Exploiting the Potential of *Agave* for Bioenergy in Marginal Lands

Dalal Bader Al Baijan

A Thesis submitted for the Degree of

Doctor of Philosophy

School of Biology

Newcastle University

July 2015

Declaration

I hereby certify that this thesis is the result of my own investigations and that no part of it has been submitted for any degree other than Doctor of Philosophy at the Newcastle University. All references to the work of others have been duly acknowledged.

Dalal B. Al Baijan

“It is not the strongest of the species that survives, nor the most intelligent, but the one most responsive to change”

Charles Darwin

Origin of Species, 1859

Abstract

Drylands cover approximately 40% of the global land area, with minimum rainfall levels, high temperatures in the summer months, and they are prone to degradation and desertification. Drought is one of the prime abiotic stresses limiting crop production. *Agave* plants are known to be well adapted to dry, arid conditions, producing comparable amounts of biomass to the most water-use efficient C3 and C4 crops but only require 20% of water for cultivation, making them good candidates for bioenergy production from marginal lands. *Agave* plants have high sugar contents, along with high biomass yield. More importantly, *Agave* is an extremely water-use efficient (WUE) plant due to its use of Crassulacean acid metabolism. Most of the research conducted on *Agave* has centered on A. *tequilana* due to its economic importance in the tequila production industry. However, there are other species of *Agave* that display higher biomass yields compared to *A. tequilana*. These include A. *mapisaga* and A. *salmiana* and A. *fourcroydes Lem* has been reported to possess high fructan content making it a promising plant for biofuel feedstock. Also, fructans act as osmo-protectants by stabilizing membranes during drought and other abiotic stress.

This project set out to examine several hypotheses. In the first experimental chapter (Chapter 2), the central aim was to start identifying traits for the improvement of *Agave* species for biomass production on arid lands by first examining if the capacity of CAM, and fructan accumulation are linked traits. To address this question 3 species of *Agave* varying in succulence were compared under different water regimes. Measurements were made of leaf, gas exchange and titratable acidities as markers of CAM and of soluble sugar and fructan content using high performance liquid chromatography (HPLC). High leaf succulence is associated with increased magnitude of CAM, manifested as higher ΔH^+ and nocturnal CO_2 uptake and fructan accumulation also increased with leaf succulence in *Agave*. Sucrose provided most, if not all of the substrate required for dark CO_2 uptake. At the leaf level, highest CAM activity was found in the tip region whilst most fructan accumulation occurred in the base of the leaf. These results indicate that CAM and fructan accumulation are subject to contrasting anatomical and physiological control processes.

In Chapter 3, the aim was to test 4 hypotheses relating to succulence and biochemical capacity for C3 and C4 carboxylation in Agave. The first hypothesis tested the abundance of PEPC and its variation between species in relation to leaf succulence and age and will vary along the leaf, in line with differences in CAM activity. The second hypothesis looked into the abundance of Rubisco and Rubisco activase and its variation between species in relation to leaf succulence and age and will vary along the leaf, in line with differences in CAM activity. The third hypothesis the more succulent *Agave* species, drought will have less impact on the abundance of PEPC, Rubisco and Rubisco activase compared to the less succulent species. And the abundance of Rubisco activase will vary over the diel cycle, particularly in leaves of more succulent species of *Agave*. Results showed that leaf succulence influenced the abundance of PEPC. Thus, the optimal anatomy for nocturnal malic acid accumulation is accompanied by high PEPC abundance in leaves with higher vacuolar storage capacity. In contrast, the abundances of Rubisco and Rubisco activase showed an inverse relationship to succulence and CAM activity.

The aim of Chapter 4, was to identify other species of Agave that could be exploited as sources of biofuel from semi-arid marginal lands. Some 14 different species of Agave that showed varying levels of succulence were compared, evaluating the capacity for CAM, fructan content, carbohydrate composition, osmotic pressure and the relationship with succulence. Results demonstrated that Inter-specific variations in the magnitude of expression of CAM in *Agave* are dependent on leaf succulence. Also, *Agave* displays flexibility in the use of carbohydrate source pools to sustain dark CO_2 uptake. Some species appear to use fructans and others sucrose as substrate for dark CO_2 uptake.

The final experimental Chapter's aim was to develop a method to identify vacuolar sugar transporters in *Agave* related to sucrose turnover and fructan accumulation. First, identifying the tonoplast by testing activity of ATPase and PP_iase of leaf vesicles of *Agave Americana marginata*, and its sensitivity to inhibition by known ATPase inhibitors. Second, was to use a proteomics approach, analysing of the purified tonoplast involved fractionation of the proteins by SDS-PAGE and analysis by LC-MS/MS, to identify vacuolar sugar transporter proteins which are hypothesized to play a key regulatory role in determining sucrose turnover for CAM and fructan accumulation and as such,

could represent future targets for genetic engineering of increased sugar content for plants grown for bioenergy. The capacity of the vacuole as a sink for carbohydrate maybe an important determinant of CAM expression and has important implications for plant growth and productivity. Combining tonoplast proteomics with the interrogation of diel transcriptome data is a potentially powerful approach to identify candidate vacuolar sugar transporters in *Agave*.

Acknowledgements

The process of giving credit to all the people who have contributed to my research is essential to the completion of this thesis. The effort to put it in written form is always filled with a certain apprehension that someone will be inadvertently be left out. Sometimes historical quotations can provide the needed inspiration. There is a saying attributed to the Prophet Muhammed that says, "**He has not** thanked **Allah who has not** thanked **people** ".

There are many people and institutions that I should like to thank for making it possible to complete my thesis. I am truly thankful for their kind support and thoughtful assistance.

First and foremost, I would like to thank my supervisor Prof Anne M Borland and Prof Jeremy Barnes for their encouragement, technical expertise, enthusiasm and development throughout my study in the past 4 years. I'm very grateful for their guidance and they never hesitated to provide me with their help in academic and personal matters.

Within the School of Biology I have received encouragement, friendship and support from many people, which contributed to an enjoyable experience and friendly environment in Newcastle University. In particular, I would like to thank Dr Taybi, Dr Aayush Sharma, Gillian Davis, Dr Eileen Power, Helen Martin and Dr Matthew Peake for their advice.

A mention must also be given to people who inspired me and contributed to help me develop new skills and knowledge and who collaborated with me in this thesis; Dr Achim Treumann and Samantha Baker from NUPPA.

My acknowledgement would not be complete without mentioning Shreya and Aysha. Thank you for your friendship, and sharing with me the ups and downs of this PhD, and being my gym buddies.

This thesis wouldn't be possible without the love, support and blessings of my parents Sandra and Bader, my twin sister Dina, my aunt Marye and my nephews Mohammed and Bader Jr; I dedicate this achievement to them. I'm very thankful to my husband Suhail for his love, patience and constant support through the distance. Thank you for being a positive force in my life. I would wish to thank Judy for her visits to Newcastle.

My sincere thank you goes to HE Sheikha Hussah Sabah al-Salim Al-Sabah, Director-General of Dar al-Athar al-Islamiyyah (Islamic Art Museum, Kuwait) not only for her support and constant encouragement but also for being an impressive role-model

This thesis was funded by Kuwait Institute for Scientific Research.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 5

Abbreviations

APS	Ammonium persulphate
ATP	Adenosine 5’-triphosphate
ATPase	Adenosine 5’-triphosphatase
BSA	Bovine serum albumin
C3	Photosynthesis with RuBisCO as primary fixator of CO2
C4	Photosynthesis with PEPC as primary fixator of CO2
CAM	Crassulacean acid metabolism
DP	Degree of polymerization
DTT	Dithiothreitol
E-64	Trans-epoxysuccinyl-L-leucylamido-(4-guanidino) butane
ECL	Enhanced chemiluminescence
EDTA	Ethylenediamine tetracetate
EPI	Environment productivity index
ESI	Electrospray ionization
FGEC	First generation bioenergy crops
GHG	Green-house gases
GNP	Gross national product
H^+-ATPase	Vacuolar type H^+ proton adenosine triphosphatase
H+-PP_iase	Vacuolar type H^+ proton pyrophosphatase
HPLC	High-performance liquid chromatography
kDa	Kilo Dalton
KWH	Kilo-watt
LCA	Life cycle analysis
LC-MS	Liquid-chromatography Mass spectrometry
LWGB	Lower gel buffer
MDH	Malate dehydrogenase
MEW	Ministry of Electricity and water, Kuwait

