harmsii</i>	17.95 $\pm$ 0.31

Supplementary Tables – Chapter 4

Supplementary Table 1. Summary of all the *ppc-1* genes used in the present study. In total 35 species; representing 7 orders and 13 families.

Species	Code	Source	Family	Order
<i>Agave deserti</i>	Locus3392v1rpkm88.29	Gross <i>et al.</i> (2013)	Asparagaceae	Asparagales
<i>Agave deserti</i>	Locus4288v1rpkm70.34	Gross <i>et al.</i> (2013)	Asparagaceae	Asparagales
<i>Agave deserti</i>	Locus59v1rpkm1185.23	Gross <i>et al.</i> (2013)	Asparagaceae	Asparagales
<i>Agave tequilana</i>	KXSK2015805	1 KP	Asparagaceae	Asparagales
<i>Agave tequilana</i>	KXSK2026723	1 KP	Asparagaceae	Asparagales
<i>Agave tequilana</i>	Locus1448v1rpkm157.97	Gross <i>et al.</i> (2013)	Asparagaceae	Asparagales
<i>Agave tequilana</i>	Locus346v1rpkm398.98	Gross <i>et al.</i> (2013)	Asparagaceae	Asparagales
<i>Aloe vera</i>	JVBR2007701	1 KP	Xanthorrhoeaceae	Asparagales
<i>Aloe vera</i>	JVBR2013795	1 KP	Xanthorrhoeaceae	Asparagales
<i>Amborella trichopoda</i>	URDJ2038130	1 KP	Amborellaceae	Amborellales
<i>Ananas comosus</i>	Aco010025	Ming <i>et al.</i> (2015)	Bromeliaceae	Poales
<i>Ananas comosus</i>	Aco018093	Ming <i>et al.</i> (2015)	Bromeliaceae	Poales
<i>Asparagus densiflorus</i>	FGRF2003144	1 KP	Asparagaceae	Asparagales
<i>Brocchinia reducta</i>	BYPY2003532	1 KP	Bromeliaceae	Poales
<i>Brocchinia reducta</i>	BYPY2011704	1 KP	Bromeliaceae	Poales
<i>Cinnamomum camphora</i>	BCGB2004675	1 KP	Canellaceae	Canellales
<i>Disporopsis pernyi</i>	UZXL2028496	1 KP	Asparagaceae	Asparagales
<i>Disporopsis pernyi</i>	UZXL2028497	1 KP	Asparagaceae	Asparagales
<i>Drakaea elastica</i>	XZME2000153	1 KP	Orchidaceae	Asparagales
<i>Elaeis guineensis</i>	XP10924612	Phytozome	Arecaceae	Arecales
<i>Goodyera pubescens</i>	VTUS2011331	1 KP	Orchidaceae	Asparagales
<i>Ludisia discolor</i>	LELS2002838	1 KP	Orchidaceae	Asparagales
<i>Ludisia discolor</i>	LELS2013941	1 KP	Orchidaceae	Asparagales
<i>Masdevallia yuagensis</i>	JSAG2080449	1 KP	Orchidaceae	Asparagales
<i>Masdevallia yuagensis</i>	JSAG2080458	1 KP	Orchidaceae	Asparagales
<i>Musa accuminata</i>	Achr9T06420	Phytozome	Musaceae	Zingiberales
<i>Oryza sativa</i>	Os01t0758300-01	Phytozome	Poaceae	Poales
<i>Oryza sativa</i>	Os08t0366000-01	Phytozome	Poaceae	Poales
<i>Phalaenopsis equestris</i>	PEQU07008	Cai <i>et al.</i> (2014)	Orchidaceae	Asparagales
<i>Phalaenopsis equestris</i>	PEQU14315	Cai <i>et al.</i> (2014)	Orchidaceae	Asparagales
<i>Phormium tenax</i>	FCEL2009430	1 KP	Xanthorrhoeaceae	Asparagales
<i>Phormium tenax</i>	FCEL2021170	1 KP	Xanthorrhoeaceae	Asparagales
<i>Pinus radiata</i>	DZQM2011359	1 KP	Pinaceae	Pinales
<i>Pistia stratiotes</i>	MFIN2004570	1 KP	Araceae	Alismatales
<i>Pistia stratiotes</i>	MFIN2009548	1 KP	Araceae	Alismatales
<i>Pitcairnia albiflos</i>	comp30390 c0seq1	Palma-Silva <i>et al.</i> (2016)	Bromeliaceae	Poales
<i>Pitcairnia albiflos</i>	comp30390 c0seq2	Palma-Silva <i>et al.</i> (2016)	Bromeliaceae	Poales
<i>Platanthera clavellata</i>	MTHW2054712	1 KP	Orchidaceae	Asparagales
<i>Puya alpestris ssp. alpestris</i>	<i>ppc-1a</i>	The present study	Bromeliaceae	Poales
<i>Puya alpestris ssp. alpestris</i>	<i>ppc-1b</i>	The present study	Bromeliaceae	Poales
<i>Puya alpestris ssp. zoellneri</i>	<i>ppc-1a</i>	The present study	Bromeliaceae	Poales
<i>Puya alpestris ssp. zoellneri</i>	<i>ppc-1b</i>	The present study	Bromeliaceae	Poales
<i>Puya chilensis</i>	<i>ppc-1a</i>	The present study	Bromeliaceae	Poales
<i>Puya chilensis</i>	<i>ppc-1b</i>	The present study	Bromeliaceae	Poales
<i>Puya coerulea</i>	<i>ppc-1a</i>	The present study	Bromeliaceae	Poales
<i>Puya coerulea</i>	<i>ppc-1b</i>	The present study	Bromeliaceae	Poales
<i>Puya ferruginea</i>	<i>ppc-1a</i>	The present study	Bromeliaceae	Poales
<i>Puya ferruginea</i>	<i>ppc-1b</i>	The present study	Bromeliaceae	Poales
<i>Puya mirabilis</i>	<i>ppc-1a</i>	The present study	Bromeliaceae	Poales
<i>Puya mirabilis</i>	<i>ppc-1b</i>	The present study	Bromeliaceae	Poales
<i>Puya raimondii</i>	<i>ppc-1a</i>	The present study	Bromeliaceae	Poales
<i>Puya raimondii</i>	<i>ppc-1b</i>	The present study	Bromeliaceae	Poales
<i>Puya spathacea</i>	<i>ppc-1b</i>	The present study	Bromeliaceae	Poales
<i>Puya spathacea</i>	<i>ppc-1a</i>	The present study	Bromeliaceae	Poales
<i>Spirodela polyrrhyza</i>	Spipo19G0008900	Phytozome	Araceae	Alismatales
<i>Typha latifolia</i>	BRUD2000428	1 KP	Thypaceae	Poales
<i>Typha latifolia</i>	BRUD2011009	1 KP	Thypaceae	Poales
<i>Vanilla planifolia</i>	THDM2007603	1 KP	Orchidaceae	Asparagales
<i>Vanilla planifolia</i>	THDM2009257	1 KP	Orchidaceae	Asparagales
<i>Zingiber officinale</i>	BDJQ2004809	1 KP	Zingiberaceae	Zingiberales

Supplementary Table 2. Isoforms of 25 genes involved in CAM photosynthesis in the eight species of *Puya*.

Gene name	<i>Ananas comosus</i> genes	<i>Puya alpestris</i> ssp. <i>zoellneri</i> genes	<i>Puya alpestris</i> genes	<i>Puya chilensis</i> genes	<i>Puya mirabilis</i> genes	<i>Puya raimondii</i> genes	<i>Puya spathacea</i> genes	<i>Puya coerulea</i> genes	<i>Puya ferruginea</i> genes
Aluminium-activated malate transporter	Aco003023.1	c11269_g1_i1	c10892_g1_i1	c16906_g1_i2, c16906_g1_i4, c16906_g1_i1, c16906_g2_i1, c16906_g1_i3	c7291_g2_i1, c7291_g2_i2	c11528_g1_i3, c11528_g1_i2, c11528_g1_i1	c21281_g1_i6, c21281_g1_i3, c21281_g1_i1, c21281_g1_i4, c21281_g1_i2, c21281_g1_i7, c21281_g1_i5	c19209_g1_i1, c19209_g1_i2, c19209_g1_i3	c24017_g2_i1
Aluminium-activated malate transporter	Aco010725.1	c21476_g1_i1, c21476_g1_i4, c21476_g1_i3, c21476_g1_i2, c21476_g2_i1	c16381_g1_i1, c16381_g1_i3, c16381_g1_i2, c2315_g1_i1	c18548_g1_i1, c18548_g1_i2, c18548_g1_i3, c18548_g1_i4, c18548_g1_i5, c18548_g1_i6, c18548_g1_i7, c18548_g2_i1	c4563_g1_i2, c4563_g1_i1	c10086_g1_i1, c17199_g1_i1	c17803_g1_i1, c17803_g1_i2	c21175_g1_i1, c21175_g1_i2, c21175_g1_i3, c21175_g1_i4, c21175_g1_i5, c21175_g1_i6, c23390_g1_i1	c27023_g1_i1, c35282_g1_i1, c29244_g1_i1
beta carbonic anhydrase 2	Aco005402.1	c19949_g1_i1, c19949_g1_i2, c19949_g1_i3	c13863_g1_i4, c13863_g1_i2, c13863_g1_i3, c13863_g1_i1	c18396_g1_i1, c18396_g1_i2, c18396_g1_i3, c18396_g1_i4, c18396_g1_i5, c18396_g1_i6, c18396_g1_i7, c18396_g1_i8, c18396_g1_i9, c18396_g1_i10	c10938_g1_i1, c10938_g1_i2, c10938_g1_i3, c10938_g1_i4	c21737_g1_i1	c18362_g1_i1, c18362_g1_i2	c13509_g1_i1, c13509_g1_i2	c20591_g1_i2, c20591_g1_i1
beta carbonic anhydrase 2	Aco006181.1	c13543_g1_i2, c13543_g1_i1	c14180_g1_i2, c14180_g1_i1	c36101_g1_i1	c171_g1_i1	c7676_g1_i1	c36516_g1_i1	c3073_g1_i1	c13111_g1_i1
beta carbonic anhydrase 5	Aco002732.1	c21152_g1_i1	c6769_g1_i1	c3830_g1_i1	c11540_g1_i1, c11540_g1_i2, c11540_g1_i3	c4769_g1_i1	c18688_g1_i2, c18688_g1_i1	c23124_g1_i2, c23124_g1_i1, c23124_g1_i3	c21020_g1_i2, c21020_g1_i1, c21020_g1_i3
Malate dehydrogenase	Aco002883.1	-	-	-	-	-	-	-	-
Malate dehydrogenase	Aco004349.1	c16081_g1_i1, c21039_g2_i1	c10912_g1_i1, c20069_g1_i1	c12831_g1_i1, c12831_g2_i1, c22695_g1_i1, c258_g1_i2, c258_g1_i1	c22063_g1_i1, c7758_g1_i1	c17390_g1_i1, c7777_g1_i1	c29555_g1_i1, c20800_g1_i1, c20800_g1_i2	c20675_g1_i1, c20675_g1_i2, c20675_g1_i3, c20675_g1_i4, c20675_g1_i5	c16330_g1_i1, c19981_g1_i1
Malate dehydrogenase	Aco004996.1	c1369_g1_i1	c15835_g1_i1	c23351_g1_i1	c30197_g1_i1	c13011_g1_i1	c17789_g1_i1	c20810_g1_i1	c19943_g1_i1
Malate dehydrogenase	Aco006122.1	c21039_g2_i1	c20069_g1_i1	c12831_g2_i1, c12831_g1_i1	c7758_g1_i1	c7777_g1_i1, c8293_g1_i5, c8293_g1_i4, c8293_g1_i3, c8293_g1_i2, c8293_g1_i1	c20800_g1_i1, c20800_g1_i2	c20675_g1_i5, c20675_g1_i4, c20675_g1_i3, c20675_g1_i2, c20675_g1_i1	c19981_g1_i1, c17046_g1_i1
Malate dehydrogenase	Aco007734.1	c8710_g1_i1, c8710_g2_i1	c14467_g1_i1, c14467_g1_i2, c14467_g1_i3, c10999_g1_i1	c7864_g1_i3, c7864_g1_i1, c7864_g1_i2	c13460_g1_i1, c13763_g1_i1, c30272_g1_i1, c1879_g1_i1, c25788_g1_i1	c9648_g1_i3, c9648_g1_i6, c9648_g1_i4, c9648_g1_i1	c20947_g1_i2, c20947_g1_i1, c15603_g1_i1	c19750_g1_i5, c19750_g1_i4, c19750_g1_i3, c19750_g1_i2, c19750_g1_i1	c24603_g2_i1, c24603_g1_i1, c24603_g1_i2
Malate dehydrogenase	Aco008626.1	c26198_g1_i1	c15705_g1_i2, c15705_g1_i1	c8213_g2_i1	c26021_g1_i1	c98_g1_i1	c19409_g1_i1, c22479_g3_i3	c16036_g1_i1, c9993_g1_i1	c12283_g1_i1
Malate dehydrogenase	Aco010232.1	c18149_g1_i2, c18149_g1_i3, c18149_g1_i1	c16165_g1_i1, c16165_g1_i2	c21290_g1_i1, c21290_g1_i3, c21290_g1_i2	c34821_g1_i1	c13011_g1_i1	c17072_g1_i1	c23875_g1_i1, c23875_g1_i2	c21315_g1_i1
Malate dehydrogenase	Aco013933.1	c8710_g2_i1, c8710_g1_i1	c14467_g1_i1, c14467_g1_i2, c14467_g1_i3	c7864_g1_i1, c7864_g1_i2, c7864_g1_i3	c13460_g1_i1	c9648_g1_i1, c9648_g1_i3, c9648_g1_i4, c9648_g1_i6	c20947_g1_i1, c20947_g1_i2	c19750_g1_i1, c19750_g1_i2, c19750_g1_i3, c19750_g1_i4, c19750_g1_i5	c24603_g1_i2, c24603_g1_i1, c24603_g2_i1
Phosphoenolpyruvate carboxylase kinase 1	Aco010093.1	c30798_g1_i1	c3480_g1_i1, c3480_g1_i2, c642_g1_i1	c15834_g1_i2, c15834_g1_i1, c3046_g1_i1	c6625_g1_i3, c6625_g1_i2, c6625_g1_i4	c38083_g1_i1, c13753_g1_i1	c6903_g1_i1, c28313_g1_i1	c36897_g1_i1, c10423_g1_i2, c10423_g1_i1, c24364_g2_i3, c24364_g2_i1, c24364_g2_i2	c26857_g1_i1, c18443_g1_i2, c18443_g1_i1, c14360_g1_i3, c14360_g1_i2, c14360_g1_i1
Vacuolar sucrose transporter	Aco015779.1	c21998_g1_i2, c21998_g1_i3	c19779_g1_i1, c19779_g1_i2	c31779_g1_i1, c1792_g1_i2, c1792_g1_i1	c2395_g1_i1, c2395_g1_i2	c3066_g1_i1, c3066_g1_i2	c22113_g1_i1, c12606_g1_i2, c12606_g1_i1	c23430_g1_i1, c23430_g1_i2	c23410_g1_i1, c1088_g1_i1