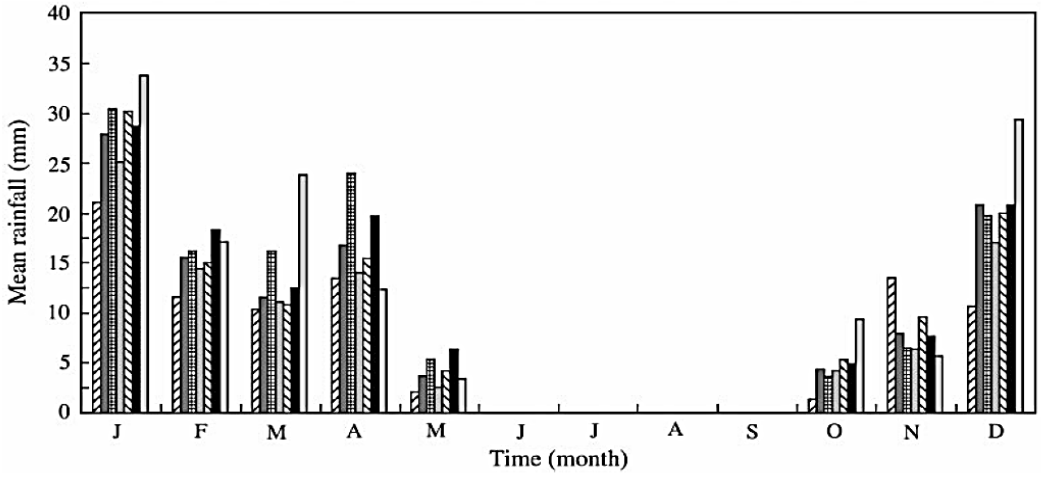
MOP	Ministry of Planning, Kuwait
MST	Monosaccharide transporter
MW	Mega-watt
NAD^+	Nicotinaamide adenine dinucleotide (oxidised form)
NADH	Nicotinaamide adenine dinucleotide (reduced form)
NAD-ME	NAD^+-malic enzyme (EC 1.1.1.39)
$NADP^+$	Nicotinaamide adenine dinucleotide phosphate (oxidised form)
NADPH	Nicotinaamide adenine dinucleotide phosphate (reduced form)
NADP-ME	$NADP^+$- malic enzyme
NCBI	National centre for biotechnology information
Nkat	Nano katal
OAA	Oxaloacetate
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvtae
PEPC	Phosphoenolpyrovate carboxylase
PMSF	Phenylmethylsulphonyl fluoride
Ppm	Parts per million
PVP	Polpyrrolidone
RuBisCO	Ribulose 1, 5-biphosphate carboxylase/oxygenase(EC.4.1.1.39)
SDS	Sodium dodecyl sulphate
SGEC	Second generation bioenergy crops
TA	Titratable acidity
TBE	Tris-borate EDTA buffer
TBS	Tris buffered saline
TBST	Tris buffered saline +Tween 20
TEMED	N,N,N',N'-tetra-methyl-ethylenediamine
TMT	Tonoplast monosaccharide transporter
Tris	2-amino-2(hydroxymethyl)1'3 propanediol

Tris-HCl	Tris-Hydrochloride
Tween-20	Polyoxyethylenesorbitan monolaurate
UPGB	Upper gel buffer (polyacrylamide gel electrophoresis)
WUE	Water use efficiency
$\delta^{13}C$	Carbon isotope ratio ($‰_0$)

Chapter 1 General Introduction

1. Introduction

Kuwait is located in the north eastern part of the Arabian Peninsula; between 28° 33N and 30° 05N latitude and 46° 33N and 48° 30E longitude. The total land area of the mainland and nine islands is approximately 17,344 km^2 (Roy and Grealish, 2004), and they are surrounded by the Arabian Gulf on the East, Iraq on the north and Kingdom of Saudi Arabia from the West and South. Summers in Kuwait are hot and dry, ranging between 42°-49°C; winters are short, from December to February, and cool, averaging 10°-30°C (MOP, 1998), with limited rainfall. Annual rainfall is about 120 mm and mean annual rainfall is 115 mm, with great variability from year to year (28-260mm) and from place to place (Roy and Grealish, 2004). Some 80% of rainfall occurs in the winter months from December through March. Evaporation ranges from 3.0 mm d^{-1} in January to 14.1mm d^{-1} in July. The relative humidity is generally low, and strong, dry and hot, north-westerly winds prevail during summer, particularly in the months of June and July (Roy and Grealish, 2004). These climatic conditions pose a number of challenges for sustainable agriculture. This thesis examines the physiological and biochemical characteristics of a drought tolerant plant genus (*Agave*) that has potential to be cultivated for the production of biomass and high value products under climatic conditions of high temperatures and low water availability.



Distribution of monthly rainfall in Kuwait. (▨) Umm Al-Aish; (▭) Al-Omariyah; (▦) Ahmadi; (▭) Mena Al-Ahmadi; (▧) Shuwaikh; (■) KuwaitAP; (▭) Failaka Island.

Figure 1.1 Distribution of monthly rainfall covering different areas in Kuwait (Nasrallah *et al.*, 2001)

1.1 Agriculture in Kuwait

Agricultural production in Kuwait is associated with two areas; Abdali farms in the north and Wafra farms in the south of Kuwait, both of which use open fields and agricultural units in crop production. Agricultural production is very low, representing less than 0.4% of the country's gross national product (GNP) (Omar, 2001). The country produces only about 20% of its need for a few selected vegetables, mainly winter cropping of vegetables production and some summer crops such as water melon and sweet melon and farming of semi perennial crops such as *alfalfa*. As a result, the country imports a great majority of its food for both human and animal consumption. Kuwait has no food security and is unable to exploit the business and commercial potentials with its agricultural production base. The future expansion of the agriculture sector in Kuwait is guided by the Agricultural Master Plan (1995-2015), with a major emphasis on sustainable utilization of available land and water resources in agriculture (Roy and Grealish, 2004).

There are many constrains to agricultural development in Kuwait, some of which are outlined below.

1.1.1 Physical constraints

Water: Ground water is brackish, with dissolved salt content up to 9000 ppm. The use of brackish water for irrigation imposes physiological stress in plants and increases soil salinity. Over 60 % of the field irrigation and all of the landscape irrigation in Kuwait is from groundwater (Abd El-Hafez, 1990). Two types of treated waste water are suitable for irrigation: municipal wastewater and industrial waste water. The quality of the municipal wastewater has markedly improved with the opening of the tertiary treatment plant in June 1985. Lately, desalinated water (fresh water) has only been used for protected agriculture, using green houses.

Soil: The native soils are predominately sandy with low cation-exchange capacity and low organic matter, low water holding capacity and low available phosphorus. When a gatch layer, which is a local name of consolidated sediment of a massive calcrete type found in many parts of Kuwait at variable depths but generally, about 2m below the surfaces is present, it obstructs natural drainage and causes water logging and salinity problems.

Harsh weather: High summer temperatures, low rainfall, high evaporation rates and sand and dust storms.

The Ministry of Planning (1988) recorded several types of crops being cultivated in Kuwait with the following percentage production rates: fruits and leafy vegetables 26%; bulbs and tubers 12 %; pulses 51%; agronomic crops 8%; and green fodder 54%.

1.1.2 Water use in agriculture

Water consumption in Kuwait is high. Some 54% of water is used for agriculture, 44% for municipal purposes, and 2% for industrial purposes (Figure 1.2). For the water withdrawn for agriculture purposes, 80% was used for productive agriculture, 9% for landscape greening, and 11% for garden watering (Frenken, 2009).

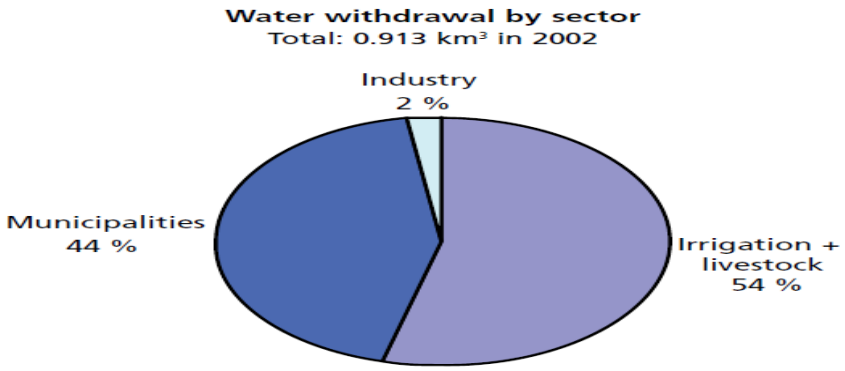


Figure 1.2 Water withdrawal by sector. Kuwait (Frenken, 2009)

1.2 Kuwait's energy scenario

Kuwait's major energy source is from fossil fuel (oil & gas). This finite natural resource is vulnerable and diminishing. Kuwait has the highest annual energy consumption per head of population in equivalent barrels of oil in the Arab world (Croome, 1991). In particular, Kuwait's per capita electricity consumption is amongst the highest in the world according to Encyclopedia of Earth (Cleveland, 2007), at about 14,000 KWH. The extreme weather conditions in Kuwait are the main reason behind the high electricity demand for air conditioning which reaches more than 9,000 mega-watts (MW) in July and August. In fact, according to government sources, an increase of 1^{o}C in ambient temperature causes an increase of 150 MW of electricity demand in the summer.

In Kuwait, the government subsidizes 85 % of the cost of electricity. In addition, the customer pays a fixed figure cost that is 2 fils/kWh (0.006 $/kWh). This has led to an escalation in the demand for electrical energy (Al-Ragom, 2004). As recorded by the Ministry of Electricity and Water, electricity peak demand in Kuwait has been increasing at an alarming rate since the fifties; 32% in the 50's, 26% in the 60's, 15% in the 70's, 8% in the 80's and 90's (MEW, 1999). These rates are considered much higher than the average increase in industrial nations, which have an energy-use rate that does not exceed more than 2%-3%. In Kuwait, the energy consumption increased from 27.0 million MWh in 1999 to 33.1 million MWh in 2003 (MEW, 2003); Figure 1.3)

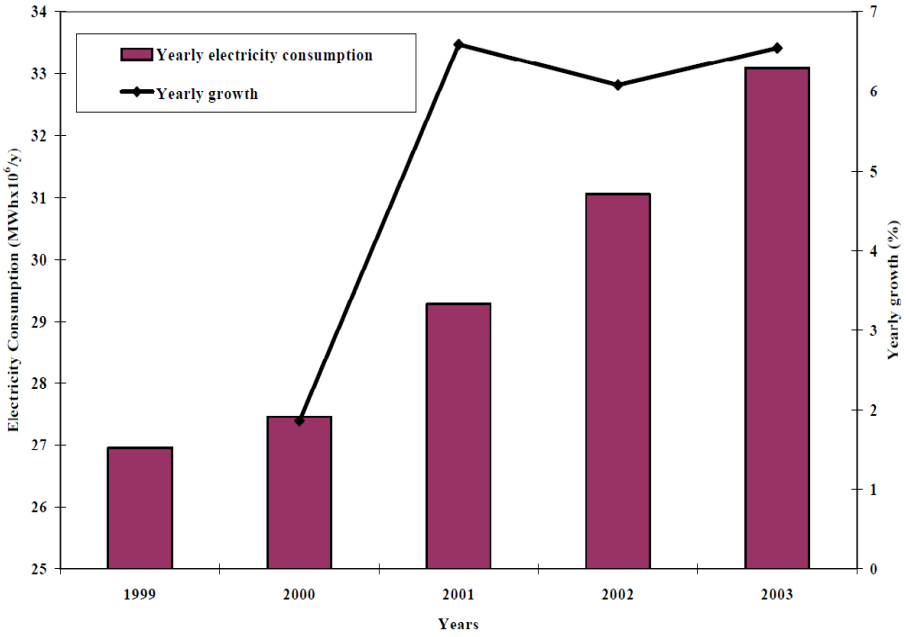


Figure 1.3 Growth of annual electrical consumption in Kuwait from 1999 to 2003 (Hajiah, 2006)

In the summer of 2006 in Kuwait, frequent power cuts were experienced due to equipment failure, giving the country a wake up call in addressing the problem and for the government to adopt a national energy efficient operation campaign. Kuwait is an energy intensive country among other Middle-Eastern countries. If it wants a place in the global economy, Kuwait must improve energy conservation and efficiency that will lead to less green-house gas (GHG) emissions, leading to a better environment. The country also needs to invest in sustainable renewable energy sources.

1.2.1 Future energy plans in Kuwait

Despite holding substantial oil reserves, Kuwait is stepping up its efforts to develop alternative sources of energy. The Shagaya Renewable Energy Park initiative was adopted by Kuwait Institute for Scientific Research, diversifying Kuwait's energy supply by exploring the viability of proven and emerging solar in photovoltaic panels (10 kilowatts) and wind energy (wind turbines, 6 kilowatts) technologies that are capable of overcoming the challenges of Kuwait's harsh climate. The target is to supply 15% of the country's electricity demand by the year 2030 (El-Katiri and Husain, 2014).

1.3 Bioenergy

The biofuel industry is driven by government policies aimed at mitigating climate change, energy security and as a strategy to support rural development. Bioenergy is renewable, non-fossil energy obtained from biomass combustion. Liquid biofuels are either bioethanol or biodiesel. Liquid biofuels can replace petrol and diesel for use in transportation, electricity, cooking and lighting. Biofuels can be defined as first, second and third generation biofuels according to their technological development (Rosegrant, 2008). First Generation Biofuels are derived from food crops such as maize, sugarcane and sugar beet, for the extraction of sugar to produce bioethanol. First generation bioenergy crops (FGEC) compete with food for fertile land.

Second Generation Bioenergy Crops (SGEC) provide fuel from cellulose and non-oxygenated pure hydrocarbon fuels like biomass to liquid fuel (Oliver *et al.*, 2009). SGEC are expected to be more efficient than FGEC, have more energy content (GJ/HA/Yr) and have the potential in reducing cost in the long term (Petersen, 2008). However, there are technical issues in fuel production and growing SGEC which depends on the type of feedstock and when and where they are produced. The net of GHG from cellulosic ethanol is less than ethanol from grain producing FGEC (Carpita and McCann, 2008; Carroll and Somerville, 2009).Third Generation Bioenergy Crops include boreal plants, crassulacean acid metabolism (CAM) plants, and micro algae (Patil *et al.*, 2008). CAM plants are potential sources of feedstock for direct cellulose fermentation (Carere *et al.*, 2008; Borland *et al.*, 2009).

Bioethanol is the most used biofuel in the transportation sector. In fact, transportation is responsible for 30% of global energy usage, and accounts for 21% of total GHG emissions (Watson *et al.*, 1996). There is an increasing demand for bioethanol which will grow by more than a third during 2005 to 2030, most of it coming from the transport sector.

Biofuels have shown a reduction of GHG emissions when compared with fossil fuel. This information is obtained by conducting Life Cycle Analysis (LCA) to calculate CO_2 emissions and uptake at each step of ethanol production and use processes. These steps include; growing of feedstock crop, land use, transporting the crop to production plant, producing ethanol, distribution of ethanol and burning ethanol in vehicles.

When comparing biofuels with gasoline, corn based ethanol reduces GHG emissions by 19% to 52%, depending on the source of energy used during ethanol production. Cellulosic ethanol shows an even greater benefit by reducing GHG emissions up to 86% (Figure 1.4).

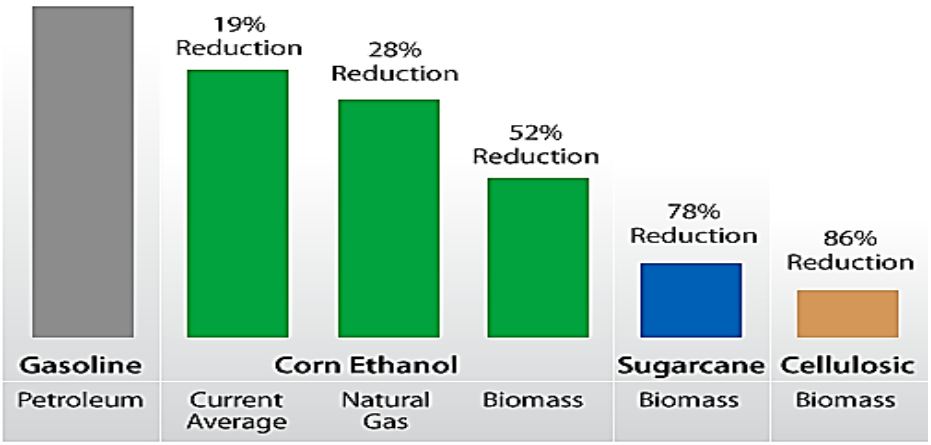


Figure 1.4 GHG emission of transportation fuels (Wang *et al.*, 2007)

1.3.2 Bioenergy feedstocks for Kuwait: the case for *Agave*

The hot, water limited conditions that are found in Kuwait will require that crops grown as potential bioenergy feedstocks in this country have great heat and drought durability in order to ensure a sustainable biomass production system. Succulent species of *Agave* (Agavaceae), which show high water-use efficiency and drought durability represent potential bioenergy feedstocks for semi-arid, abandoned, or degraded agricultural lands and could also help with soil

stabilisation and the reclamation of drylands (Cushman *et al.*, 2008; Borland *et al.*, 2009).Unlike some other drought tolerant biofuel stocks such as maize and sugarcane, *Agave* is a non-food crop and thus, could be grown as a dedicated bioenergy feedstock. In addition to their drought and heat tolerance, *Agave* leaves have high cellulose and sugar contents, and the plants are capable of high biomass yields (Garcia‐Moya *et al.*, 2011)

To date, most research on *Agave* has revolved around A. *tequilana* due to its economic importance in the tequila production industry. The swollen leaf bases (piña's) of *A. tequilana* contain high levels of fructans, fructose polymers which are stored in the leaf vacuole (Davis *et al.*, 2011b).There are other species of *Agave* that display yields greater than *A. tequilana*, such as A. *mapisaga* and A. *salmiana (Davis et al., 2011b).* Also *A. fourcroydes Lem* has a high fructan content and ethanol can be produced from both the leaves and pina's making it a promising plant for biofuel feedstock (MartÍNez‐Torres *et al.*, 2011).