Gene name	<i>Ananas comosus</i> genes	<i>Puya alpestris ssp. zoellneri</i> genes	<i>Puya alpestris</i> genes	<i>Puya chilensis</i> genes	<i>Puya mirabilis</i> genes	<i>Puya raimondii</i> genes	<i>Puya spathacea</i> genes	<i>Puya coerules</i> genes	<i>Puya ferruginea</i> genes
Malate dehydrogenase	Aco014690.1	c21741_g1_i1	c17638_g1_i1, c27311_g1_i1	c18686_g1_i1	c13622_g1_i1	c35105_g1_i1	c20078_g1_i1	c23109_g1_i1	c24037_g1_i1
Malate dehydrogenase	Aco017525.1	c18267_g1_i2, c18267_g1_i1	c21379_g1_i1	c21106_g1_i1	c22589_g1_i1	c12626_g1_i2, c12626_g1_i1	c21835_g1_i2, c21835_g1_i1	c3136_g1_i1, c29455_g1_i1	c23520_g1_i1, c23520_g1_i2
Malate dehydrogenase	Aco017526.1	c18267_g1_i2, c18267_g1_i1	c21379_g1_i1	c21106_g1_i1	c22589_g1_i1	c12626_g1_i2, c12626_g1_i1	c21835_g1_i2, c21835_g1_i1	c3136_g1_i1, c29455_g1_i1	c23520_g1_i1, c23520_g1_i2
Malate dehydrogenase	Aco017527.1	c18267_g1_i2, c18267_g1_i1	c21379_g1_i1	c21106_g1_i1	c22589_g1_i1	c12626_g1_i2, c12626_g1_i1	c21835_g1_i2, c21835_g1_i1	c3136_g1_i1, c29455_g1_i1	c23520_g1_i1, c23520_g1_i2
Malate dehydrogenase	Aco017528.1	c18267_g1_i2, c18267_g1_i1	c21379_g1_i1	c21106_g1_i1	c22589_g1_i1	c12626_g1_i2, c12626_g1_i1	c21835_g1_i2, c21835_g1_i1	c3136_g1_i1, c29455_g1_i1	c23520_g1_i1, c23520_g1_i2
Malate dehydrogenase	Aco019631.1	-	c29130_g1_i1	-	c35233_g1_i1, c31228_g1_i1	c22304_g1_i1	c43104_g1_i1	c30672_g1_i1	c29496_g1_i1
NAD-dependent malic enzyme	Aco016569.1	c22349_g1_i1	c18607_g1_i1	c10253_g1_i2, c10253_g1_i1	c10239_g1_i2, c10239_g1_i1	c7368_g1_i1	c21826_g1_i1, c21826_g1_i2, c21826_g1_i3, c21826_g1_i4, c21826_g1_i5, c21826_g1_i6, c21826_g1_i7, c21826_g1_i8	c22719_g1_i1, c22719_g1_i2, c22719_g1_i3, c22719_g1_i4, c22719_g1_i5	c25838_g3_i2, c25838_g3_i1, c25838_g2_i1
NAD-dependent malic enzyme 2	Aco007622.1	c22550_g1_i1	c18541_g1_i1	c8263_g1_i1	c11506_g1_i1	c7721_g1_i2, c7721_g1_i1	c26318_g1_i1	c22267_g2_i1	c23342_g1_i1
NADP-malic enzyme 1	Aco009967.1	c13752_g1_i1	c15202_g1_i1	c18426_g1_i1	c7236_g1_i1, c7236_g1_i2	c16894_g1_i1	c16768_g1_i1	c6814_g1_i3, c6814_g1_i1, c6814_g1_i2	c24222_g2_i1
NADP-malic enzyme 3	Aco005989.1	c12399_g1_i1	c15816_g1_i1	c19447_g1_i1	c34322_g1_i1	c21549_g1_i1, c15422_g1_i1	c40251_g1_i1	c16399_g1_i1, c33375_g1_i1, c33994_g1_i1	c24222_g1_i1
NADP-malic enzyme 4	Aco005631.1	c12399_g1_i1, c7025_g1_i2, c7025_g1_i1	c15816_g1_i1	c19447_g1_i1, c8383_g1_i1, c8383_g1_i2	c34322_g1_i1, c3081_g1_i1, c3081_g1_i2	c21549_g1_i1, c15422_g1_i1, c16894_g1_i1	c40251_g1_i1	c16399_g1_i1	c24222_g1_i1, c24222_g2_i1
Phosphoenolpyruvate carboxykinase 1	Aco017762.1	c21746_g2_i2, c21746_g2_i3, c21746_g2_i4, c21746_g2_i1	c8930_g1_i1	c14920_g1_i1, c14920_g1_i2, c14920_g1_i3	c9259_g1_i1, c9259_g1_i2	c9835_g1_i1, c9835_g1_i2, c9835_g1_i3, c9835_g1_i4, c9835_g1_i5, c9835_g1_i6	c8320_g1_i1, c8320_g1_i2, c8320_g1_i3	c20156_g1_i1, c20156_g1_i2, c20156_g1_i3	c20298_g1_i1, c7495_g1_i2, c7495_g1_i1, c15161_g1_i1
phosphoenolpyruvate carboxylase 3	Aco010025.1	c12362_g1_i1	c10999_g1_i1	c16560_g1_i1	c25547_g1_i1	c7709_g1_i1, c7709_g1_i2	c10345_g1_i1	c22339_g1_i1	c24427_g2_i1
phosphoenolpyruvate carboxylase 3	Aco018093.1	c20452_g2_i1, c12362_g1_i1, c28967_g1_i1, c20452_g1_i1	c5947_g1_i1, c10999_g1_i1, c9767_g1_i1, c4518_g1_i1	c18980_g1_i1, c16560_g1_i1, c12461_g1_i1, c9045_g1_i1	c21517_g1_i1, c25547_g1_i1, c22549_g1_i1, c30143_g1_i1	c30586_g1_i1, c7709_g1_i1, c7709_g1_i2, c16757_g1_i1	c17537_g1_i1, c10345_g1_i1, c16022_g2_i1, c9906_g1_i1, c16022_g1_i1	c1701_g1_i1, c22339_g1_i1, c10207_g2_i1, c9163_g1_i1, c10207_g1_i1, c6813_g1_i2, c6813_g1_i1	c17316_g1_i1, c24427_g2_i1, c29156_g1_i1, c9103_g1_i1
phosphoenolpyruvate carboxylase 4	Aco016429.1	c22224_g1_i1	c9210_g1_i1, c980_g1_i1	c27924_g1_i1	c30926_g1_i1	c13693_g1_i1	c20250_g1_i1	c22976_g1_i1	c24361_g1_i1

Gene name	<i>Ananas comosus</i> genes	<i>Paya alpestris</i> ssp. <i>zoellneri</i> genes	<i>Paya alpestris</i> genes	<i>Paya chilensis</i> genes	<i>Paya mirabilis</i> genes	<i>Paya raimondii</i> genes	<i>Paya spathacea</i> genes	<i>Paya coerulea</i> genes	<i>Paya ferruginea</i> genes
Phosphoenolpyruvate carboxylase kinase 1	Aco013938.1	c11031_g1_i2, c11031_g1_i3	c3480_g1_i2, c3480_g1_i1	c10074_g1_i3, c10074_g1_i2, c10074_g1_i1, c15834_g1_i2, c15834_g1_i1	c6625_g1_i4, c26135_g1_i1, c6625_g1_i1, c6625_g1_i2, c6625_g1_i3	c7607_g1_i3, c7607_g1_i2	c14082_g1_i2, c14082_g1_i3, c14082_g1_i1, c28313_g1_i1	c24364_g2_i2, c24364_g2_i3, c24364_g2_i1	c14360_g1_i1, c14360_g1_i3, c14360_g1_i2
Phosphoenolpyruvate carboxylase-related kinase 1	Aco022525.1	c18840_g1_i3, c18840_g1_i2, c18840_g1_i1	c7090_g1_i2, c7090_g1_i1, c7090_g1_i3	c23112_g2_i9, c23112_g2_i4, c23112_g2_i3, c23112_g2_i2, c23112_g2_i6, c23112_g2_i5, c23112_g2_i12, c23112_g2_i8	c9653_g1_i3, c9653_g1_i2, c9653_g1_i1, c9653_g1_i4	c30435_g1_i1	c13645_g1_i3, c13645_g1_i2, c13645_g1_i1	c24400_g1_i15, c24400_g1_i6, c24400_g1_i4, c24400_g1_i9	c23349_g1_i1, c23349_g1_i3, c23349_g1_i2
Phosphoenolpyruvate carboxylase-related kinase 2	Aco001261.1	c19714_g2_i1, c19714_g2_i2	c14921_g1_i1	c12715_g1_i2, c12715_g1_i1, c12715_g1_i3	c9001_g1_i1, c9001_g1_i2, c8281_g1_i1	c10840_g1_i6, c10840_g1_i2, c10840_g1_i1, c10840_g1_i3, c10840_g1_i4, c10840_g1_i5	c13616_g1_i5, c13616_g1_i1, c13616_g1_i2, c13616_g1_i4, c13616_g1_i3	c11705_g1_i1, c11705_g1_i2, c31830_g1_i1	c20282_g1_i3, c20282_g1_i5, c20282_g1_i1, c20282_g1_i6, c20282_g1_i4
PPDK regulatory protein	Aco014488.1	c20954_g2_i1, c20954_g2_i3, c25588_g1_i1	c13746_g1_i1, c13746_g1_i2, c13746_g1_i3, c13746_g1_i4	c19831_g1_i3, c32118_g1_i1	c6070_g1_i1, c6070_g1_i2	c6923_g1_i1, c6923_g1_i2	c16133_g1_i1, c16133_g1_i2, c33318_g1_i1, c31856_g1_i1, c25175_g1_i1	c18963_g1_i1, c18963_g1_i2	c25629_g1_i1, c25629_g1_i2
Pyrophosphate:fructose-6-phosphate 1-phosphotransferase	Aco017133.1	c23261_g1_i1	c6743_g1_i1, c6743_g1_i2	c22465_g1_i1, c12213_g1_i1	c8380_g1_i1	c3893_g1_i1, c7815_g1_i1	c19334_g1_i1, c19334_g1_i2, c19334_g1_i3	c24320_g1_i1	c12959_g1_i1, c9606_g1_i1
Pyrophosphate:fructose-6-phosphate 1-phosphotransferase	Aco018526.1	c20458_g1_i1, c20458_g1_i2, c37922_g1_i1, c14703_g3_i1	c35140_g1_i1, c35859_g1_i1, c30683_g1_i1, c3657_g1_i1, c11667_g1_i1	c21049_g1_i2, c12618_g1_i1, c12618_g1_i2, c21049_g1_i1, c47454_g1_i1, c23091_g2_i2, c23091_g2_i1, c23091_g3_i1	c7791_g1_i1, c6795_g1_i1, c18959_g1_i1, c7791_g1_i2, c6795_g2_i1, c22344_g1_i1	c30590_g1_i1, c6639_g1_i2, c4152_g1_i1, c14969_g1_i2, c37095_g1_i1, c4152_g1_i2, c12872_g1_i1, c31427_g1_i1, c6639_g1_i1	c14969_g1_i3, c6632_g1_i1, c14969_g1_i2, c14969_g1_i1, c12872_g1_i1, c5085_g1_i1, c13962_g1_i1	c7744_g2_i2, c7744_g2_i1, c17470_g1_i3, c17470_g1_i1, c17470_g1_i2, c28041_g1_i1, c9597_g1_i1, c7744_g1_i1	c24608_g1_i1, c34131_g1_i1, c9428_g1_i1, c9957_g1_i1
Pyruvate, orthophosphate dikinase	Aco024818.1	c17710_g1_i1	c17091_g1_i1	c41455_g1_i1	c34907_g1_i1	c17040_g1_i1	c17992_g1_i1	c22109_g1_i1	c19394_g1_i1
Tonoplast decarboxylate transporter	Aco000795.1	-	c28062_g1_i1	c17691_g1_i2, c17691_g1_i1	c7920_g1_i1, c7920_g1_i2	c4972_g1_i1, c23403_g1_i1, c15404_g1_i1, c4972_g1_i2	-	c37895_g1_i1, c36769_g1_i1, c33427_g1_i1	-
V-ATPase A subunit	Aco008300.1	c13538_g1_i1	c14890_g1_i1, c14890_g1_i2	c36172_g1_i1	c29998_g1_i1	c21307_g1_i1	c1353_g1_i1	c19562_g1_i1	c18194_g1_i1
V-ATPase B subunit	Aco015226.1	c18871_g1_i1	c15625_g1_i1	c36217_g1_i1	c34501_g1_i1	c30485_g1_i1	c9874_g1_i1	c19253_g2_i1	c19328_g1_i1
V-PPiase (K ⁺ -dependent)	Aco008167.1	c22888_g1_i1, c22888_g1_i2	c19124_g1_i1, c19124_g2_i1	c17422_g1_i1, c17422_g1_i2, c15329_g1_i1	c5900_g2_i1	c7675_g1_i1, c7675_g1_i2, c3865_g1_i1	c19318_g1_i1, c19318_g1_i2, c19318_g1_i3	c22170_g1_i1, c22170_g1_i2, c22170_g1_i3, c22170_g3_i1, c33366_g1_i1	c25520_g1_i1, c25520_g1_i2, c8369_g1_i1
Vacuolar hexose transporter	Aco005379.1	c22384_g1_i1	c1425_g1_i1, c19789_g2_i2	c22905_g2_i1, c8682_g1_i2, c8682_g1_i1	c13188_g4_i1, c13188_g7_i1, c13188_g3_i2, c13188_g3_i1	c12364_g4_i1, c12364_g3_i1	c19193_g2_i1, c19193_g1_i1	c18627_g2_i1, c18627_g1_i1	c22164_g1_i1
Vacuolar invertases	Aco017533.1	c23713_g3_i1, c23713_g3_i2, c23713_g3_i3	c20123_g3_i1, c20123_g3_i2, c20123_g3_i3, c20123_g3_i4	c21145_g1_i1, c21145_g1_i2, c21145_g1_i3, c21145_g1_i4, c21145_g1_i5, c21145_g1_i6, c21145_g1_i7	c9789_g1_i1, c9789_g1_i2	c11398_g2_i1, c11398_g2_i2, c11398_g2_i3, c11398_g2_i4, c11398_g2_i5, c11398_g2_i6	c11919_g1_i1, c11919_g1_i2, c11919_g1_i3	c23715_g1_i1, c23715_g1_i2, c23715_g1_i3, c23715_g1_i4	c24639_g1_i1, c24639_g1_i2
Vacuolar invertases	Aco023030.1	c21391_g1_i1, c21391_g1_i2, c21391_g1_i3	c16068_g1_i1, c16068_g2_i1, c16068_g2_i2	c20483_g1_i1	c31174_g1_i1	c7414_g1_i1, c7414_g1_i2, c7414_g1_i3	c24055_g1_i1, c22581_g1_i1, c35757_g1_i1	c17401_g1_i1	c20662_g1_i1, c20662_g1_i2
Vacuolar sucrose transporter	Aco011916.1	c21998_g1_i2, c21998_g1_i1, c21998_g1_i3	c14055_g1_i1	c21145_g1_i1, c21145_g1_i2, c21145_g1_i3, c21145_g1_i4	c7339_g1_i1, c7339_g1_i2, c7339_g1_i3	c7451_g1_i1, c12880_g1_i1	c21631_g1_i1, c21631_g1_i2, c21631_g1_i3, c21631_g1_i4	c23668_g1_i1, c23668_g1_i2, c23668_g1_i3	c25402_g1_i1, c25402_g1_i2

Supplementary Material

Supplementary Material – Chapter 4

Supplementary Material 1 – RNA extraction protocol for species with high polysaccharides and phenols after Gehrig *et al.* (2000)

Materials

For the solutions **always** use water treated with diethyl pyrocarbonate (DEPC), 0.1 % (v/v) incubated for at least 2 h (ideally overnight) and then autoclaved to inactivate DEPC.

Grind the tissue in a pestle and mortar using liquid nitrogen. **Do not let the tissue thaw.**

Work always in the fume hood and keep your samples on ice.

SDS buffer

2% (w/v) Cetyl trimethylammonium bromide (CTAB)

0.1 M Tris-HCl (pH 8.0)

1.4 M NaCl

4% (w/v) Polyethylene glycol (PEG; high-molecular weight)

20 mM Ethylenediaminetetraacetic acid (EDTA)

1% (w/v) Dimethyl-carbamic acid DECA

4% (w/v) L-Ascorbic acid

5% (v/v) β -mercaptoethanol

Trizol buffer

0.5% (w/v) *N*-lauryl-sarcosine

Trizol®

Other reagents

SEVAG (Chloroform: isoamyl alcohol (24:1, v/v))

Ethanol (97%, v/v)

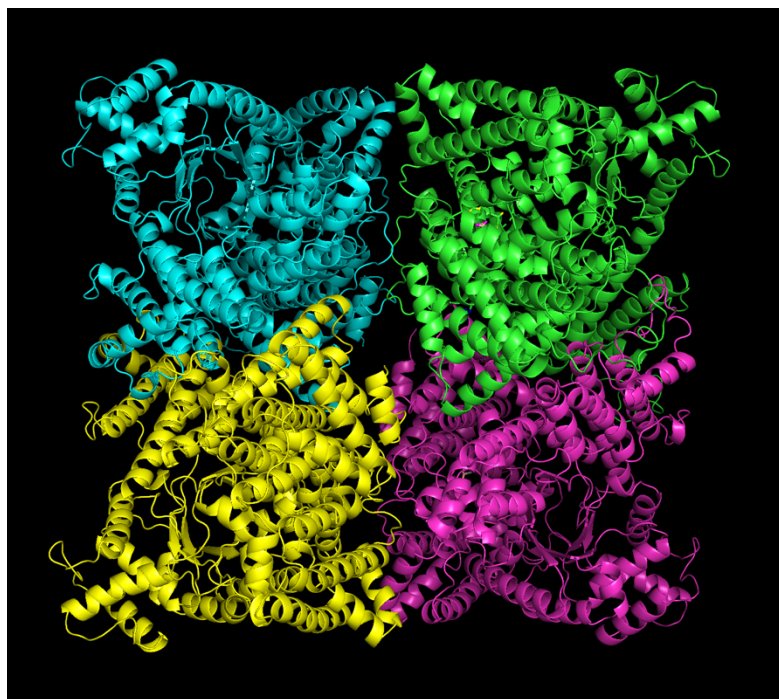
Isopropanol

Protocol

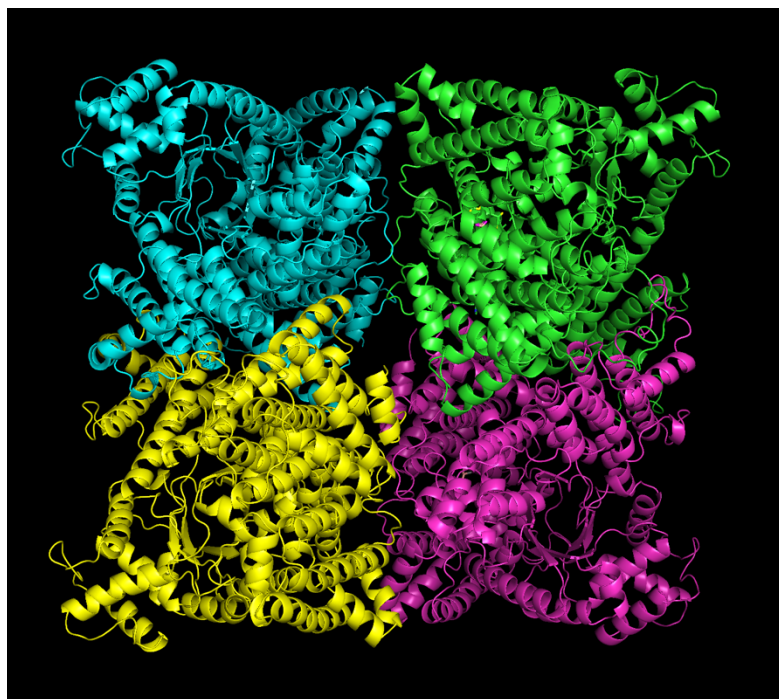
- 1)** Grind 100 mg of fresh tissue in a pestle and mortar previously cooled with liquid nitrogen.
- 2)** In a 2 ml Eppendorf tube add 450 μ l of CTAB buffer and mix well (vortex for 20 s), incubate for **10 min at 60°C**. (Reduce the time if the samples are rich in starch.) If you have more than one sample after adding the CTAB buffer, rest the tubes on ice until you finish adding all the buffer so as to guarantee that each tube has the same incubation time.
- 3)** Centrifuge at $15\,000 \times g$ for 10 min at 4°C.
- 4)** Recover the supernatant (this should be clear and not viscous) in a new tube and add 500 μ l SEVAG.
- 5)** Centrifuge at $13\,000 \times g$ for 10 min at 4°C.
- 6)** Recover the supernatant) in a new tube and add 500 μ l of Trizol buffer and 500 μ l of SEVAG.
- 7)** Centrifuge at $13\,000 \times g$ for 10 min at 4°C.
- 8)** Recover the supernatant (if this is not clear, add 900 μ l of SEVAG and repeat step 7)
- 9)** Add at least 500 μ l of chilled isopropanol, if it is possible fill up the tube with isopropanol.
- 10)** Incubate for 10 min at -20°C and spin for $12\,000 \times g$ for 20 min at 4°C.
- 11)** Pour off the supernatant and wash the pellet with 1 ml of 97 % (v/v) ethanol; vortex until pellet does not stick any more.
- 12)** Centrifuge at $9\,000 \times g$ for 10 min at 4°C.
- 13)** Pour off the ethanol (if the pellet does not stick on the bottom of the tube, the ethanol should be removed by pipetting very carefully).
- 14)** Centrifuge the tube with the pellet for 1 min at $15\,000 \times g$ at 4°C and pipette off the excess of ethanol.
- 15)** Let the tube dry for 2 min at room temperature.
- 16)** Add 20–50 μ l of DEPC water.
- 17)** Proceed with DNase protocol.

Supplementary Material 2 – Effect of changes in the residue 679 in the CAM isoform of *Ananas comosus* from arginine to histidine, and possible functional implications of CAM polymorphism Arg679.

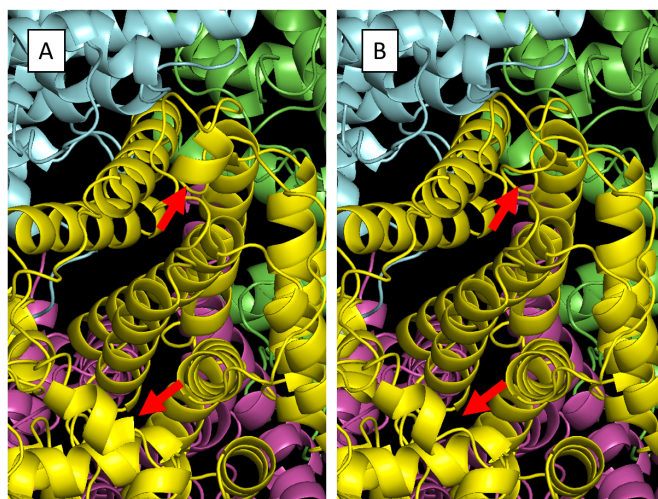
Once the polymorphism at the residue 679 of PEPC (*Ananas comosus* sequence position) was discovered in Poales and Asparagales, the next question was whether that polymorphism has a functionality, for example produces kinetic differences between the CAM and C₃ isoforms. Based on functional polymorphisms in bacteria and C₄ PEPC isoforms, one of the most common kinetic differences associated such substitutions is a change in sensitivity of the enzyme by malate (Takeya *et al.*, 2017). A 3D-model of the *Ananas comosus* PEPC CAM (Arg679) and the mutated protein *Ananas comosus* PEPC CAM (His679) were generated to understand whether there is a predicted structural difference between the two proteins. The theoretical 3D models were constructed on SWISS-MODEL SERVER (Biasini *et al.*, 2014) using the amino acid sequence Aco010025 from Ming *et al.* (2015). Both models were predicted as a homo-tetrameric protein with bound aspartate ligand; this is because the majority of X-ray structures used aspartate to investigate the binding of this allosteric inhibitor (Paulus *et al.*, 2013). The *Ananas comosus* PEPC CAM isoform model and the *Ananas comosus* CAM-mutated (His679) model are described in Figs. 1 and 2, respectively. The small differences between the models are shown in Fig. 3.



Supplementary Material 2 – Figure 1. *Ananas comosus* phosphoenolpyruvate carboxylase (PEPC) CAM isoform model. Theoretical model of the PEPC CAM isoform of *Ananas comosus*, using the amino acid sequence on SWISS-MODEL SERVER. The model predicted a homotetramer with bound aspartate ligand. Each different colour represents a single subunit of 965 amino acids.



Supplementary Material 2 – Figure 2. *Ananas comosus* CAM-Mutated (His679) model. Theoretical model of the PEPC CAM isoform of *Ananas comosus*, using the amino acid sequence on SWISS-MODEL SERVER. The residue Arg679 was modified to His679 before the model simulation. The model predicted a homo-tetramer with an aspartate ligand. Each colour represents the PEPC monomer, 965 amino acids.



Supplementary Material 2 – Figure 3. *Ananas comosus* CAM-mutated (His679) model. Theoretical model of the PEPC CAM isoform of *Ananas comosus*, using the amino acid sequence on SWISS-MODEL SERVER. The residue Arg679 was changed to His679 before the model simulation. The model predicted a homotetramer with a bound aspartate ligand. Each different colour represents a single subunit of 965 amino acids.

The functional residues in the PEPC C₄ isoform of maize (Kai *et al.*, 2003) and in the C₄ isoform of *Flaveria trinervia* (Paulus *et al.*, 2013) were identified in the *Ananas comosus* CAM isoform (Table 1). The sequences of the C₄ isoform of maize and *Flaveria trinervia* were retrieved from the Protein Data Bank. The distance between the residue Arg679 (Fig. 1) or His679 (Fig. 2) located in the monomer A and the residues identified in Table 1 in the monomer A, B, C, and D were measured using the *get_distance* function in PyMol 2.0 (Schrödinger). A summary of the different distances between residues in the *Ananas comosus* CAM isoform model and the *Ananas comosus* CAM-mutated (His679) model are described in Table 2 and Table 3, respectively. Table 4 showed a summary of the differences in distance between the two models.

Supplementary Material 2 – Table 1. Functional residues of PEPC based on the numbering of maize C₄ isoform, pineapple CAM isoform, and *Flaveria trinerva* C₄ isoform. Functional residues that differ in the three species are highlighted in grey. The role is divided by: allosteric inhibitor (Asp); allosteric activator (G6P); substrate binding (HCO₃⁻) and (PEP).

Maize C₄ PEPC	Pineapple CAM isoform	<i>Flaveria trinerva</i> C₄ isoform	Role	Publication
R647	R641	R641	Asp	Kai <i>et al.</i> (2003)
K835	K829	K829	Asp	Kai <i>et al.</i> (2003)
G890	R884	G884	Asp	Paulus <i>et al.</i> (2013)
R894	R888	R888	Asp	Kai <i>et al.</i> (2003)
N968	N962	N964	Asp	Kai <i>et al.</i> (2003)
F679	Q673	Q673	Asp	Paulus <i>et al.</i> (2013)
S780	A774	S774	C ₄ -Polymorphysim	Kai <i>et al.</i> (2003)
H685	R679	H679	CAM-Polymorphism	Kai <i>et al.</i> (2003)
H177	H172	H172	Catalytic base	Kai <i>et al.</i> (2003)
R183	R178	R178	G6P	Kai <i>et al.</i> (2003)
R184	R179	R179	G6P	Kai <i>et al.</i> (2003)
R231	R226	R226	G6P	Kai <i>et al.</i> (2003)
R372	R366	R366	G6P	Kai <i>et al.</i> (2003)
K606	K600	K600	HCO ₃ ⁻	Kai <i>et al.</i> (2003)
K762	K756	K756	HCO ₃ ⁻	Kai <i>et al.</i> (2003)
R763	R757	R757	HCO ₃ ⁻	Kai <i>et al.</i> (2003)
R764	K758	K758	HCO ₃ ⁻	Kai <i>et al.</i> (2003)
W288	W283	W283	Hydrophobic pocket	Kai <i>et al.</i> (2003)
L564*	L558	L558	Hydrophobic pocket	Kai <i>et al.</i> (2003)
M598	M592	M592	Hydrophobic pocket	Kai <i>et al.</i> (2003)
E566	E560	E560	Mn ⁺²	Kai <i>et al.</i> (2003)
D603	D597	D597	Mn ⁺²	Kai <i>et al.</i> (2003)
R456	R450	R450	PEP	Kai <i>et al.</i> (2003)
R647	R641	R641	PEP	Kai <i>et al.</i> (2003)
R759	R753	R753	PEP	Kai <i>et al.</i> (2003)
R773	R767	R767	PEP	Kai <i>et al.</i> (2003)
K12	R8	K8	Phosphorylation	Kai <i>et al.</i> (2003)
S15	S11	S11	Phosphorylation	Kai <i>et al.</i> (2003)
E493	E487	E487	Tetramer formation	Kai <i>et al.</i> (2003)
R498	R492	R492	Tetramer formation	Kai <i>et al.</i> (2003)

*In Kai *et al.* (2003) the residue quoted is L565, however based on the sequenced in the Protein Data Bank that residue corresponds to L564.

Supplementary Material 2 – Table 2. Summary of the distance between the residue Arg679 in subunit A and the residues listed in Table 1 in subunits A, B, C and D of the *Ananas comosus* CAM PEPC isoform model. The distances were measured from the alpha-carbon of Arg679 to the alpha-carbon of each residue. Distances are in Ångstroms. The role is divided by: allosteric inhibitor (Asp); allosteric activator (G6P); substrate binding (HCO_3^-) and (PEP).

Residue	A	B	C	D	Role
R679	0.000	38.987	26.385	46.914	CAM-polymorphism
Q673	10.813	43.478	34.015	53.182	Asp
R641	17.268	43.895	42.870	58.061	Asp –PEP
R884	18.258	35.877	37.890	47.068	Asp
R179	19.428	56.226	31.600	60.421	G6P
N962	20.596	47.234	44.281	60.079	Asp
R178	21.631	57.280	34.570	61.989	G6P
R888	22.229	41.391	39.698	50.641	Asp
H172	22.304	55.806	42.691	65.410	Catalytic base
K829	25.639	43.283	47.142	56.691	Asp
W283	25.715	61.204	46.623	71.981	Hydrophobic pocket
M592	30.194	59.397	54.914	74.789	Hydrophobic pocket
S11	31.798	38.010	37.304	39.033	Phosphorylation
R8	32.026	NA	41.696	NA	Phosphorylation
D597	32.508	58.970	57.072	74.265	Mn^{+2}
R366	32.700	65.903	31.813	66.802	G6P
A774	32.869	57.342	54.040	69.671	C ₄ -polymorphysim
R226	32.906	64.361	44.832	69.500	G6P
R450	32.950	66.559	54.523	78.986	PEP
L558	33.788	63.011	58.110	78.326	Hydrophobic pocket
R767	34.835	62.884	57.474	76.491	PEP
K600	35.363	58.875	59.877	74.916	HCO_3^-
E560	35.586	62.215	60.435	78.343	Mn^{+2}
R753	35.731	67.030	55.407	77.607	PEP
K756	40.905	69.357	62.019	83.334	HCO_3^-
R757	44.241	69.308	65.305	86.858	HCO_3^-
K758	46.132	71.549	65.457	87.948	HCO_3^-
E487	53.072	85.120	73.802	99.588	Tetramer formation
R492	56.806	90.347	75.459	103.37	Tetramer formation

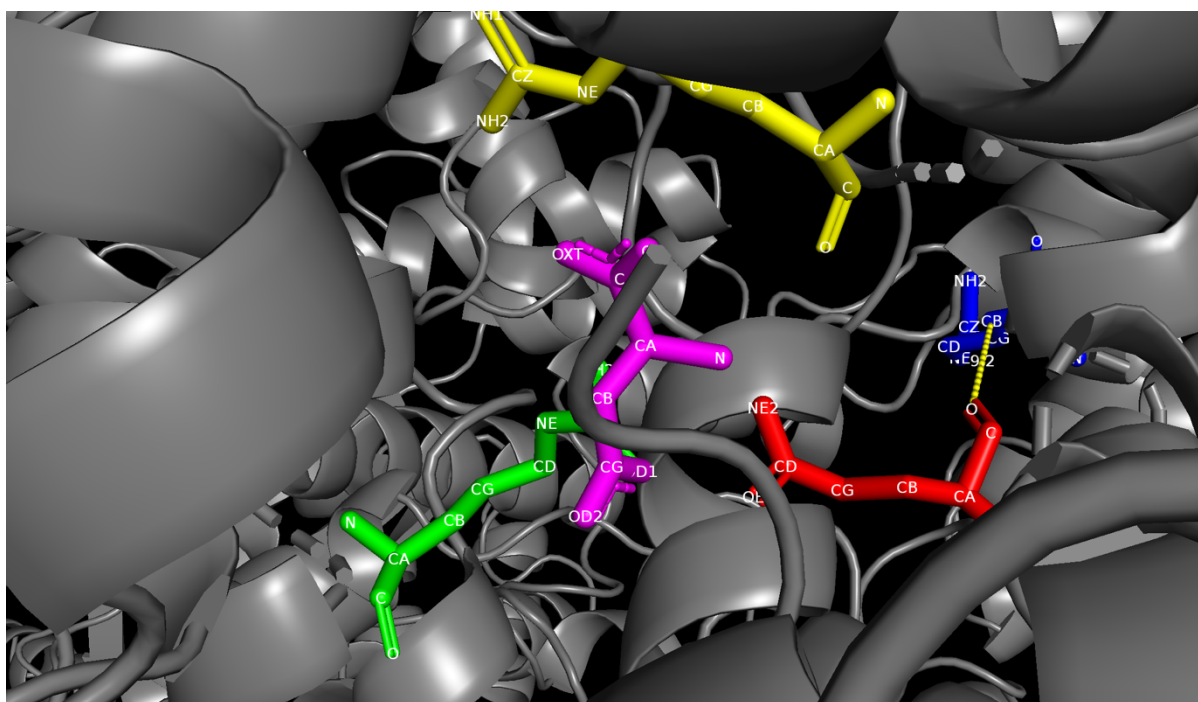
Supplementary Material 2 – Table 3. Summary of the distance between the residue His679 in subunit A and the residues listed in Table 1 in subunits A, B, C and D of the *Ananas comosus* CAM-mutated (His679) PEPC model. The distances were measured from the alpha-carbon of His679 to the alpha-carbon of each residue. Distances are in Ångstroms. The role is divided by: allosteric inhibitor (Asp); allosteric activator (G6P); substrate binding (HCO_3^-) and (PEP).