When considering the economic viability of *Agave* as a dedicated bioenergy feedstock, production costs of *Agave* per year in Mexico were lower than those associated with sugarcane production (Sanchez, 2009). In general, *Agave* produces more ethanol per hectare than sugar cane, even with low biomass production due to the high fructan content of the leaves. Thus, *Agave* shows economic and environmental advantages over other widely adopted bioethanol producing crops (Table 1.1). *Agave* is sustainable because it is an environmentally friendly crop in many ways; it has high water use efficiency, it is a non-food crop and doesn't compete with food crops over fertile land and it restrains soil erosion and desertification by carbon sequestration. Furthermore, *Agaves* are considered as low-input perennial crops, similar to *Miscanthus* and switchgrass, that exhibit lower GHG emissions and nitrogen leaching during production than maize (Davis *et al.*, 2015).

Table 1.1 Impact comparison of sugarcane, maize and *Agave mezcalero* in terms of cost biomass production, and ethanol potential (Sanchez, 2009)

Crop	Sugarcane(Mexico)	Maize (USA)	Agave mescalero, (Mexico)
Years to harvest	1	1	6
Yield ton/ha	73.18	12	81.25
Ethanol (Litre)/ha	4	3.785	9.462
Labor	High	High	Low
Water Use	Very high	High	Low
Environmental impact	High	Very high	Low
Need as Food	High	Very high	Low
Sugar Content (%)	8-12	5-10	23-30
Soil/Fertilizer needs	High	Very high	Low
Reduction of GHG emissions from transportation (%)	78	52	86

A key factor underpinning the potential of *Agave* as a sustainable bioenergy feedstock is the fact that the species uses the specialised photosynthetic pathway of crassulacean acid metabolism (CAM) for fixation of carbon. The CAM pathway engenders *Agave* with physiological characteristics that allow these species to operate at near maximum productivity with relatively low water requirements (Borland *et al.*, 2009; Borland *et al.*, 2011). In general, CAM crops such as *Agave* only require 20% of water for cultivation, when compared to calculated values of crop water demand with the most water efficient crops with C3 and C4 photosynthesis (Borland *et al.*, 2009). Table 1.2, indicates the crop water demand for the different photosynthetic pathways, biomass productivity and water use efficiency. The precipitation input from a 100mm rain event equals to 100 Mg H_2O ha^{-1}.

Water use efficiency (WUE) is defined as the ratio of moles of CO_2 fixed and assimilated to moles of water lost by transpiration (Nobel, 2010). CAM plants have high water use efficiencies since they open stomata at night when the temperatures are lower to take up CO_2 and subsequently close them during the day (Garcia‐Moya *et al.*, 2011). High WUE is one of the greatest physiological benefits of CAM photosynthesis (Osmond, 1978; Nobel, 2003) and the evolution and success of CAM plants rely on the defining WUE trait (Gil, 1986; Lüttge, 2006).

Table 1.2 Comparison of the different photosynthetic pathways with different agronomic traits (Borland et.al 2009)

Agronomic Traits	Photosynthetic Pathways		
	CAM	C_3	C_4
Above ground water productivity[Mg (tones)ha^{-1} year^{-1}]	43	35	49
Water use efficiency (mmol CO_2 per mol H_2O)	4-10	0.5-1.5	1-2
Crop water demand (Mg H_2O ha^{-1} year^{-1}	2580-6450	14000-42000	14000-28000

After the liberation of Kuwait in 1991, a plant palette was conducted to evaluate plants which survived the forced neglect for 12-18 months, especially deficiency of irrigation water. Approximately 70 species were included in the initial database. Among these were *Agave americana L* and *Agave americana v. marginata Aurea L*, both of which showed a medium to high tolerance to salinity (640-3200 mg/l), high drought tolerance and required low irrigation (Suleiman and Abdal, 2002).

In conclusion, it would seem that *Agave* could represent a potential bioenergy feedstock for Kuwait. A key aim of this thesis was to compare the potential of a number of different *Agave* species as potential bioenergy feedstocks. Key attributes examined were capacity for CAM, water-use efficiency and sugar accumulation. The following sections provide background on the taxonomy, diversity and productivity of *Agave* before going on to consider in detail, the physiological and biochemical components of CAM and carbohydrate metabolism/sugar accumulation.

1.4 The *Agave* genus

Agaves are keystone species, of arid and semi-arid regions, with Mexico being the geographic centre of origin. Natural populations spread from the south-western United States through Central America, Northern South America and the Caribbean (Garcia‐Moya *et al.*, 2011). The genus *Agave* is the largest in the family *Agavaceae* (García Mendoza, 2002).

1.4.1 Taxonomy, morphology, leaf anatomy and distribution

Agave plants are perennial, belonging to the *Asparagales* order within the monocotyledon family *Agavaceae* with more than 200 species and 47 intraspecific categories (García Mendoza, 2002; Nava-Cruz *et al.*, 2014). Approximately 75% of *Agave* species are found in Mexico which has at least 135 endemic species (Narváez-Zapata and Sánchez-Teyer, 2010). Evidence from molecular clock studies with two different genes evolving at different rates, indicated that the *Agave* genus had a peak in speciation rates that coincided with increasingly dry conditions in central Mexico. The same study indicated that the genus *Agave* emerged 8-10 million years ago (García Mendoza, 2002; Good-Avila *et al.*, 2006).

All *Agave* species are xerophytes but range in size from a few cm to 4 m in height (Valenzuela-Zapata, 1985; Gentry, 2004) Figure 1.5 C). *Agaves* consist of a basal rosette, evergreen succulent leaves which are usually lanceolate in shape with a terminal spine. Some species have leaves with spiny margins. The leaves have a waxy epidermis, sunken stomata which occur on both surfaces of the leaves (amphi-stomatous), and large storage vacuoles in the mesophyll (Blunden *et al.*, 1973). The plants have retractile roots that shrink in response to low soil water potential (Alejandra *et al.*, 2013) which isolate the plant hydraulically from dry air and dry soil, aiding in the maintenance of high water content through long periods of drought (Davis *et al.*, 2011a). The stem is thick and fibrous with a flower emerging as the stem grows. When the growth cycle of the plant nears its end, the flower appears and life span is from 8 to 20 years (Martínez Salvador *et al.*, 2005) Figure 1.5 B). The plants are propagated by seeds with the assistance of pollinators such as insects and nectarvorous bats, (Figure 1.5 D). (Gómez-Pompa, 1963) stated that sexual reproduction is limited or absent, and seeds on average have a 33% germination success rate.