Residue	A	B	C	D	Role
R679	0.000	38.928	26.206	46.759	CAM-Polymorphism
Q673	10.840	43.442	33.936	53.058	Asp
R641	17.326	43.890	42.788	57.951	Asp
R884	18.233	35.882	37.819	46.957	Asp
R179	19.381	56.158	31.546	60.304	G6P
N962	20.661	47.251	44.225	59.976	Asp
R178	21.640	57.293	34.553	61.947	G6P
R888	22.188	41.404	39.637	50.544	Asp
H172	22.356	55.808	42.668	65.327	Catalytic base
K829	25.635	43.290	47.071	56.562	Asp
W283	25.800	61.213	46.581	71.911	Hydrophobic pocket
M592	30.286	59.403	54.836	74.688	Hydrophobic pocket
S11	31.712	38.003	37.273	38.908	Phosphorylation
R8	31.906	NA	34.401	NA	Phosphorylation
D597	32.583	59.006	56.998	74.144	Mn^{+2}
R366	32.694	65.808	31.769	66.762	G6P
A774	32.868	57.341	53.970	69.543	C4-Polymorphysim
R226	32.881	64.372	44.824	69.382	G6P
R450	33.047	66.581	54.466	78.907	PEP
L558	33.876	63.031	58.038	78.235	Hydrophobic pocket
R767	34.886	62.884	57.426	76.384	PEP
K600	35.419	58.890	59.804	74.805	HCO_3^-
E560	35.657	62.241	60.345	78.225	Mn^{+2}
R753	35.750	67.042	55.354	77.495	PEP
K756	40.943	69.340	61.961	83.252	HCO_3^-
R757	44.292	69.320	65.240	86.782	HCO_3^-
K758	46.193	71.627	65.405	87.858	HCO_3^-
E487	53.187	85.157	73.757	99.527	Tetramer formation
R492	56.894	90.351	75.399	103.282	Tetramer formation

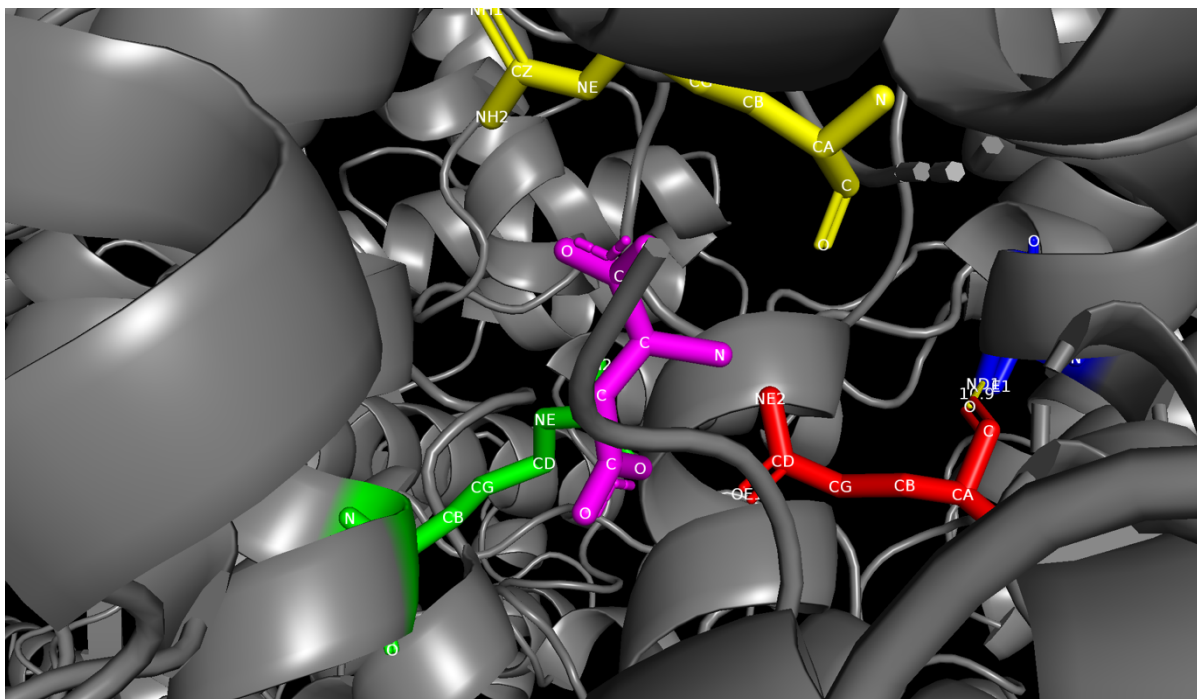
Supplementary Material 2 – Table 4. Differences between distances calculated used the *Ananas comosus* CAM isoform model and the *Ananas comosus* CAM-mutated (His679) isoform model, from subunit A to subunits A, B, C and D. Distances are in Ångstroms. The role is divided by: allosteric inhibitor (Asp); allosteric activator (G6P); substrate binding (HCO_3^-) and (PEP).

Residue	A	B	C	D	Role
R8	0.120	NA	7.295	NA	Phosphorylation
S11	0.086	0.007	0.031	0.125	Phosphorylation
R179	0.047	0.068	0.054	0.117	G6P
R888	0.041	-0.013	0.061	0.097	Asp
R226	0.025	-0.011	0.008	0.118	G6P
R884	0.025	-0.005	0.071	0.111	Asp
R366	0.006	0.095	0.044	0.040	G6P
K829	0.004	-0.007	0.071	0.129	Asp
A774	0.001	0.001	0.070	0.128	C4-Polymorphysim
R679	0.000	0.059	0.179	0.155	CAM-Polymorphism
R178	-0.009	-0.013	0.017	0.042	G6P
R753	-0.019	-0.012	0.053	0.112	PEP
Q673	-0.027	0.036	0.079	0.124	Asp
K756	-0.038	0.017	0.058	0.082	HCO_3^-
R757	-0.051	-0.012	0.065	0.076	HCO_3^-
R767	-0.051	0.000	0.048	0.107	PEP
H172	-0.052	-0.002	0.023	0.083	Cataytic base
K600	-0.056	-0.015	0.073	0.111	HCO_3^-
R641	-0.058	0.005	0.082	0.110	PEP
R641	-0.058	0.005	0.082	0.110	Asp
K758	-0.061	-0.078	0.052	0.090	HCO_3^-
N962	-0.065	-0.017	0.056	0.103	Asp
E560	-0.071	-0.026	0.090	0.118	Mn+2
D597	-0.075	-0.036	0.074	0.121	Mn+2
W283	-0.085	-0.009	0.042	0.070	Hydrophobic pocket
R492	-0.088	-0.004	0.060	0.088	Tetramer formation
L558	-0.088	-0.020	0.072	0.091	Hydrophobic pocket
M592	-0.092	-0.006	0.078	0.101	Hydrophobic pocket
R450	-0.097	-0.022	0.057	0.079	PEP
E487	-0.115	-0.037	0.045	0.061	Tetramer formation

The great majority of homologous residues in the PEPC sequence of *Ananas comosus*, *Flaveria trinervia* and *Zea mays* showed the same amino acid. This probably is the product of purifying selection on these sites reducing non-synonymous substitutions, given the importance of these residues to enzyme function (Kai *et al.*, 2003). However, there are five residues that were not the same in the three species. The C₄-polymorphism is not found in *A. comosus*, and the C₄ species did not show the CAM polymorphism. Another different residue was the residue R8 (*A. comosus* numbering) that differs from K12 or K8 in *Z. mays* and *F. trinervia*, respectively. Another interesting residue is R884 (*A. comosus* numbering) and G884 (*F. trinervia* numbering). Paulus *et al.* (2013) demonstrated that G884 in C₄ *Flaveria* species decreases the malate inhibition in comparison with the R884 found in C₃ *Flaveria* species. This is because G884 does not interact with aspartate or malate, while R884 does with other residues such as Arg888, Lys829, Asn964 and Arg641 (in *F. trinervia* numbering). The homologous residue in *A. comosus* is R884; possibly there are other residues interacting with R884 that makes the R884–aspartate interaction weaker. Given the relatively short distances between the R641, Q673 and R884 with Arg679, the distance between other carbons of Arg679 with other atoms of Q673, was explored. A summary of is shown in Figs. 4 and 5.



Supplementary Material 2 – Figure 4. Interaction of residues R641 (yellow), Q673 (red) and R884 (green) with aspartate (magenta) based on the *Ananas comosus* CAM PEPC isoform model. The residue Arg679 is shown in blue. The distance between the carbon (CB) in Arg679 and the oxygen (O) in Q673 is 9.2 Ångstroms.



Supplementary Material 2 – Figure 5. Interaction of the residues R641 (yellow), Q673 (red) and R884 (green) with aspartate (magenta) based on the *Ananas comosus* CAM-mutated (His679) model. The residue His679 is shown in blue. The distance between the carbon (ND) in His679 and the oxygen (O) in Q673 is 10.9 Ångstroms. His679 is facing the opposite direction from aspartate.

According to Table 4, residues Arg679 and Q673 are closer than in the hypothetical case of His679 and Q673. This was confirmed when the distance was measured from other atoms of the 679 residue. Q673 is important in the interaction with aspartate (Paulus *et al.*, 2013). As shown in Figs. 4–5, in Q673 the OE1 and NE2 atoms interact with aspartate while the O atom interacts with other residues. Given the relative proximity of Arg679 and Q673, it is probable that if Arg679 is of functional importance this will most likely be relevant to malate/aspartate inhibition.

Arginine and histidine are considered positively charged amino acids. However, the positive charge of histidine is only apparent when the pH drops below 6.0. This is because the pK_a of histidine is approximately 6.0. In contrast, the pK_a of arginine is 12.49. These differences suggest that at pH above 7 the charge of histidine will be neutral,

and the charge of arginine will be positive. It is believed that the cytosolic pH in CAM cells at night time is above pH 7, given that the pH optimum of PEPC is around 8.0 (Buchanan-Bollig *et al.*, 1984). It is likely that the positively charged Arg679 residue interacts with malate and/or aspartate more than the neutral His679.

To test the functionality of Arg679, it is necessary to clone the PEPC CAM isoform and C₃ isoform and induce mutation in residue Arg679 and His680 to His679 and Arg680, respectively. First, high-quality RNA from *Ananas comosus* was extracted following the protocol in Supplementary material 1 of this chapter. Once high-quality RNA was obtained, a cDNA synthesis was conducted using the Maxima H Minus First Strand cDNA Synthesis Kit® (Thermo Scientific) according to the manufacturer's protocol. After cDNA synthesis, RT-PCR was conducted using Phusion High-Fidelity DNA Polymerase® (2 U/μL) (Thermo Scientific). The reaction contained 1X Phusion HF Buffer, 200 μM dNTPs, 0.5 μM Primer F, 0.5 μM Primer R, 10% (v/v) template, and 0.002 U/μl. The primers are described in Table 5.

Supplementary Material 2 – Table 5. Primers used to amplify the *ppc-1* CAM and C₃ isoforms in *Ananas comosus*.

Primer*	Sequence 5'–3'	<i>ppc-1</i> isoform
BromCAM_Forward	TA <u>GGATCC</u> ATGGCGAAGAGTTCGGT	CAM
BromCAM_Reverse	GGGGTACCCTAACCAGTGTTCTCAATC	CAM
BromC ₃ _Forward	TA <u>GGATCC</u> ATGTCGCGGAACGTGGTG	C ₃
BromC ₃ _Reverse	GGGGTACCCTAACCAGTGTTCTGCATAC	C ₃

*The forward primers contained a restriction site *Bam*HI at the 5' end, while the reverse primers contained a restriction site *Kpn*I at 3'. The restriction site in each case is underlined.

The band at ~2900 bp was purified using the GeneJET Gel Extraction Kit® (ThermoFisher Scientific) and the fragment was incorporated in the vector pQE-30®

(Qiagen) using temperature shock. The vector pQE-30 contains a with a 6-His tag at the N-terminus for protein purification using nickel affinity chromatography. The colonies containing the *ppc-1* sequences were evaluated using a PCR with the primers used to amplify the fragment originally. Positive colonies were then sequenced to verify that there were no insertions, deletions or mutations in the sequence. The sequences of *ppc-1* CAM and *ppc-1* C₃ isoforms of *Ananas comosus* were successfully isolated. Once the *ppc-1* insert was verified, the protein was expressed and purified. Protein expression was performed as described by Paulus *et al.* (2013). An initial culture of ~10 µl of bacteria (or the tip of the colony) in 5 ml of 2YT media (5 g l⁻¹ sodium chloride, 10 g l⁻¹ yeast extract and 16 g l⁻¹ peptone) with 1% (w/v) ampicillin and 1% (w/v) glucose. The culture was left over night at 37 °C shaking at 250 rpm. An aliquot of ~1 ml was added to 1 litre of 2YT media with 1% (w/v) ampicillin and 1% (w/v) glucose at 37 °C at 250 rpm until reaching OD₆₀₀ = 0.8. The expression of recombinant proteins was then induced with 1 mM IPTG (final concentration) and cells were grown for 18 h at 16 °C. After harvesting, cells were suspended in lysis buffer (50 mM Tris, 10 mM MgCl₂, 150 mM NaCl, lysozyme, pH = 8.0) and disrupted by sonication. The extract was then centrifuged at 11 000 *g* for 30 min and the supernatant was equilibrated with the nickel column (15 ml total volume of the column, containing 3 ml of Ni-NTA Agarose®; Qiagen) for 30 min at 4 °C. After equilibration, the column was open in both ends and the supernatant was filtered. An aliquot of the elution (20 µl) was collected and analysed by SDS-PAGE. Then a series of washing steps (10 ml) was performed using a washing buffer (50 mM NaH₂PO₄, 300 mM NaCl, 20 mM imidazole, pH = 8.0). After each washing step an aliquot of the eluate was collected, analysed by SDS-PAGE, and the absorbance monitored until it reached the absorbance of the washing buffer. Then 2.5 ml of elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 250 mM imidazole, pH = 8.0) was added and the whole elution was saved.