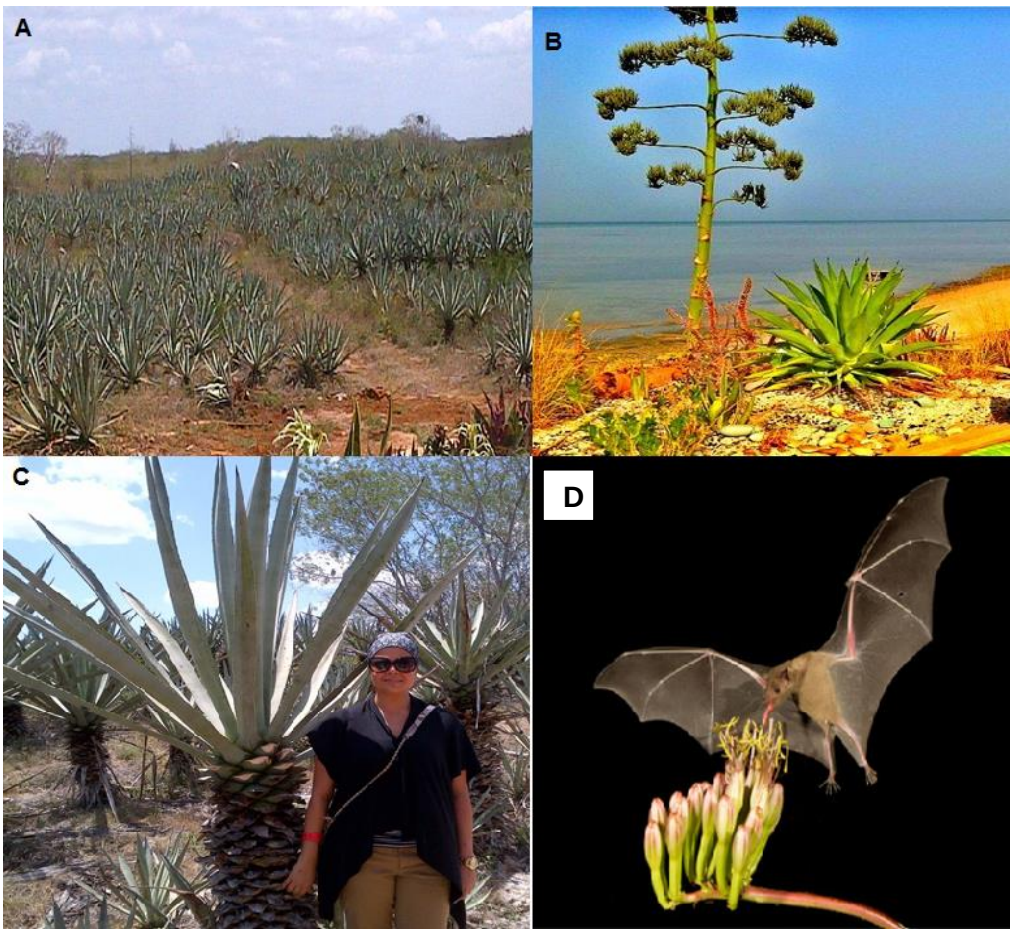


Figure 1.5 Photograph taken of *Agave sisalana* in Merida ,Mexico 2012 (A). Flowering of *Agave americana*. Photograph taken in Nuwaiseeb, Kuwait 2013 (B). In photograph (C) standing beside *Agave angustifolia* in Merida, Mexico 2012. Agave pollinator, the lesser long-nosed bat (*Leptonycteris yerbabuenae*), feeding on *Agave* flower, Amado, Arizona (D). This bat is listed as vulnerable. Photograph taken by Roberta Olenick/Corbis.

Asexual cultivation of *Agave* is common with vegetative stems derived from rhizomes emitted from after the first year of plantation, as illustrated in Figure 1.6. The physiological, morphological and metabolic characteristics of *Agave*, allow them to survive under extreme conditions, and species can be found in valleys, plains, hills and high altitude mountains, some growing in specific areas and others found widely distributed (Nava-Cruz *et al.*, 2014).

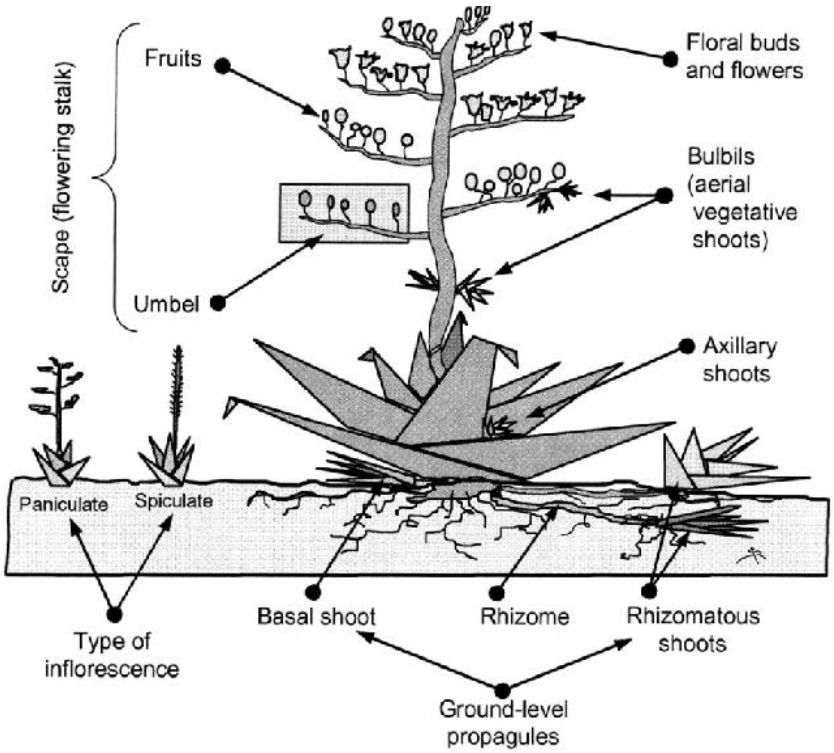


Figure 1.6 Simplified morphology of a rosette of a paniculate *Agave*. (Arizaga and Ezcurra, 2002)

Leaf succulence is a wide spread feature of *Agaves*. Succulence is required for the operation of CAM with leaves possessing large cells containing a central vacuole (Gibson, 1982; Smith *et al.*, 1996; Winter and Smith, 1996) for storage of nocturnal organic acids and water (Borland *et al.*, 1998). Heterogenous chlorenchyma is arranged in a thin layer surrounding photosynthetic cells above a large volume of water storage parenchyma (WSP) (Borland *et al.*, 2000). Large cell size reduces internal air space as a result of tightly packed cells. Succulence would serve to buffer long term changes in water availability, maximizing nocturnal CO_2 uptake and extending the duration of atmospheric CO_2 acquisition duration, particularly under conditions of drought (Pimienta-Barrios et al., 2001).

Agave plants create a microhabitat hosting bacteria, fungi and invertebrates. Originally discovered on *Agave* leaves, the bacterium *Zymomonas mobilis* has the potential as a fermentative organism with high ethanol tolerance (Davis *et al.*, 2011a). A number of parasitic organisms benefiting from *Agave* are the weevil *Schyphohorus acupunctatus* and the fungus *Fusarium spp w*hich causes severe necrosis in xylem tissue (González

et al., 2007). The rhinoceros beetle, *Strategus spp* can kill *Agave* within 24 h by eating the root system (González *et al.*, 2007). Increasing genetic diversity in *Agave* crops will aid in pest resistance or selecting new resistant clones (Zapata and Nabhan, 2003).

1.4.2 Traditional uses and products of *Agave*

Historically in the Americas, *Agave* species have served as a source of food, fibre, shelter, beverages and artisanal speciality products (Colunga-García Marín *et al.*, 2007; Escamilla-Treviño, 2012). The most consumed national alcohol beverage in Mexico is tequila which is distilled and fermented from sugars (fructans) of *A.tequilana Weber* var. azul (López-Alvarez *et al.*, 2012). Tequlia, can only be produced in certain areas of Mexico, for it has protected designation of origin. *Agave* plants are harvested for beverage production when they are between 8-10 years old. Farmers remove the inflorescence in order for sugars to concentrate in the stem and avoid sugar consumption by scavengers such as koyotes. Other species of *Agave* such *as A. angustifolia*, *A. esperrimia, A. weberii, A. potatorum*, *A. salmiana*, are used for production of aquamiel (honey water), nectar or syrup, sweeteners and mescal (Nobel, 2010; Nunez *et al.*, 2011; Escamilla-Treviño, 2012). *Agave fourcryodes* and *A. lechuguilla* are grown for fibres used in cordage and textiles and also for sugars for alcoholic beverages, in countries such as the Philippines, Columbia, Cuba, Nicaragua (MartÍNez‐Torres *et al.*, 2011; Nunez *et al.*, 2011; Valenzuela, 2011). Sisal fibres are derived from *A. sisalana*, and grown in Brazil, Kenya and Tanzania (FAO, 2012); see Figure 1.9). By-products such as biomass from harvested leaves, waste fibre and bagasse from juice extraction can be utilised as compost, animal feed and combustible fuel (Iñiguez-Covarrubias *et al.*, 2001; Chávez-Guerrero and Hinojosa, 2010; Chávez-Guerrero, 2013). *Agave* uses are shown in Figure 1.7