An aliquot of 1–3 μ l was used to measure the protein based on the absorbance at 280 and 260 nm. The final elute was mixed with 10% (v/v) glycerol and kept at -80°C until enzymatic assays.

The goal is to compare the kinetics of the recombinant PEPC CAM (His679) and C₃ (Arg680) isoforms to establish the possible functionality of the polymorphism in the residue 679 in monocots. This is currently being investigated. At the moment, the predicted 3D protein model of PEPC suggests the possibility that the polymorphism could be associated with inhibition of enzyme activity by the negatively charged allosteric effector malate.

References

Adams III WW, Demmig-Adams B. 1996. Energy dissipation and the xanthophyll cycle in CAM plants. In Winter K and Smith JAC [eds.], *Crassulacean acid metabolism: biochemistry, ecophysiology and evolution*, 97–114. Heidelberg: Springer-Verlag.

Alvarado-Cárdenas LO, Martínez-Meyer E, Feria TP, Eguiarte LE, Hernández HM, Midgley G, Olson ME. 2013. To converge or not to converge in environmental space: testing for similar environments between analogous succulent plants of North America and Africa. *Annals of Botany* 111: 1125–1138.

Amidon WH, Fisher GB, Burbank DW, Ciccioli PL, Alonso RN, Gorin AL, Silverhart PH, Kylander-Clark AR, Christoffersen MS. 2017. Mio-Pliocene aridity in the south-central Andes associated with Southern Hemisphere cold periods. *Proceedings of the National Academy of Sciences of the United States of America* 114: 6474–6479.

Antonelli A, Nylander JA, Person C, Sanmartín I. 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences of the United States of America* 106: 9749–9754.

Antonelli A, Sanmartín I. 2011. Why are there so many plant species in the Neotropics? *Taxon* 60: 403–414.

Arakaki M, Christin PA, Nyffeler R, Lendel A, Eggli U, Ogburn, RM, Spriggs E, Moore MJ, Edwards EJ. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences of the United States of America* 108: 8379–8384.

Araújo MB, Guisan A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography*. 31: 1677–1688.

Araújo MB, Pearson RG, Thuiller W, Erhard M. 2005. Validation of species–climate impact models under climate change. *Global Change Biology* 11: 1504–1513.

Arnold ML, Kunte K. 2017. Adaptive genetic exchange: a tangled history of admixture and evolutionary innovation. *Trends in Ecology and Evolution*. 32: 601–611.

Asai N, Nakajima N, Tamaoki M, Kamada H, Kondo N. 2000. Role of malate synthesis mediated by phosphoenolpyruvate carboxylase in guard cells in the regulation of stomatal movement. *Plant and Cell Physiology* 41:10–15.

Austin M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling* 200: 1–19

Bandurski RS, Greiner CM. 1953. The enzymatic synthesis of oxalacetate from phosphoryl-enolpyruvate and carbon dioxide. *Journal of Biological Chemistry* 204: 781–786.

Baresch A, Smith JAC, Winter K, Valerio AC, Jaramillo C. 2011. *Karatophyllum bromelioides* L.D. Gómez revisited: a probable CAM fossil bromeliad. *American Journal of Botany* 98: 1905–1908.

Barfuss MH, Till W, Leme EM, Pinzon JP, Manzanares JM, Halbritter H, Samuel R, Brown GK. 2016. Taxonomic revision of Bromeliaceae subfam. Tillandsioideae based on a multi-locus DNA sequence phylogeny and morphology. *Phytotaxa* 279: 1–97.

Beck E, Senser M, Scheibe R, Steiger H, Pongratz P. 1982. Frost avoidance and freezing tolerance in Afroalpine ‘giant rosette’ plants. *Plant, Cell and Environment* 5: 215–222.

Beck E. 1994. Cold tolerance in tropical alpine plants. In Rundel P, Smith AP, Meinzer FC [eds.], *Tropical alpine environments. Plant form and function*, 77–110. Cambridge: Cambridge University Press.

Begon M, Townsend CR, Harper JL. 2006. *Ecology: From Individuals to Ecosystems* 4th Edition. Oxford: Blackwell Publishing.

Beltrán JD, Lasso E, Madriñán S, Virgo A, Winter K. 2013. Juvenile tank-bromeliads lacking tanks: do they engage in CAM photosynthesis? *Photosynthetica* 51: 55–62.

Bender M, Rouhani I, Vines M, Black CC Jr. 1973. $^{13}\text{C}/^{12}\text{C}$ ratio changes in crassulacean acid metabolism plants. *Plant Physiology* 52: 427–430.

Bendix J, Rollenbeck R, Reudenbach C. 2005. Diurnal patterns of rainfall in a tropical Andean valley of southern Ecuador as seen by a vertically pointing K-band Doppler radar. *International Journal of Climatology* 26: 829–846.

Benito-Garzón M, Sanchez de Dios R, Sainz Ollero H. 2008. Effects of climate change on the distribution of Iberian tree species. *Applied Vegetation Science* 11: 169–178.

Benzing DH. 2000. *Bromeliaceae: profile of an adaptive radiation*. Cambridge: Cambridge University Press.

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* 42: 252–258.

Blomberg SP, Garland JrT, Ives AR. 2003. Testing for phylogenetic signal in comparative data: behavioural traits are more labile. *Evolution* 57: 717–745.

Blonde JD, Plaxton WC. 2003. Structural and kinetic properties of high and low molecular mass phosphoenolpyruvate carboxylase isoforms from the endosperm of developing castor oilseeds. *Journal of Biological Chemistry* 278:11867–11873.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Bolton CT, Hernández-Sánchez MT, Fuertes MÁ, González-Lemos S, Abrevaya L, Mendez-Vicente A, Flores JA, Probert I, Giosan L, Johnson J, Stoll HM. 2016. Decrease in coccolithophore calcification and CO₂ since the middle Miocene. *Nature Communications* 7: 1–13.

Bolton CT, Stoll HM. 2013. Late Miocene threshold response of marine algae to carbon dioxide limitation. *Nature* 500: 558–562.

Borland AM, Guo HB, Yang X, Cushman JC. 2016. Orchestration of carbohydrate processing for crassulacean acid metabolism. *Current Opinion in Plant Biology* 31: 118–124.

Borland AM, Hartwell J, Weston D, Schlauch K, Tschaplinski T, Tuskan G, Yang X, Cushman JC. 2014. Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends in Plant Science* 19: 327–338.

Borland AM, Hartwell J, Weston DJ, Schlauch KA, Tschaplinski TJ, Tuskan GA, Yang X, Cushman JC. 2014. Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends in Plant Science* 5: 327–338.

Borland AM, Zambrano VAB, Ceusters K, Shorrock K. 2011. The photosynthetic plasticity of crassulacean acid metabolism: an evolutionary innovation for sustainable productivity in a changing world. *New Phytologist* 191: 619–633.

Breiman L. 2001. Random forests. *Machine Learning* 45: 5–32.

Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B, Pellissier L, Yoccoz NG, Thuiller W, Fortin MJ, Randin C, Zimmermann NE, Graham CH. 2012.

Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography* 21: 481–497.

Buchanan-Bollig IC, Kluge M. 1981. Crassulacean acid metabolism (CAM) in *Kalanchoë daigremontiana*: temperature response of phosphoenolpyruvate (PEP)-carboxylase in relation to allosteric effectors. *Planta* 152: 181–188.

Buchanan-Bollig IC, Smith JAC. 1984. Circadian rhythms in crassulacean acid metabolism: phase relationships between gas exchange, leaf water relations and malate metabolism in *Kalanchoë daigremontiana*. *Planta* 161: 314–319.

Buchanan-Bollig IC, Kluge M, Müller D. 1984. Kinetic changes with temperature of phosphoenolpyruvate carboxylase from a CAM plant. *Plant, Cell and Environment*. 7: 63–70.

Cai J, Liu X, Vanneste K, Proost S, Tsai WC, Liu KW, Chen LJ, He Y, Xu Q, Bian C, Zheng Z. 2015. The genome sequence of the orchid *Phalaenopsis equestris*. *Nature Genetics* 47: 65–72

Cernusak LE, Ubierna N, Winter K, Holtum JAM, Marshall JD, Farquhar GD. 2013. Environmental and physiological determinants of carbon isotope discrimination in terrestrial plants. *New Phytologist* 200: 950–965.

Ceusters J, Borland AM, Taybi T, Frans M, Godts C, De Proft MP. 2014. Light quality modulates metabolic synchronization over the diel phases of crassulacean acid metabolism. *Journal of Experimental Botany* 65: 3705–3714.

Cheviron ZA, Connaty AD, McClelland GB, Storz JF. 2014. Functional genomics of adaptation to hypoxic cold-stress in high-altitude deer mice: transcriptomic plasticity and thermogenic performance. *Evolution* 68: 48–62.

Chiuri R, Maiorano G, Rizzello A, Del Mercato LL, Cingolani R, Rinaldi R, Maffia M, Pompa PP. 2009. Exploring local flexibility/rigidity in psychrophilic and mesophilic carbonic anhydrases. *Biophysical Journal* 96: 1586–1596.

Chollet, R., Vidal, J. and O’Leary, M. 1996. Phosphoenolpyruvate carboxylase: a ubiquitous, highly regulated enzyme in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 47: 273–298.

Christin PA, Arakaki M, Osborne CP, Bräutigam A, Sage RF, Hibberd JM, Kelly S, Covshoff S, Wong GK, Hancock L, Edwards EJ. 2014. Shared origins of a key enzyme during the evolution of C₄ and CAM metabolism. *Journal of Experimental Botany* 65: 3609–3621.

Christin PA, Besnard G. 2009. Two independent C₄ origins in Aristidoideae (Poaceae) revealed by the recruitment of distinct phosphoenolpyruvate carboxylase genes. *American Journal of Botany* 96: 2234–2239.

Christin PA, Boxall SF, Gregory R, Edwards EJ, Hartwell J, Osborne CP. 2013. Parallel recruitment of multiple genes into C₄ photosynthesis. *Genome Biology and Evolution* 5: 2174–2187.

Christin PA, Freckleton RP, Osborne CP. 2010a. Can phylogenetics identify C₄ origins and reversals? *Trends in Ecology and Evolution* 25: 403–409.

Christin PA, Salamin N, Savolainen V, Besnard G. 2007. A phylogenetic study of the phosphoenolpyruvate carboxylase multigene family in Poaceae: Understanding the molecular changes linked to C₄ photosynthesis evolution. *Kew Bulletin* 1: 455–462.

Christin PA, Weinreich DM, Besnard G. 2010b. Causes and evolutionary significance of genetic convergence. *Trends in Genetics* 26: 400–405.

Christopher JT, Holtum JAM. 1996. Patterns of carbohydrate partitioning in

the leaves of crassulacean acid metabolism species during deacidification. *Plant Physiology* 112: 393–399.

Cockburn W. 1985. Variation in photosynthetic acid metabolism vascular plants: CAM and related phenomena. *New Phytologist* 101: 3–24.

Crayn DM, Winter K, Schulte K, Smith JAC. 2015. Photosynthetic pathways in Bromeliaceae: the phylogenetic and ecological significance of CAM and C₃ based on carbon isotope ratios for 1893 species. *Botanical Journal of the Linnean Society* 178: 169–221.

Crayn DM, Winter K, Smith JAC. 2004. Multiple origins of crassulacean acid metabolism and the epiphytic habit in the Neotropical family Bromeliaceae. *Proceedings of the National Academy of Sciences of the United States of America* 101: 3703–3708.

de Villemereuil P, Gaggiotti OE, Mousterde M, Till-Bottraud I. 2016. Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity* 116: 249–254.

Deblauwe V, Droissart V, Bose R, Sonké B, Blach-Overgaard A, Svenning JC. 2016. Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. *Global Ecology and Biogeography* 25: 443–454.

Deng H, Zhang LS, Zhang GQ, Zheng BQ, Liu ZJ, Wang Y. 2016. Evolutionary history of PEPC genes in green plants: Implications for the evolution of CAM in orchids. *Molecular Phylogenetics and Evolution* 94: 559–564.

Dodd AN, Borland AM, Haslam RP, Griffiths H, Maxwell K. 2002. Crassulacean acid metabolism: plastic, fantastic. *Journal of Experimental Botany* 53: 569–580.

Donoghue MJ. 2008. A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences of the United States of America* 105:

11549–11555.

Dray S, Dufour AB, Chessel D. 2007. The ade4 package-II: Two-table and K-table methods. *R News* 7: 47–52.

Droc G, Larivière D, Guignon V, Yahiaoui N, This D, Garsmeur O, Dereeper A, Hamelin C, Argout X, Dufayard JF, Lengelle J, Baurens FC, Cenci A, Pitollat B, D'Hont A, Ruiz M, Rouard M, Bocs S. 2013. The banana genome hub. *Database: the journal of biological databases and curation* 2013: 1–14.

Duellman WE. 1999. *Patterns of Distribution of Amphibians: A Global Perspective*. Baltimore: Johns Hopkins University Press.

Dynesius M, Jansson R. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences of the United States of America* 97: 9115–9120.

Earnshaw MJ, Winter K, Ziegler H, Stichler W, Cruttwell N, Kerenga K, Cribb PJ, Wood J, Croft JR, Carver KA, Gunn TC. 1987. Altitudinal changes in the incidence of crassulacean acid metabolism in vascular epiphytes and related life forms in Papua New Guinea. *Oecologia* 73: 566–572.

Edgar RC. 2004. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460–2461.

Edwards EJ, Díaz M. 2006. Ecological physiology of *Pereskia guamacho*, a cactus with leaves. *Plant, Cell and Environment* 29: 247–256.

Edwards EJ, Nyffeler R, Donoghue MJ. 2005. Basal cactus phylogeny: implications of *Pereskia* (Cactaceae) paraphyly for the transition to the cactus life form. *American Journal of Botany* 92:1177–1188.

Edwards EJ, Ogburn RM. 2012. Angiosperm responses to a low-CO₂ world: CAM and C₄ photosynthesis as parallel evolutionary trajectories. *International Journal of Plant Sciences* 173: 724–733.

Elias M, Wieczorek G, Rosenne S, Tawfik DS. 2014. The universality of enzymatic rate–temperature dependency. *Trends in Biochemical Sciences* 39: 1–7.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17: 43–57.

Ellenberg H. 1981. Reasons for stem succulents being present or absent in the arid regions of the world (in German). *Flora* 171: 114–169.

Elton C. 1927. *Animal Ecology*. London: Sedgwick and Jackson.

Engel Z, Skrzypek G, Chuman T, Šefrna L, Mihaljevič M. 2014. Climate in the Western Cordillera of the Central Andes over the last 4300 years. *Quaternary Science Reviews* 99: 60–77.

Falvey M, Garreaud RD. 2009. Regional cooling in a warming world: Recent temperature trends in the southeast Pacific and along the west coast of subtropical South America (1979–2006). *Journal of Geophysical Research: Atmospheres* 114: 1–16.

Farquhar GD, Ehleringer JR, Hubick KT. 1989. Carbon isotope discrimination and photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology* 40: 503–537.

Feller G, Gerday C. 2003. Psychrophilic enzymes: hot topics in cold adaptation. *Nature Reviews. Microbiology* 1: 200–208.

Feller G. 2003 Molecular adaptations to cold in psychrophilic enzymes. *Cellular and Molecular Life Sciences* 60: 648–662.

Feller G. 2010. Protein stability and enzyme activity at extreme biological temperatures. *Journal of Physics: Condensed Matter* 22: 323101.

Feller G. 2013. Psychrophilic enzymes: from folding to function and biotechnology. *Scientifica* 2013: 1–28.

Fields PA, Somero GN. 1988. Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A₄ orthologs of Antarctic notothenioid fishes. *Proceedings of the National Academy of Sciences of the United States of America* 95: 11476–11481.

Foster GL, Lear CH, Rae JW. 2012. The evolution of pCO₂, ice volume and climate during the middle Miocene. *Earth and Planetary Science Letters* 341: 243–254.

Fox J, Weisberg S. 2011. *Multivariate linear models in R. An R Companion to Applied Regression*. Los Angeles: Thousand Oaks.