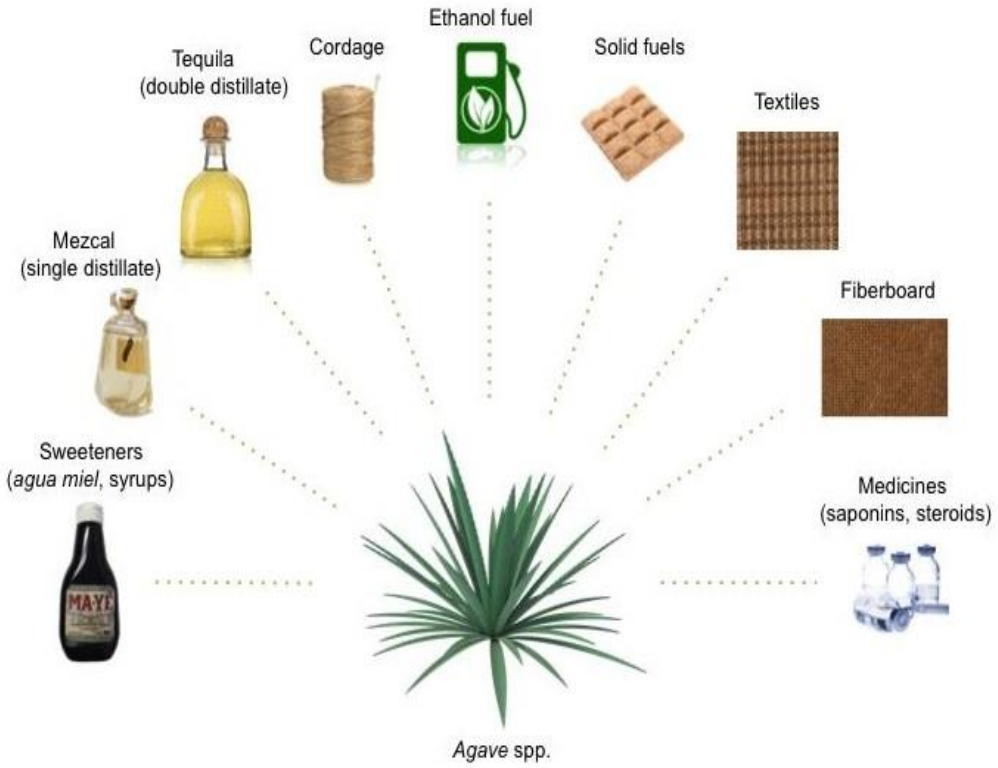


Figure 1.7 Multiple uses & productions derived from *Agave spp*. ranging from beverages, fibres to biofuel (Taken from Cushman et al, 2015)

The increased awareness of recycling fibres has given *Agave* a purpose for this goal (Elenga *et al.*, 2009). *Agave* fibres are biodegradable and recyclable, and have a low density and cost. Thus *Agave* fibres have many advantages over synthetic fibres (Flores-Sahagun *et al.*, 2013).

Figure 1.8 Process to obtain natural fibres from sisal. (A &B) Leaves of *A. sisalana* collected from the field. C) Decortication process. D) Juice extraction and bagasse used as fertilizer and animal feed. E) Drying of sisal fibres. F) Packing of natural sisal fibre. G) End product after compression of sisal. All photographs were taken in Sotuta De Peon Hacienda, Mexico, 2012.

1.4.3 *Agave* as a source of prebiotics and bioactive compounds

Agave species have been used to cure many bacterial diseases and oxidative stress (Ahumada-Santos *et al.*, 2013). Additionally, antifungal (Verástegui *et al.*, 2008), anti-inflammatory (da Silva *et al.*, 2002), antiseptic (Orestes Guerra *et al.*, 2008) and anti-hypertensive activities (Duncan *et al.*,

1999) have been observed. Some organic extracts of *Agave* demonstrated antibacterial activity against *Streptococcus group A-4, Salmonella enterica typhi, Shigella dysenteriae, Escherichia coli 25922, Pseudomonas aeruginosa 27853, Enterococcus faecalis 29212, Staphylococcus aureus 3, Escherichia coli A011,* and *Staphylococcus aureus 29213*; with action from *A.tequilana* (Ahumada-Santos *et al.*, 2013). The *Agavaceae* family is also recognised as an important source of sapogenins with steroidal nature and primarily saponins, which have applications as antifungal, antibacterial, anti-cancer and anti-hemolytic activity (Güçlü-Üstündağ and Mazza, 2007)

1.4.4 *Agave b*iomass characteristics and composition

Water soluble carbohydrates (WSC) are found in high concentrations in *Agave* species and are concentrated in the *piña* (in Spanish due to the resemblance of the harvested stems to pineapples). The *piña* are the swollen stem bases which are rich in non-structural carbohydrates (Figure 1.9).

Tissue composition differs among *Agave* species and varieties and changes over the lifetime of the plants (Arrizon *et al.*, 2010). The most abundant sugar found in *Agave* plant tissue is fructose and much of this fructose is found in fructo-oligosaccharides (fructans) which are stored in the vacuole. Total sugar content in the *piña* ranges from 12-28% (fresh weight) (Yan et al., 2011). Fructan concentrations in the *piña* range from 36 to 73% (dry weight) of tissue at maturity depending on species (Davis *et al.*, 2011a). Fructans are oligomers composed mainly of fructose units attached to a sucrose molecule, which is easily degradable by thermal or enzymatic treatments (Narváez-Zapata and Sánchez-Teyer, 2010).

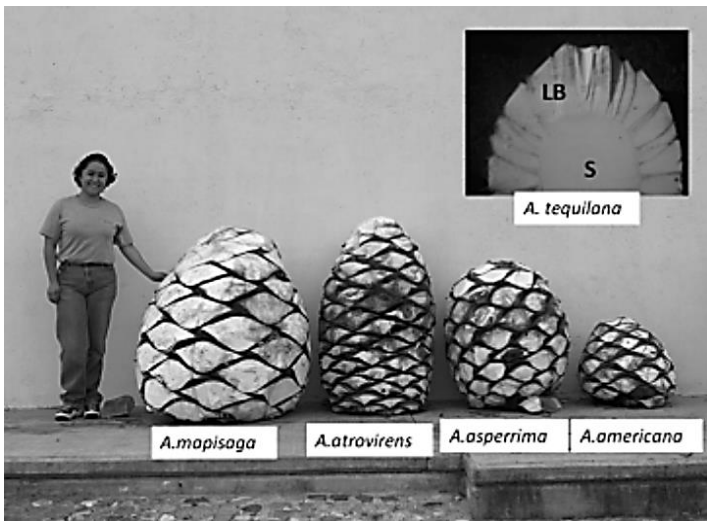


Figure 1.9 Examples of harvested *piñas* from different *Agave* species. From left to right are *A. mapisaga* (diameter=310 cm, weight=471.85 kg), *A. atrovirens* (diameter=215 cm, weight= 280.4 kg), *A. asperrima* (diameter=225 cm, weight= 222.5 kg), *A. americana* (diameter=172 cm, weight=76.2 kg). Stems taken close to maturity. Guanajuato, Mexico. The inset shows a dissected *Agave tequilana* stem. S= Stem, LB=Leaf Base (Simpson *et al.*, 2011b).

Depending on the linkage type between the fructosyl residues and the position of the glucose residue, different types of fructans may be found (Lewis, 1984). *Agave* fructans are formed from a basic sucrose molecule by β (2-1) and β (2-6) linkages between fructose residues to form 1-ketose by sucrose:sucrose 1 fructosyl transferase (6-SFT). Neoketose is formed by 1-ketose by adding fructan:fructan 6G fructosyltransferase (6G-FFT) and bifurcose from 1-ketose by adding fructose in a β (2-6) linkage by 6-SFT. The enzyme fructan:fructan 1-fructosyltransferase (1-FFT) is necessary in completing the synthesis of long and complex fructan structures (agavins and graminans) (Figure 1.10) (Simpson *et al.*, 2011a).