Frankl A, Zwertvaegher A, Poesen J, Nyssen J. 2013. Transferring Google Earth observations to GIS-software: example from gully erosion study. *International Journal of Digital Earth* 6: 196–201.

Franklin J. 2010. *Mapping species distributions: spatial inference and prediction*. Cambridge: Cambridge University Press.

Freschi L, Takahashi CA, Cambui CA, Semprebom TR, Cruz AB, Miotto PT, de Melo Versieux L, Calvente A, Latansio-Aidar SR, Aidar MP, Mercier H. 2010. Specific leaf areas of the tank bromeliad *Guzmania monostachia* perform distinct functions in response to water shortage. *Journal of Plant Physiology* 167: 526–33.

García-Meneses PM, Ramsay PM. 2012. Pollinator response to within- patch spatial context determines reproductive output of a giant rosette plant. *Basic and Applied Ecology* 13: 516–523.

García-Meneses PM, Ramsay PM. 2014. *Puya hamata* demography as an indicator of recent fire history in the paramo of El Angel and Volcán Chiles, Ecuador-Colombia. *Caldasia* 36: 53–69.

Garreaud RD. 2009. The Andes climate and weather. *Advances in Geosciences* 7: 3–11.

Garzione CN, Hoke GD, Libarkin JC, Withers S, MacFadden B, Eiler J, Ghosh P, Mulch A. 2008. Rise of the Andes. *Science* 320: 1304–1307.

Garzón MB, de Dios RS, Ollero HS. 2008. Effects of climate change on the distribution of Iberian tree species. *Applied Vegetation Science*. 11: 169–178.

Gaut BS, Muse SV, Clark WD, Clegg MT. 1992. Relative rates of nucleotide substitution at the *rbcl* locus in monocotyledonous plants. *Journal of Molecular Evolution* 35: 292–303.

GBIF. 2012. Global Biodiversity Information Facility, *Puya*, 3744 records, accessed on 20 September 2014: <http://data.gbif.org>

Gehrig H, Heute V, Kluge M. 2001. New partial sequences of phosphoenolpyruvate carboxylase as molecular phylogenetic markers. *Molecular Phylogenetics and Evolution* 20: 262–274.

Gehrig H, Taybi T, Kluge M, Brulfert J. 1995. Identification of multiple PEPC isogenes in leaves of the facultative crassulacean acid metabolism (CAM) plant *Kalanchoe blossfeldiana*. *Federation of European Biochemical Societies Letters* 377: 399–402.

Gehrig HH, Aranda J, Cushman MA, Virgo A, Cushman JC, Hammel BE, Winter K. 2003. Cladogram of Panamanian *Clusia* based on nuclear DNA: Implications for the origins of crassulacean acid metabolism. *Plant Biology* 5: 59–70.

Gehrig HH, Heute V, Kluge M. 1998. Toward a better knowledge of the molecular evolution of phosphoenolpyruvate carboxylase by comparison of partial cDNA sequences. *Journal of Molecular Evolution* 46: 107–114.

Gehrig HH, Winter K, Cushman J, Borland A, Taybi T. 2000. An improved RNA isolation method for succulent plant species rich in polyphenols and polysaccharides. *Plant Molecular Biology Reporter* 18: 369–376.

Gehrig HH, Wood J, Cushman MA, Virgo A, Cushman JC, Winter K. 2005. Large gene family of phosphoenolpyruvate carboxylase in the crassulacean acid metabolism plant *Kalanchoe pinnata* (Crassulaceae). *Functional Plant Biology* 32: 467–472.

Georlette D, Blaise V, Collins T, D'Amico S, Gratia E, Hoyoux A, Marx JC, Sonan G, Feller G, Gerday C. 2004. Some like it cold: biocatalysis at low temperatures. *Federation of European Microbiology Societies Microbiology Reviews* 28: 25–42.

Givnish TJ, Barfuss MHJ, Van Ee B, Riina R, Schulte K, Horres R, Gonsiska PA, Jabaily RS, Crayn DM, Smith JAC, Winter K, Brown GK, Evans TM, Holst BK, Luther H, Till W, Zizka G, Berry PE, Sytsma KJ. 2011. Adaptive radiation and diversification in Bromeliaceae: insights from a 7-locus plastid phylogeny. *American Journal of Botany* 98: 872–895.

Givnish TJ, Barfuss MHJ, Van Ee B, Riina R, Schulte K, Horres R, Gonsiska PA, Jabaily RS, Crayn DM, Smith JAC, Winter K, Brown GK, Evans TM, Holst BK, Luther H, Till W, Zizka G, Berry PE, Sytsma KJ. 2014. Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution* 71: 55–78.

Givnish TJ, Millam KC, Berry PE, Sytsma KJ. 2007. Phylogeny, adaptive radiation, and historical biogeography of Bromeliaceae inferred from *ndhF* sequence

data. In J.T. Columbus, E.A. Friar, C.W. Hamilton, J.M. Porter, L.M. Prince, M.G. Simpson [eds.], *Monocots: comparative biology and evolution: Poales*, 3–26. Claremont, CA: Rancho Santa Ana Botanic Garden.

Goldstein G, Rada F, Azocár A. 1985. Cold hardiness and supercooling along an altitudinal gradient in Andean giant rosette species. *Oecologia* 68:147–152.

Gómez Romero, S. E. & A. Grau. 2009. Las especies de *Puya* (Bromeliaceae) en la Argentina. *Boletín de la Sociedad Argentina de Botánica* 44: 175–208.

Gouda EJ, Butcher D, Gouda CS. 2017. Encyclopaedia of Bromeliads, Version 3. <http://encyclopedia.florapix.nl/>. University Botanic Gardens, Utrecht.

Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* 27: 221–224.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

Grafen A. 1989. The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 326: 119–157.

Graham EA, Mulkey SS, Kitajima K, Phillips NG, Wright SJ. 2003. Cloud cover limits net CO₂ uptake and growth of a rainforest tree during tropical rainy seasons. *Proceedings of the National Academy of Sciences of the United States of America* 100: 572–576.

Grau A, Gómez–Romero E, Aráoz E. 2010. Puyas andinas. *Ciencia Hoy*. 20: 1–18.

Graves S, Piepho HP, Selzer L, Dorai-Raj S. 2015. multcompView: Visualizations of Paired Comparisons. R package version 0.1-7. <https://CRAN.R-project.org/package=multcompView>

Gregory-Wodzicki KM. 2000. Relationships between leaf morphology and climate, Bolivia: Implications for estimating paleoclimate from fossil floras. *Paleobiology* 26: 668–688.

Griffiths H, Smith JAC. 1983. Photosynthetic pathways in the Bromeliaceae of Trinidad: relations between life-forms, habitat preference and the occurrence of CAM. *Oecologia* 60: 176–184.

Griffiths H. 1988. Crassulacean acid metabolism: a re-appraisal of physiological plasticity in form and function. *Advances in Botanical Research* 15: 43–92.

Grinnell J. 1917. The niche-relationships of the California Thrasher. *The Auk* 34: 427–433.

Gross SM, Martin JA, Simpson J, Abraham-Juarez MJ, Wang Z, Visel A. 2013. De novo transcriptome assembly of drought tolerant CAM plants, *Agave deserti* and *Agave tequilana*. *BMC Genomics*. 14: 563–577.

Guillet C, Just D, Bénard N, Destrac-Irvine A, Baldet P, Hermould M, Causse M, Raymond P, Rothan C. 2002. A fruit-specific phosphoenolpyruvate carboxylase is related to rapid growth of tomato fruit. *Planta* 214: 717–726.

Guisan A, Thuiller W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993–1009.

Guisan A, Zimmermann N. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186.

Gurevitch J, Scheiner SM, Fox GA. 2006. The ecology of plants. Sunderland: Sinauer Associates UK.

Guy CL. 1990. Cold acclimation and freezing stress tolerance: role of protein metabolism. *Annual Review of Plant Biology*. 41: 187–223.

Hancock L, Edwards EJ. 2014. Phylogeny and the inference of evolutionary trajectories. *Journal of Experimental Botany*. 65: 3491–3498.

Hartwell J, Gill A, Nimmo GA, Wilkins MB, Jenkins GI, Nimmo HG. 1999. Phosphoenolpyruvate carboxylase kinase is a novel protein kinase regulated at the level of gene expression. *The Plant Journal* 20: 333–342.

Hartwell J, Smith LH, Wilkins MB, Jenkins GI, Nimmo HG. 1996. Higher plant phosphoenolpyruvate carboxylase kinase is regulated at the level of translatable mRNA in response to light or a circadian rhythm. *The Plant Journal* 10: 1071–1078.

Hartwell J. 2005a. The circadian clock in CAM plants. In Hall AJW, McWatters HG [eds.], *Endogenous plant rhythms. Annual Plant Reviews. Vol. 21*, 211–236. Oxford: Blackwell.

Hartwell J. 2005b. The co-ordination of central plant metabolism by the circadian clock. *Biochemical Society Transactions* 33: 945–948.

Hastie T, Fithian W. 2013. Inference from presence-only data; the ongoing controversy. *Ecography*. 36: 864–867.

Hedberg O. 1964. Features of afroalpine plant ecology. *Acta Phytogeographica Suecica* 49: 1–144.

Herrera A, Martin CE, Tezara W, Ballestrini C, Medina E. 2010. Induction by drought of crassulacean acid metabolism in the terrestrial bromeliad, *Puya floccosa*. *Photosynthetica* 48: 383–388.

Heyduk K, Burrell N, Lalani F, Leebens-Mack J. 2016a. Gas exchange and leaf anatomy of a C₃-CAM hybrid, *Yucca gloriosa* (Asparagaceae). *Journal of Experimental Botany* 67: 1369–1379.

Heyduk K, McKain MR, Lalani F, Leebens-Mack J. 2016b. Evolution of a CAM anatomy predates the origins of Crassulacean acid metabolism in the Agavoideae (Asparagaceae). *Molecular Phylogenetics and Evolution* 105: 102–113.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.

Hijmans RJ, Phillips S, Leathwick J, Elith J. 2015. dismo: Species distribution modeling. R package version 1.0. The R Foundation for Statistical Computing, Vienna <http://cran.r-project.org>.

Hijmans RJ, van Etten J. 2014. raster: Geographic data analysis and modeling. R package version 1.0. The R Foundation for Statistical Computing, Vienna <http://cran.r-project.org>.

Hill L, Hector A, Hemery G, Smart S, Tanadini M, Brown N. 2017. Abundance distributions for tree species in Great Britain: A two-stage approach to modeling abundance using species distribution modeling and random forest. *Ecology and Evolution* 7: 1043–56.

Holtum JAM, Hancock LP, Edwards EJ, Crisp MD, Crayn DM, Sage R, Winter K. 2016. Australia lacks stem succulents but is it depauperate in plants with crassulacean acid metabolism (CAM)? *Current Opinion in Plant Biology* 31: 109–117.

Holtum JAM, Smith JAC, Neuhaus H. 2005. Intracellular transport and pathways of carbon flow in plants with crassulacean acid metabolism. *Functional Plant Biology* 32: 429–449.

Horn JW, Xi Z, Riina R, Peirson JA, Yang Y, Dorsey BL, Berry PE, Davis CC, Wurdack KJ. 2014. Evolutionary bursts in *Euphorbia* (Euphorbiaceae) are linked with photosynthetic pathway. *Evolution* 68: 3485–3504.

Hornung-Leoni C, Sosa V, Lopez MG. 2007. Xylose in the nectar of *Puya raimondii* (Bromeliaceae), the Queen of the Puna. *Biochemical Systematics and Ecology* 35: 554–556.

Hornung-Leoni CT, González-Gómez PL, Troncoso AJ. 2013. Morphology, nectar characteristics and avian pollinators in the Andean *Puya* species (Bromeliaceae). *Acta Oecologica* 51: 54–61.

Hornung-Leoni CT, Sosa V. 2008. Morphological phylogenetics of *Puya* subgenus *Puya* (Bromeliaceae). *Botanical Journal of the Linnean Society* 156: 93–110.

Hothorn T, Buehlmann P, Dudoit S, Molinaro A, Van Der Laan M. 2006. Survival Ensembles. *Biostatistics* 7: 355–373.

Hutchinson GE. 1957. Concluding remarks. In Cold Spring Harbor Symposia on Quantitative Biology. Cold Spring. *Harbor Laboratory Press* 22: 415–427.

Ibaraki M. 1997. Closing of the Central American Seaway and Neogene coastal upwelling along the Pacific coast of South America. *Tectonophysics* 281: 99–104.

Izui K, Matsumura H, Furumoto T, Kai Y. 2004. Phosphoenolpyruvate carboxylase: a new era of structural biology. *Annual Review of Plant Biology* 55: 69–84.

Jabaily RS, Sytsma KJ. 2010. Phylogenetics of *Puya* (Bromeliaceae): placement, major lineages, and evolution of Chilean species. *American Journal of Botany* 97: 337–356.

Jabaily RS, Sytsma KJ. 2013. Historical biogeography and life-history evolution of Andean *Puya* (Bromeliaceae). *Botanical Journal of the Linnean Society* 171: 201–224.

Jackson ST, Overpeck JT. 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology* 26: 194–220.

James FC, Johnston RF, Warner NO, Niemi G, Boecklen W. 1984. The Grinnellian niche of the Wood Thrush. *The American Naturalist* 124: 17–47.

Janeba Z. 2017. A New Species of *Puya* (Bromeliaceae) from Coastal Peru. *Cactus and Succulent Journal* 89: 176–184.

Janitza S, Strobl C, Boulesteix AL. 2013. An AUC-based permutation variable importance measure for random forests. *BMC Bioinformatics*. 14: 1–11.

Jaramillo C, Montes C, Cardona A, Silvestro D, Antonelli A, Bacon CD. 2017. Comment (1) on “Formation of the Isthmus of Panama” by O’Dea *et al.* *Science Advances*. 3: e1602321.

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics*. 8: 275–282.

Jones R, Ougham H, Thomas H, Waaland S. 2013. The Molecular Life of Plants 1st Edition. Wiley-Blackwell, Oxford, UK.

Kai Y, Matsumura H, Izui K. 2003. Phosphoenolpyruvate carboxylase: three-dimensional structure and molecular mechanisms. *Archives of Biochemistry and Biophysics*. 414: 170–179.

Keeley JE, Keeley SC. 1989. Crassulacean acid metabolism (CAM) in high elevation tropical cactus. *Plant, Cell and Environment* 12: 331–336.

Keeley JE. 1998. CAM photosynthesis in submerged aquatic plants. *The Botanical Review* 64: 121–75.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26: 1463–1464.

Killeen T, Douglas M, Consiglio T, Jorgensen PM, Mejia J. 2007. Dry spots and wet spots in the Andean hotspot. *Journal of Biogeography* 34: 1357–1373.

Kluge M, Ting IP. 1978. Crassulacean acid metabolism. Analysis of an ecological adaptation. Berlin, Germany: Springer.

Konowalik K, Proćków M, Proćków J. 2017. Climatic niche of *Selinum alatum* (Apiaceae, Selineae), a new invasive plant species in Central Europe and its alterations according to the climate change scenarios: Are the European mountains threatened by invasion? *PLOS One*. 12: e0182793.

Körner C, Farquhar GD, Roksandic Z. 1988. A global survey of carbon isotope discrimination in plants from high altitude. *Oecologia* 74: 623–632.

Körner C, Farquhar GD, Wong SC. 1991. Carbon isotope discrimination by plants follows latitudinal and altitudinal trends. *Oecologia* 88: 30–40.

Körner C. 2012. Alpine treelines: functional ecology of the global high elevation tree limits. Springer Science & Business Media.

Körner Ch, Farquhar GD, Roksandic Z. 1988. A global survey of carbon isotope discrimination in plants from high altitude. *Oecologia* 74: 623–632.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33: 1870–1874.

Kunitake G, Stitt C, Saltman P. 1959. Dark fixation of CO₂ by tobacco leaves. *Plant Physiology* 34: 123–127.

Larcher W, Kainmuller C, Wagner J. 2010. Survival types of high mountain plants under extreme temperatures. *Flora* 205: 3–18.

Law M, Collins A. 2013. Getting to know ArcGIS for desktop 3th Edition. ESRI press.

Lawler JJ, Shafer SL, White D, Kareiva P, Maurer EP, Blaustein AR, Bartlein PJ. 2009. Projected climate-induced faunal change in the Western Hemisphere. *Ecology* 90: 588–597.

Ledru MP, Jomelli V, Bremond L, Ortuño T, Cruz P, Bentaleb I, Sylvestre F, Kuentz A, Beck S, Martin C, Paillès C. 2013. Evidence of moist niches in the Bolivian Andes during the mid-Holocene arid period. *The Holocene* 23: 1547–1559.

Legendre P, Legendre LFJ. 2012. Numerical ecology. Elsevier.

Leibold, M. 1996. The niche concept revisited: mechanistic models and community context. *Ecology* 76: 1371–1382.

Leuschner C. 2000. Are high elevations in tropical mountains arid environments

for plants? *Ecology* 81: 1425–1436.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12: 323–339.

Li X, Yang Y, Ma L, Sun X, Yang S, Kong X, Hu X, Yang Y. 2014. Comparative proteomics analyses of *Kobresia pygmaea* adaptation to environment along an elevational gradient on the central Tibetan Plateau. *PLOS One* 9: e98410.

Littlejohn RO, Williams III GJ. 1983. Diurnal and seasonal variations in activity of crassulacean acid metabolism and plant water status in a northern latitude population of *Opuntia erinacea*. *Oecologia* 59: 83–87.

Loik ME, Nobel PS. 1991. Water relations and mucopolysaccharide increases for a winter hardy cactus during acclimation to subzero temperatures. *Oecologia* 88: 340–346.

Loik ME, Nobel PS. 1993a. Exogenous abscisic acid mimics cold acclimation for cacti differing in freezing tolerance. *Plant Physiology* 103: 871–876.

Loik ME, Nobel PS. 1993b. Freezing tolerance and water relations of *Opuntia fragilis* from Canada and the United States. *Ecology* 74: 1722–1732.

Londono C, Cleef A, Madrinán S. 2014. Angiosperm flora and biogeography of the páramo region of Colombia, northern Andes. flora-morphology, distribution, functional ecology of plants. *Flora* 209: 81–87.

Luteyn, JL. 1999. Páramos: A Checklist of Plant Diversity, Geographic Distribution and Botanical Literature. New York, NY: The New York Botanical Garden Press.

Lüttge U. 2004. Ecophysiology of Crassulacean acid metabolism (CAM). *Annals*

of Botany 93: 629–652.

Lüttge U. 2010. Ability of crassulacean acid metabolism plants to overcome interacting stresses in tropical environments. *AoB Plants* 2010: 1–15.

Ma L, Yang L, Zhao J, Wei J, Kong X, Wang C, Zhang X, Yang Y, Hu X. 2015a. Comparative proteomic analysis reveals the role of hydrogen sulfide in the adaptation of the alpine plant *Lamiophlomis rotata* to altitude gradient in the Northern Tibetan Plateau. *Planta* 241: 887–906.

Ma L, Sun X, Kong X, Galvan JV, Li X, Yang S, Yang Y, Yang Y, Hu X. 2015b. Physiological, biochemical and proteomics analysis reveals the adaptation strategies of the alpine plant *Potentilla saundersiana* at altitude gradient of the northwestern Tibetan Plateau. *Journal of Proteomics* 112: 63–82.

MacArthur, R. 1968. The theory of the niche. In Lewontin KC [eds.], *Population Biology and Evolution*, 159–176. Syracuse: Syracuse University Press.

Madriñán S, Cortés AJ, Richardson JE. 2013. Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Frontiers in Genetics* 4: 1–7.

Madriñán S. 2015. A new species of *Puya* (Bromeliaceae) from the páramos near Bogotá, Colombia. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales* 39: 389–398.

Mahalingam S, McClelland GB, Scott GR. 2017. Evolved changes in the intracellular distribution and physiology of muscle mitochondria in high-altitude native deer mice. *The Journal of Physiology* 595: 4785–4801.

Maire V, Wright IJ, Prentice IC, Batjes NH, Bhaskar R, Bodegom PM, Cornwell WK, Ellsworth D, Niinemets Ü, Ordóñez A, Reich PB. 2015. Global effects of soil and

climate on leaf photosynthetic traits and rates. *Global Ecology and Biogeography* 24: 706–717.

Males J, Griffiths H. 2017. Functional types in the Bromeliaceae: relationships with drought resistance traits and bioclimatic distributions. *Functional Ecology* 31: 1868–1880.

Males, J. 2017. Secrets of succulence. *Journal of Experimental Botany* 68: 2121–2134.

Malhi Y, Aragão LE, Galbraith D, Huntingford C, Fisher R, Zelazowski P, Sitch S, McSweeney C, Meir P. 2009. Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences of the United States of America* 106: 20610–20615.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal* 17: 1–10.

Martins E, Hansen T. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *The American Naturalist* 149: 646–667.

Matasci N, Hung LH, Yan Z, Carpenter EJ, Wickett NJ, Mirarab S, Nguyen N, Warnow T, Ayyampalayam S, Barker M, Burleigh JG, Gitzendanner MA, Wafula E, Der JP, dePamphilis CW, Roure B, Philippe H, Ruhfel BR, Miles NW, Graham SW, Mathews S, Surek B, Melkonian M, Soltis DE, Soltis PS, Rothfels C, Pokorny L, Shaw JA, DeGironimo L, Stevenson DW, Villarreal JC, Chen T, Kutchan TM, Rolf M, Baucom RS, Deyholos MK, Samudrala R, Tian Z, Wu X, Sun X, Zhang Y, Wang J, Leebens-Mack J, Wong GK. 2014. Data access for the 1,000 Plants (1KP) project. *GigaScience* 3: 1–10.

Maxwell K, von Caemmerer S, Evans JR. 1997. Is a low internal conductance to CO₂ diffusion a consequence of succulence in plants with crassulacean acid metabolism? *Australian Journal of Plant Physiology* 24: 777–786.

Maxwell K. 2002. Resistance is useful: diurnal patterns of photosynthesis in C₃ and crassulacean acid metabolism epiphytic bromeliads. *Functional Plant Biology* 29: 679–687.

Mc Williams EL. 1970. Comparative rates of dark CO₂ uptake and acidification in the Bromeliaceae, Orchidaceae, and Euphorbiaceae. *Botanical Gazette* 131: 285–290.

McKay JK, Bishop JG, Lin JZ, Richards JH, Sala A, Mitchell-Olds T. 2001. Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proceedings of the Royal Society of London B: Biological Sciences* 268: 1715–1721.

Medina E, Delgado M, Troughton JH, Medina JD. 1977. Physiological ecology of CO₂ fixation in Bromeliaceae. *Flora* 166: 137–152.

Medina E, Delgado M. 1976. Photosynthesis and night CO₂ fixation in *Echeveria columbiana* var. *poellnitz*. *Photosynthetica* 10: 155–163.

Medina E. 1974. Dark CO₂ fixation, habitat preference and evolution within the Bromeliaceae. *Evolution* 28: 677–686.

Meir P, Brando PM, Nepstad D, Vasconcelos S, Costa AC, Davidson E, Almeida S, Fisher RA, Sotta ED, Zarin D, Cardinot G. 2009. The effects of drought on Amazonian rain forests. In Keller M, Bustamante M, Gash J, Silva-Dias P [eds.] *Amazonia and Global Change*, 429–449. Washington, D.C.: American Geophysical Union Press.

Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F, Chang C, Lin C.

2014. e1071: Misc Functions of the Department of Statistics, Probability, R package version 1.6–8.

Middleton N, Thomas D. 1997. *World Atlas of Desertification*, UNEP, 1–182. London: Edward Arnold.

Miller GA. 1986. Pubescence, floral temperature and fecundity in species of *Puya* (Bromeliaceae) in the Ecuadorian Andes. *Oecologia* 70: 155–160.

Ming R, VanBuren R, Wai CM, Tang H, Schatz MC, Bowers JE, Lyons E, Wang ML, Chen J, Biggers E, Zhang J, Huang L, Zhang L, Miao W, Zhang J, Ye J, Miao Z, Lin Z, Wang H, Zhou H, Yim WC, Priest HD, Zheng C, Woodhouse M, Edger PP, Guyot R, Zhu X, Yang X, Smith JAC, Cushman JC, Paull RE, Yu Q. 2015. The pineapple genome and the evolution of CAM photosynthesis. *Nature Genetics* 47: 1435–1442.

Moncrieff GR, Hickler T, Higgins SI. 2015. Intercontinental divergence in the climate envelope of major plant biomes. *Global Ecology and Biogeography*. 24: 324–334.

Montes C, Cardona A, Jaramillo C, Pardo A, Silva JC, Valencia V, Ayala C, Pérez–Ángel LC, Rodríguez-Parra LA, Ramírez V, Niño H. 2015. Middle Miocene closure of the Central American seaway. *Science* 348: 226–229.

Mort ME, Soltis DE, Soltis PS, Francisco-Ortega J, Santos-Guerra A. 2001. Phylogenetic relationships and evolution of Crassulaceae inferred from *matK* sequence data. *American Journal of Botany* 88:76–91.

Nelson EA, Sage RF. 2008. Functional constraints of CAM leaf anatomy: tight cell packing is associated with increased CAM function across a gradient of CAM expression. *Journal of Experimental Botany* 59, 1841–1850.

Nelson EA, Sage TL, Sage RF. 2005. Functional leaf anatomy of plants with

crassulacean acid metabolism. *Functional Plant Biology* 32: 409–419.

Nevado B, Atchison GW, Hughes CE, Filatov DA. 2016. Widespread adaptive evolution during repeated evolutionary radiations in New World lupins. *Nature Communications*. 7: 1–9.

Neves-Chaves CJ, Leal BS, de Lemos-Filho JP. 2015. Temperature modulation of thermal tolerance of a CAM-tank bromeliad and the relationship with acid accumulation in different leaf regions. *Physiologia Plantarum* 154: 500–510.

New M, Lister D, Hulme M, Makin I. 2002. A high-resolution data set of surface climate over global land areas. *Climate Research* 21: 1–25

Nimmo HG. 2000. The regulation of phosphoenolpyruvate carboxylase in CAM plants. *Trends in Plant Science* 5: 75–80.

Nobel PS, Castañeda M, North G, Pimienta-Barrios E, Ruiz A. 1998. Temperature influences on leaf CO₂ exchange, cell viability and cultivation range for *Agave tequilana* *Journal of Arid Environments* 39: 1–9.

Nobel PS, Hartsock TL. 1986. Short-term and long-term responses of Crassulacean acid metabolism plants to elevated CO₂. *Plant Physiology* 82: 604–606.

Nobel PS, Smith SD. 1983. High and low temperature tolerances and their relationships to distribution of agaves *Plant, Cell and Environment* 6: 711–719.

Nobel PS. 1980a. Interception of photosynthetically active radiation by cacti of different morphology. *Oecologia* 45: 160–6.

Nobel PS. 1980b. Influences of minimum stem temperatures on ranges of cacti in southwestern United States and central Chile. *Oecologia* 47: 10–15.

Nobel PS. 1981. Influence of freezing temperatures on a cactus, *Coryphantha*

vivipara. *Oecologia* 48: 194–198.

Nobel PS. 1982. Low-temperature tolerance and cold hardening of cacti. *Ecology* 63: 1650–1656.

Nobel PS. 1988. Environmental biology of agaves and cacti. Cambridge, UK: Cambridge University Press.

O'Leary B, Park J, Plaxton WC. 2011. The remarkable diversity of plant PEPC (phosphoenolpyruvate carboxylase): recent insights into the physiological functions and post-translational controls of non-photosynthetic PEPCs. *Biochemical Journal*. 436: 15–34.

O'Dea A, Lessios HA, Coates AG, Eytan RI, Restrepo-Moreno SA, Cione AL, Collins LS, de Queiroz A, Farris DW, Norris RD, Stallard RF. 2016. Formation of the Isthmus of Panama. *Science Advances* 2: e1600883.

O'Donnell MS, and Ignizio DA. 2012. Bioclimatic predictors for supporting ecological applications in the conterminous United States: U.S. Geological Survey Data Series 691: 1–10.

O'Leary MH, Osmond CB. 1980. Diffusional contribution to carbon isotope fractionation during dark CO₂ fixation in CAM plants. *Plant Physiology* 66: 931–934.

Oakley JC, Sultmanis S, Stinson CR, Sage TL, Sage RF. 2014. Comparative studies of C₃ and C₄ *Atriplex* hybrids in the genomics era: physiological assessments. *Journal of Experimental Botany*. 65: 3637–3647.

Ogburn RM, Edwards EJ. 2015. Life history lability underlies rapid climate niche evolution in the angiosperm clade Montiaceae. *Molecular Phylogenetics and Evolution* 92: 181–192.

Osmond CB 2007. Crassulacean acid metabolism: now and then. *Progress in Botany* 69: 3–32.

Osmond CB. 1978. Crassulacean acid metabolism: a curiosity in context. *Annual Review of Plant Physiology* 29: 379–414.

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, Orvis J, Haas B, Wortman J, Buell CR. 2007. The TIGR Rice Genome Annotation Resource: improvements and new features., *Nucleic Acids Research*. Database issue: D883–887.

Owen NA, Fahy KF, Griffiths H. 2016. Crassulacean acid metabolism (CAM) offers sustainable bioenergy production and resilience to climate change. *GCB Bioenergy* 8: 737–749.

Owen NA, Griffiths H. 2014. Marginal land bioethanol yield potential of four crassulacean acid metabolism candidates (*Agave fourcroydes*, *Agave salmiana*, *Agave tequilana* and *Opuntia ficus-indica*) in Australia. *GCB Bioenergy*. 6: 687–703.

Pagel M. 1997. Inferring evolutionary processes from phylogenies. *Zoologica Scripta* 26: 331–348.

Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401: 887–884.

Palma-Silva C, Wendt T, Pinheiro F, Barbará T, Fay MF, Cozzolino S, Lexer C. 2011. Sympatric bromeliad species (*Pitcairnia* spp.) facilitate tests of mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs. *Molecular Ecology*. 20: 3185–3201.

Parolo G, Rossi G, Ferrarini A. 2008. Toward improved species niche modelling: *Arnica montana* in the Alps as a case study. *Journal of Applied Ecology*. 45: 1410–1418.

Paulus JK, Schlieper D, Groth G. 2013. Greater efficiency of photosynthetic carbon fixation due to single amino-acid substitution. *Nature Communications* 4: 1–7.

Pebesma EJ, Bivand RS. 2005. Classes and Methods for Spatial Data in R. *R News*. 5: 9–13.

Peterson AT. 2003. Predicting the geography of species invasions via ecological niche modeling. *The Quarterly Review of Biology* 78: 419–433.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modeling* 190: 231–259.

Pierce S, Winter K, Griffiths H. 2002a. Carbon isotope ratio and the extend of daily CAM use by Bromeliaceae. *New Phytologist* 155: 75–84.

Pierce S, Winter K, Griffiths H. 2002b. The role of CAM in high rainfall cloud forests: an *in situ* comparison of photosynthetic pathways in Bromeliaceae. *Plant Cell and Environment* 25: 1181–1189.

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2017. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-131.

Prášil I, Zámečník J. 1998. The use of a conductivity measurement method for assessing freezing injury: I. Influence of leakage time, segment number, size and shape in a sample on evaluation of the degree of injury. *Environmental and Experimental Botany* 40: 25–42.

Quan C, Han S, Utescher T, Zhang C, Liu YS. 2013. Validation of temperature–precipitation based aridity index: Paleoclimatic implications. *Palaeogeography, Palaeoclimatology, Palaeoecology*. 386: 86–95.

Quezada IM, G Zotz, E Gianoli. 2014. Latitudinal variation in the degree of

crassulacean acid metabolism in *Puya chilensis*. *Plant Biology* 16: 848–852.

Quezada IM, Saldaña A, Gianoli E. 2017. Divergent Patterns of Selection on Crassulacean Acid Metabolism Photosynthesis in Contrasting Environments. *International Journal of Plant Sciences*. 178: 398–405.