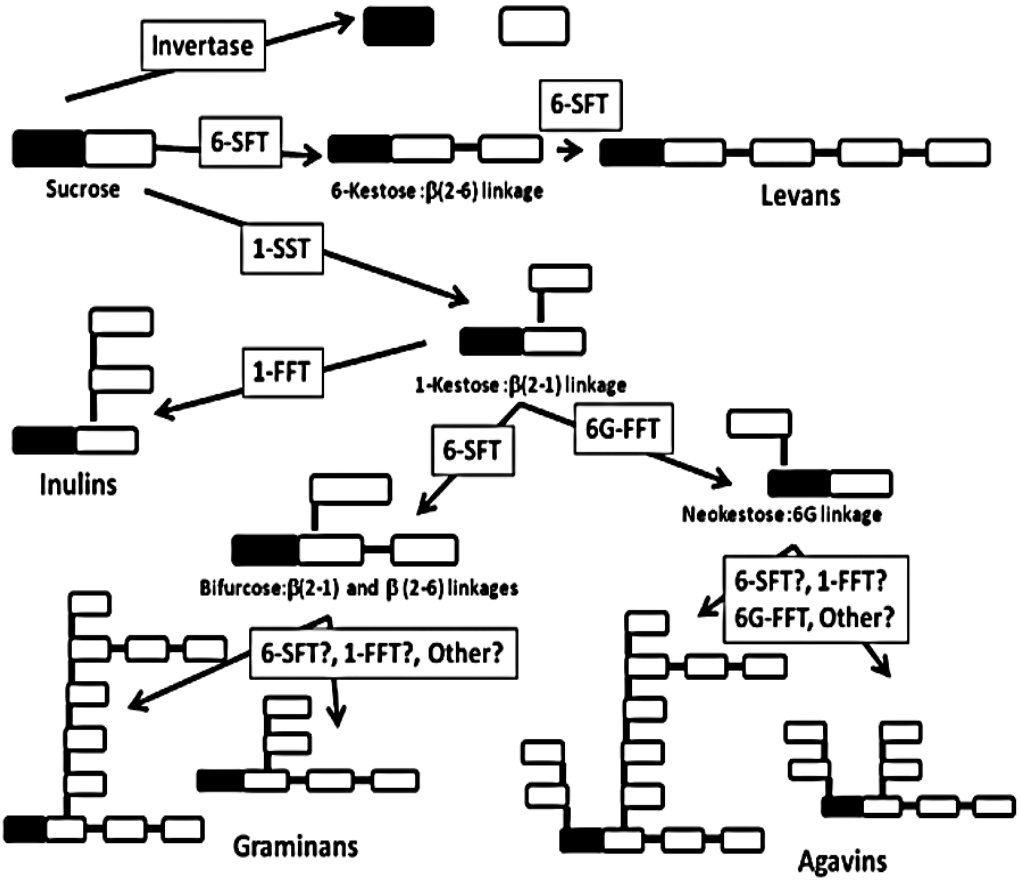


Figure 1.10 Outline of enzymes involved in fructan metabolism in *Agave*. Black boxes represent glucose residues, Open boxes are fructose residues, 1-SST, sucrose:sucrose 1-fructosyltransferase; 6-SFT, sucrose:fructan 6-fructosyl transferase; 6G-FFT, fructan:fructan 6G-fructosyltransferase;1-FFT, fructan:fructan 1-fructosyltransferase (Simpson *et al.*, 2011a).

In *Agave* there is more than one fructan structure. *Agave* fructans have a unique feature, in which the molecules of fructose have β (2-1) linkages and 3 to 29 degrees of polymerization (DP) with β (2-6) linkages which classify them as mixed fructans and neoseries fructans (López and Mancilla-Margalli, 2007). In *A.tequilana*, fructans have received the name of agavins (Muñoz-Gutiérrez *et al.*, 2009) (Figure 1.11), which have been in use for tequila production, dietary products and systems of drug delivery (Arrizon *et al.*, 2010). The production of fructans is influenced by several factors such as growth region, nutrients in the soil, climatic changes, seasonal time and water level and also differ depending on the *Agave* species and their age (Muñoz-Gutiérrez *et al.*, 2009)

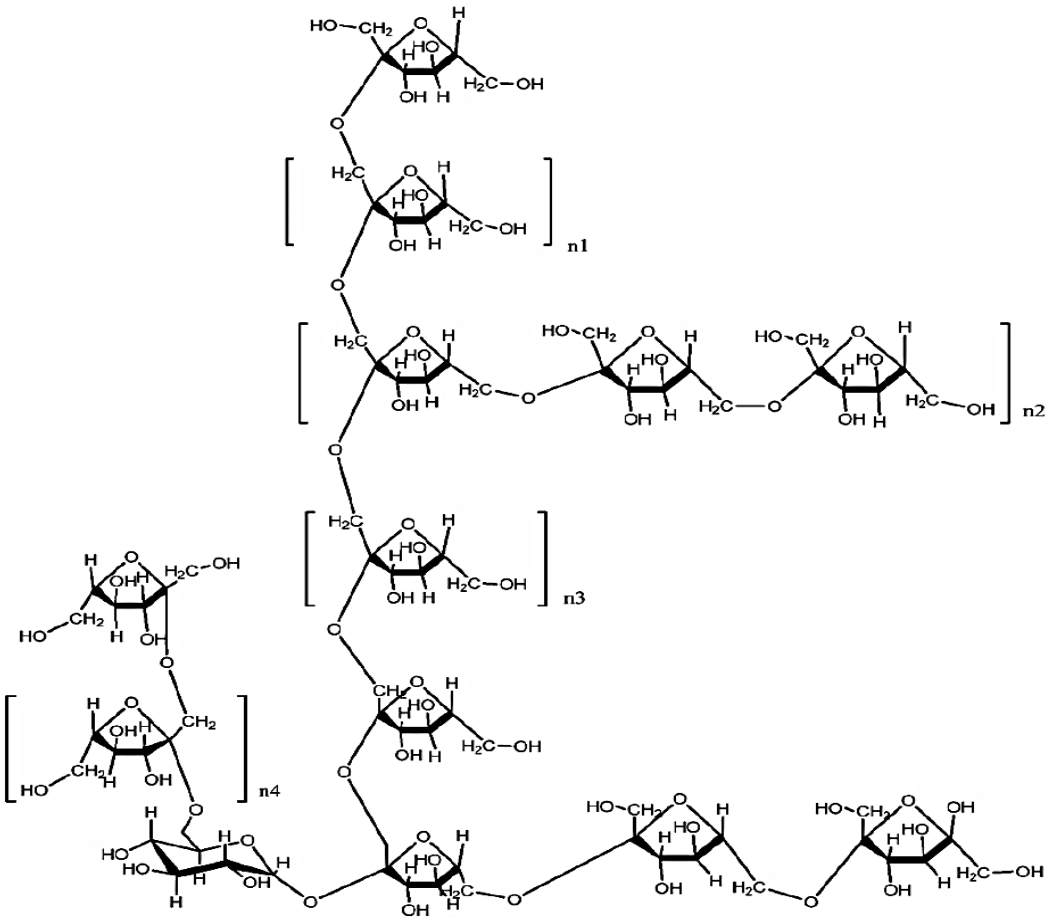


Figure 1.11 Structure of the polysaccharide agavin found in *Agave* species (López and Mancilla-Margalli, 2007)

(Mellado-Mojica and López, 2012) proposed that new possible molecular structures of agave fructans occur during the plant life cycle in the field. This suggestion was based on *A.tequilana* fructan content which increased to a maximum in 5 year old plants and remained constant up to the age of 7. The plant starts off with equal amounts of agavins and graminans and then moves toward a higher abundance of agavins with higher DP as plants age, producing isomeric forms that are complex and difficult to identify (Figure 1.12).

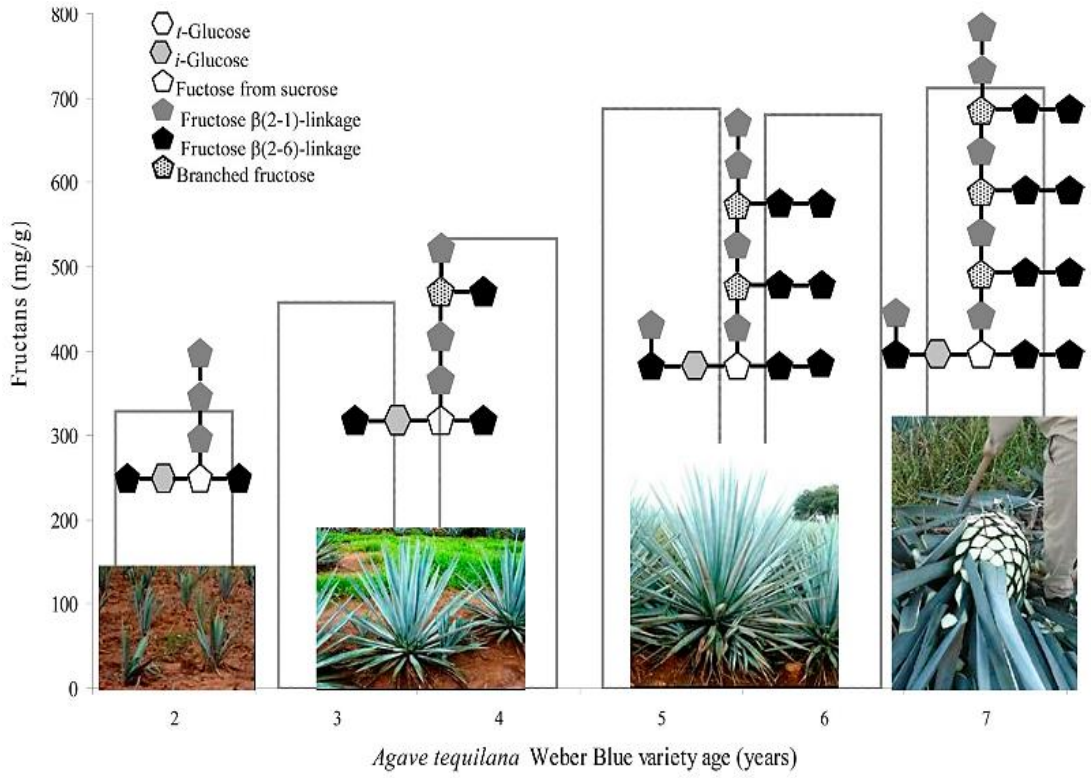


Figure 1.12 Life cycle of *A.tequilana* Weber Blue variety in the field, with fructan content and proposed molecular structures. *A. tequilana* Weber Blue exhibits changes in carbohydrate, fructan content, DP type and molecular structure (Mellado-Mojica and López, 2012)