R Development Core Team. 2017. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria

Rada F, García-Núñez C, Boero C, Gallardo M, Hilal M, Gonzalez J, Prado F, Liberman-Cruz M, Azocar A. 2001. Low-temperature resistance in *Polylepis tarapacana*, a tree growing at the highest altitudes in the world. *Plant, Cell and Environment* 24: 377–381.

Rada F, Goldstein G, Azócar A, Meinzer, F. 1985. Freezing avoidance in Andean giant rosette plants. *Plant, Cell and Environment* 8: 501–507.

Rambaut A, Drummond AJ. 2007. Tracer v1. 4: MCMC trace analyses tool. Available at <http://tree.bio.ed.ac.uk/software/tracer>.

Rao CR. 1964. The use and interpretation of principal component analysis in applied research. *Sankhyā: The Indian Journal of Statistics Series A* 26: 329–358.

Rentsch JD, Leebens-Mack J. 2012. Homoploid hybrid origin of *Yucca gloriosa*: intersectional hybrid speciation in *Yucca* (Agavoideae, Asparagaceae). *Ecology and Evolution* 2: 2213–2222.

Rohlf FJ. 2001. Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution* 55: 2143–2160.

Rohlf RV, Nielsen R. 2015. Phylogenetic ANOVA: The Expression Variance and Evolution (EVE) model for quantitative trait evolution. *Systematic Biology* 64: 695–708.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Rosnow JJ, Edwards GE, Roalson EH. 2014. Positive selection of Kranz and non-Kranz C₄ phosphoenolpyruvate carboxylase amino acids in Suaedoideae (Chenopodiaceae). *Journal of Experimental Botany*. 65: 3595–3607.

Rundel PW, Dillon MO. 1998. Ecological patterns in the Bromeliaceae of the lomas formations of Coastal Chile and Peru. *Plant Systematics and Evolution* 212: 261–278.

Rundel PW, Smith AP, Meinzer FC. 1994. Tropical alpine environments. Plant form and function. Cambridge: Cambridge University Press.

Rundel PW. 1994. Tropical alpine climates. Pages 21–44 in Rundel PW, Smith AP, Meinzer FC, [eds.], *Tropical alpine environments: plant form and function*. Cambridge: University of Cambridge Press.

Russell V. Lenth. 2016. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software* 69: 1–33.

Sarkar D. 2008. Lattice: multivariate data visualization with R. Springer Science & Business Media. (Online resource)

Schulte K, Barfuss MH, Zizka G. 2009. Phylogeny of Bromelioideae (Bromeliaceae) inferred from nuclear and plastid DNA loci reveals the evolution of the tank habit within the subfamily. *Molecular Phylogenetics and Evolution* 51: 327–339.

- Schulte K, Silvestro D, Kiehlmann E, Vesely S, Novoa P, Zizka G. 2010. Detection of recent hybridization between sympatric Chilean *Puya* species (Bromeliaceae) using AFLP markers and reconstruction of complex relationships. *Molecular Phylogenetics and Evolution* 57:1105–1119.
- Scott GR, Elogio TS, Lui MA, Storz JF, Chevignon ZA. 2015. Adaptive modifications of muscle phenotype in high-altitude deer mice are associated with evolved changes in gene regulation. *Molecular Biology and Evolution*. 32:1962–1976.
- Segurado P, Araújo MB. 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography* 31:1555–1568.
- Segurado PA, Araújo MB, Kunin WE. 2006. Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology* 43: 433–444.
- Shapiro SS, Wilk MB. 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52: 591–611.
- Siddiqui KS, Cavicchioli R. 2006. Cold-adapted enzymes. *Annual Review of Biochemistry* 75: 403–433.
- Siddiqui KS. 2015. Some like it hot, some like it cold: Temperature dependent biotechnological applications and improvements in extremophilic enzymes. *Biotechnology Advances*. 33: 1912–1922.
- Sierra-Almeida A, Reyes-Bahamonde C, Cavieres LA. 2016. Drought increases the freezing resistance of high-elevation plants of the Central Chilean Andes. *Oecologia* 181:1011–1023.
- Silsbe GM, Malkin SY. 2015. phytotools: Phytoplankton Production Tools. The R Foundation for Statistical Computing, Vienna <http://cran.r-project.org>.

Silvera K, Lasso E. 2016. Ecophysiology and Crassulacean Acid Metabolism of Tropical Epiphytes. In Goldstein g, Santiago LS. [eds.], *Tropical Tree Physiology*, 25–43. Switzerland: Springer International Publishing.

Silvera K, Neubig KM, Whitten WM, Williams NH, Winter K, Cushman JC. 2010a. Evolution along the crassulacean acid metabolism continuum. *Functional Plant Biology* 37: 995–1010.

Silvera K, Santiago LS, Cushman JC, Winter K. 2009. Crassulacean acid metabolism and epiphytism linked to adaptive radiations in the Orchidaceae. *Plant Physiology* 149: 1838–1847.

Silvera K, Santiago LS, Cushman JC, Winter K. 2010b. The incidence of crassulacean acid metabolism in the Orchidaceae derived from carbon isotope ratios: a checklist of the flora of Panama and Costa Rica. *Botanical Journal of the Linnean Society* 163: 194–222.

Silvera K, Santiago LS, Winter K. 2005. Distribution of crassulacean acid metabolism in orchids of Panama: evidence of selection for weak and strong modes. *Functional Plant Biology* 32: 397–407.

Silvera K, Winter K, Rodriguez BL, Albion RL, Cushman JC. 2014. Multiple isoforms of phosphoenolpyruvate carboxylase in the Orchidaceae (subtribe Oncidiinae): implications for the evolution of crassulacean acid metabolism. *Journal of Experimental Botany* 65: 3623–3636.

Silvestro D, Zizka G, Schulte K. 2013. Disentangling the effects of key innovations on the diversification of Bromelioideae (Bromeliaceae). *Evolution* 68: 163–175.

Simpson BB. 1983. An historical phytogeography of the high Andean flora. *Revista Chilena de Historia Natural*. 56: 109–122.

SipesDL, Ting IP. 1985. Crassulacean acid metabolism and crassulacean acid metabolism modifications in *Peperomia camptotricha*. *Plant Physiology* 77: 59–63.

Sklenář P, Dušková E, Balslev H. 2011. Tropical and temperate: evolutionary history of páramo flora. *The Botanical Review* 77: 71–108.

Sklenář P, Hedberg I, Cleef AM. 2014. Island biogeography of tropical alpine floras. *Journal of Biogeography* 41: 287–297.

Sklenář P, Kučerová A, Macek P, Mackova J. 2010. Does plant height determine the freezing resistance in the páramo plants? *Austral Ecology* 35: 929–934.

Sklenář P, Kučerová A, Macek P, Macková J. 2012. The frost-resistance mechanism in páramo plants is related to geographic origin. *New Zealand Journal of Botany* 50: 391–400.

Smith JAC, Bryce JH. 1992. Metabolite compartmentation and transport in CAM plants. In Tobin AK [ed], *Plant organelles*, 141–167. Cambridge University Press, Cambridge.

Smith JAC, Griffiths H, Lüttge U. 1986. Comparative ecophysiology of CAM and C₃ bromeliads. I. The ecology of the Bromeliaceae in Trinidad. *Plant. Cell and Environment* 9: 359–376.

Smith JAC, Lüttge U. 1985. Day-night changes in leaf water relations associated with the rhythm of crassulacean acid metabolism in *Kalanchoë daigremontiana* *Planta* 163: 272–282.

Smith JAC, Winter K. 1996. Taxonomic distribution of crassulacean acid metabolism. In: Winter K, Smith JAC [eds.], *Crassulacean acid metabolism: biochemistry, ecophysiology and evolution*, 427–436. Heidelberg: Springer-Verlag.

Smith JAC. 1989. Epiphytic bromeliads. In: Lüttge U, [ed.], *Vascular plants as epiphytes: evolution and ecophysiology*, 109–138. Berlin: Springer-Verlag.

Smith LB, Downs RJ. 1974. *Flora Neotropica, Monograph No. 14, Part 1, Pitcairnioideae (Bromeliaceae)*. New York: Hafner Press.

Smith LB, Downs RJ. 1977. *Flora Neotropica, Monograph No. 14, Part 2, Tillandsioideae (Bromeliaceae)*. New York: Hafner Press.

Smith LB, Downs RJ. 1979. *Flora Neotropica, Monograph No. 14, Part 3, Bromelioideae (Bromeliaceae)*. New York: Hafner Press.

Smith SA, Donoghue MJ. 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science* 322: 86–89.

Somero GN. 2004. Adaptation of enzymes to temperature: searching for basic “strategies”. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*. 139: 321–333.

Stamatakis E. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.

Stankowski S, Streisfeld MA. 2015. Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proceedings of the Royal Society of London B: Biological Sciences* 282: 20151666.

Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. 2008. Conditional

variable importance for random forests. *BMC Bioinformatics* 9: 1–11.

Strobl C, Boulesteix AL, Zeileis A, Hothorn T. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* 8: 1–21.

Takeya M, Hirai MY, Osanai T. 2017. Allosteric Inhibition of Phosphoenolpyruvate Carboxylases is Determined by a Single Amino Acid Residue in Cyanobacteria. *Nature Scientific Reports* 7: 41080.

Tavaré S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* 17: 57–86.

Tomlinson PB. 1970. Monocotyledons—towards an understanding of their morphology and anatomy. *Advances in Botanical Research* 3: 207–292.

Torgo L. 2010. *Data Mining using R: learning with case studies*. Boca Raton Florida: CRC Press.

Tovar C, Arnillas CA, Cuesta F, Buytaert W. 2013. Diverging responses of tropical Andean biomes under future climate conditions. *PLOS One* 8: e63634.

Trabucco A, Zomer RJ .2010. *Global Soil Water Balance Geospatial Database*. CGIAR Consortium for Spatial Information. Published online, available from the CGIAR-CSI GeoPortal at: <http://www.cgiar-csi.org>

Vandermeer J. 1972. Niche theory. *Annual Review of Ecology, Evolution, and Systematics* 3: 107–132.

Varadarajan GS, Gilmartin AJ. 1988. Seed morphology of the subfamily Pitcairnioideae (Bromeliaceae) and its systematic implications. *American Journal of Botany* 1: 808–818.

Varadarajan GS. 1990. Patterns of geographic distribution and their implications on the phylogeny of *Puya* (Bromeliaceae). *Journal of the Arnold Arboretum* 1: 527–552.

Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S*. Fourth Edition. New York: Springer.

Vuille M, Bradley RS, Keimig F. 2000. Climate variability in the Andes of Ecuador and its relation to tropical Pacific and Atlantic sea surface temperature anomalies. *Journal of Climate* 13: 2520–2535.

Wagner J, Larcher W. 1981. Dependence of CO₂ gas exchange and acid metabolism of the alpine CAM plant *Sempervivum montanum* on temperature and light. *Oecologia* 50: 88–93.

White J, Prell J, James EK, Poole P. 2007. Nutrient sharing between symbionts. *Plant Physiology* 144: 604–614.

Whittaker RH, Levin SA, Root RB. 1973. Niche, habitat, ecotope. *The American Naturalist* 95: 321–338.

Wickham H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wilson AM, Jetz W. 2016. Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. *PLOS Biology* 14: e1002415.

Wilson MF, O’Connell B, Brown C, Guinan JC, Grehan AJ. 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy* 30: 3–5.

Winter K, Aranda J, Holtum JAM. 2005. Carbon isotope composition and water-use efficiency in plants with crassulacean acid metabolism. *Functional Plant Biology* 32: 381–388.

Winter K, Holtum JA. 2015. Cryptic crassulacean acid metabolism (CAM) in *Jatropha curcas*. *Functional Plant Biology* 42: 711–717.

Winter K, Holtum JAM, Smith JAC. 2015. Crassulacean acid metabolism: a continuous or discrete trait? *New Phytologist* 208: 73–78.

Winter K, Holtum JAM. 2002. How closely do the $\delta^{13}\text{C}$ values of CAM plants reflect the proportion of CO_2 fixed during day and night? *Plant Physiology* 129: 1843–1851.

Winter K, Holtum JAM. 2014. Facultative crassulacean acid metabolism (CAM) plants: powerful tools for unravelling the functional elements of CAM photosynthesis. *Journal of Experimental Botany* 65: 3425–3441.

Winter K, Smith JAC. 1996. *Crassulacean Acid Metabolism: Biochemistry, Ecophysiology and Evolution*. Heidelberg: Springer-Verlag.

Winter K, von Willert DJ. 1972. NaCl-induzierter Crassulaceen-Säurestoffwechsel bei *Mesembryanthemum crystallinum*. *Z. Pflanzenphysiol.* 67: 166–170.

Winter K. 1985. Crassulacean acid metabolism. In: Barber J, Baker NR [eds.] *Photosynthetic mechanisms and the environment*, 329–387. Amsterdam: Elsevier, Amsterdam.

Wisniewski M, Gusta L, Neuner G. 2014. Adaptive mechanisms of freeze avoidance in plants: a brief update. *Environmental and Experimental Botany* 99: 133–140.

Wu MX, Wedding RT. 1987. Temperature effects on phosphoenolpyruvate carboxylase from a CAM and a C₄ plant: a comparative study. *Plant Physiology* 85: 497–501.

Yamori W, Hikosaka K, Way DA. 2014. Temperature response of photosynthesis in C₃, C₄, and CAM plants: temperature acclimation and temperature adaptation. *Photosynthesis Research* 119: 101–17.

Yang Z, Nielsen R, Goldman N, Pedersen AMK. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155: 431–449.

Yang Z, Wong WS, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution* 22: 1107–1118.

Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular Biology and Evolution* 15: 568–573.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24: 1586–1591.

Yao AYM. 1974. Agricultural potential estimated from the ratio of actual to potential evapotranspiration. *Agricultural Meteorology*. 13: 405–417.

Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG, McGlinn DJ, O'Meara BC, Moles AT, Reich PB, Royer DL. 2014. Three keys to the radiation of angiosperms into freezing environments. *Nature Environments* 506: 89–92.

Zhang JZ. 2003. Evolution by gene duplication: an update. *Trends in Ecology and Evolution* 18: 292–298.

Zhang L, Chen F, Zhang GQ, Zhang YQ, Niu S, Xiong JS, Lin Z, Cheng ZM, Liu ZJ. 2006. Origin and mechanism of crassulacean acid metabolism in orchids as

implied by comparative transcriptomics and genomics of the carbon fixation pathway. *The Plant Journal* 86: 175–185.

Zhao SY, Chen LY, Muchuku JK, Hu GW, Wang QF. 2016. Genetic adaptation of giant lobelias (*Lobelia aberdarica* and *Lobelia telekii*) to different altitudes in east african mountains. *Frontiers in Plant Science* 7: 488.

Zizka G, Horres R, Nelson EC, Weising K. 1999. Revision of the genus *Fascicularia* Mez (Bromeliaceae). *Botanical Journal of the Linnean Society* 129: 315–322.

Zizka G, Schmidt M, Schulte K, Novoa P, Pinto R, König K. 2009. Chilean Bromeliaceae: diversity, distribution and evaluation of conservation status. *Biodiversity and Conservation* 18: 2449–2471.

Zizka G, Schneider JV, Schulte K, Novoa P. 2013. Taxonomic revision of the Chilean *Puya* species (Puyoideae, Bromeliaceae), with special notes on the *Puya alpestris*-*Puya berteroniana* species complex. *Brittonia* 65: 387–407.

Zizka G, Trumpler K, Zöllner O. 2002. Revision of the genus *Ochagavia* (Bromeliaceae, Bromelioideae). *Willdenowia* 32: 331–350.

Zomer RJ, Bossio DA, Trabucco A, Yuanjie L, Gupta DC, Singh VP. 2007. *Trees and water: smallholder agroforestry on irrigated lands in Northern India*. Colombo: IWMI Press.

Zomer RJ, Trabucco A, Bossio DA, van Straaten O, Verchot LV. 2008. Climate change mitigation: a spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agricultural Ecosystems and Environment* 126: 67–80.

Zotz G, Ziegler H. 1997. The occurrence of crassulacean acid metabolism among vascular epiphytes from Central Panama. *New Phytologist* 137: 223–229.