Agave plants generally have low lignin content (4.9-19.3% dry weight). The low lignin content is beneficial for overcoming recalcitrance to cellulose degradation and improving saccharification for the eventual production of bioethanol (Ragauskas *et al.*, 2006). In addition to low lignin content, some *Agave* species have low crystalline cellulose content and high paracrystalline cellulose content relative to woody biomass feedstocks (Yan *et al.*, 2011; Li *et al.*, 2012b). Moreover, the high water content in *Agave piña* and leaves, ranging from 60-70-% and 78-89% respectively, could reduce water inputs needed for downstream lignocellulosic processing (Yan *et al.*, 2011). Table 1.4 exhibits structural carbohydrate composition from various *Agave* species.

Table 1.3 Comparison of biomass composition of different *Agave* feedstocks

***Agave* ssp. (fraction)**	**Structural Component (dwt %)**					
	Solubles (Extractives)	Cellulose	Hemicellulose	Lignin	Ash	Citation
A. americana (Bagasse)	14.5	n/a	n/a	8.2	7.4	(Li *et al.*, 2012a)
A. fourcroydes (Leaf fibre)	3.6	77.6	5-7	13.1	n/a	(Vieira *et al.*, 2002)
A. lechugulla (Leaf fibre)	2-4	79.8	3-6	15.3	n/a	(Vieira *et al.*, 2002)
A. salmiana (Bagasse)	n/a	47.3	12.8	10.1	n/a	(Garcia-Reyes and Rangel-Mendez, 2009)
A. salmiana (Bagasse)	17.9	n/a	n/a	9.8	6.1	(Li *et al.*, 2012a)
A. sisalana (Leaf fibre)	n/a	77.3-84.4	6.9-10.3	7.4-11.4	n/a	(Vieira *et al.*, 2002) (Martin *et al.*, 2009)
A. tequilana (Bagasse)	14	64.8	5.1	15.9	1.0	(Iñiguez-Covarrubias *et al.*, 2001b)
A. tequilana (Bagasse)	n/a	68.4	15.7	4.9	n/a	(Mylsamy and Rajendran, 2010)
A. tequilana (Bagasse)	17.4	n/a	n/a	11.9	6.4	(Li *et al.*, 2012a)
A. tequilana (Bagasse)	n/a	n/a	n/a	19.3	4.4	(Perez-Pimienta *et al.*, 2013)
A. tequilana (Bagasse)	29.7	26.6	23.4	13.1	6.1	(Yang *et al.*, 2015a)

1.4.5 *Agave* biomass production

The best productivities measured for *Agave* species are 38 and 42 Mg ha^{-1} $year^{-1}$ for *Agave mapisaga* and *A. salmiana*, respectively growing in Mexico (Nobel *et al.*, 1992; Davis *et al.*, 2011a). These yields far exceed corn, soy-bean, sorghum and wheat productivities under intensive management. Most yields have been assessed for individual experimental plants rather than production fields where yields are likely to be lower. To provide an analytical framework for evaluating environmental and edaphic factors on net CO_2 uptake and plant productivity, an Environmental Productivity Index (EPI) was developed as a powerful quantitative tool (Nobel *et al.*, 1998; Nobel, 2003). EPI helps to evaluate the agronomic potential of *Agave* by predicting productivity over wide geographical areas with diverse environmental conditions. EPI can be represented as Light index x Temperature index x Water index x Nutrient index x CO_2 index (Nobel, 2010). Individual indices vary from 0.00 which indicate complete inhibition of net CO_2 uptake up to 1.00 which is optimal. Predictions of yield using EPI have been shown to correlate with actual measurements of the rate of unfolding of new leaves from the central spike, as first shown in *A.deserti* and *A. fourcryodes* (Nobel, 1985; Nobel, 2010). Unfolding of leaves is a useful

morphological indicator of biomass productivity and varies with plant age, shading and season. An annual comparison of total number of leaves unfolding in 3 year old and 6 year old plants was 19.6 and 24.9 respectively ($p<0.05$). When shading was reduced by 30%, it reduced the number of leaves unfolding for both plant ages by 35% ($p<0.01$). Unfolding rates increase in wet summer season vs. dry winter season (Garcia - Moya *et al.*, 2011).

The predictions of *Agave* growth and productivity are important considerations for optimizing the colocation of solar panels and *Agave in* hybrid bioenergy and renewable energy production systems (Figure 1.13)

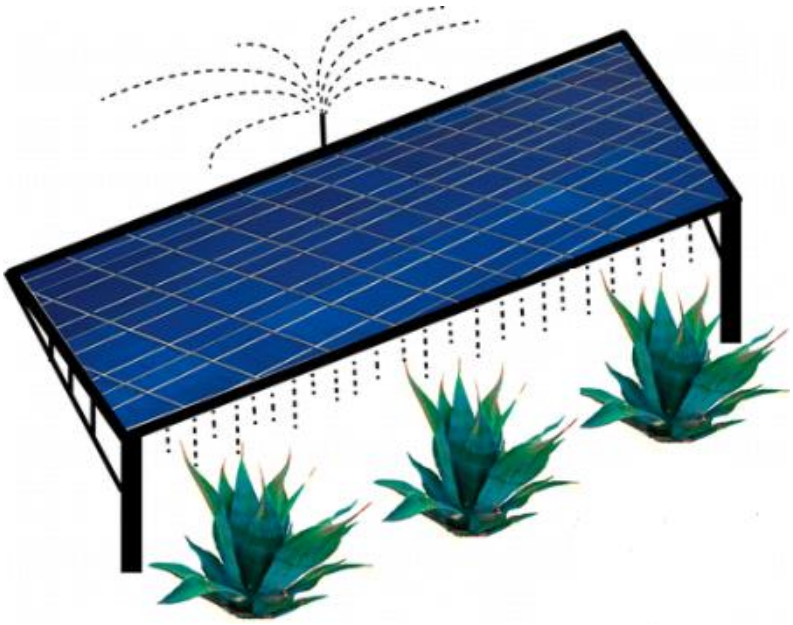


Figure 1.13 Conceptual colocation of PV solar panels with *Agave*, showing water input for cleaning solar panels and dust suppression equals water needed for annual Agave growth (Ravi *et al.*, 2012)

Solar energy installations in deserts are on the rise due to policy changes and advances in technology. This has inspired a comparative study on the water use and GHG emissions associated with solar installations and *Agave*-based biofuel production. A life cycle analysis (LCA) of hypothetical colocation resulted in higher returns per m^3 of water used than either system alone and could generate a higher rate of energy return (Ravi *et al.*, 2012) Figure 1.13). Colocation can be an advantage in water limiting environments providing attractive economic incentives and efficiency of land and water use

Agave is typically propagated asexually from bulbils. Micro-propagation is currently used in the tequila industry(Robert *et al.*, 2006; Ramírez-Malagón *et al.*, 2008). Prior to planting in the field, plantlets are grown in culture and transferred to a greenhouse for 1-2 years. A typical planting field ranges from 2000-4000 plants ha^{-1} for tequila production (Cedeño, 1995). In a regional evaluation of crops in Mexico, composition of carbohydrates extracted from the same species differed according to location subjected to different climates (Mancilla-Margalli and López, 2006). Several species of *Agave* including *A. angustifolia*, *A. potatorum* and *A. cantala* had similar carbohydrate profiles among species. This is an important indication of site selection for optimising biofuel yield.

1.4.6 Effects of global climate change on *Agave* productivity

Challenges that necessitate the search for alternatives to generate energy efficiently are of great importance with ecological sustainability and global climate change (Pimienta-Barrios *et al.*, 2001). There is a need for agricultural biofuel crops that allow effective CO_2 sequestration under the warmer and drier world that climate models predict for the next 60 years whilst producing high sugar contents that are readily convertible to alcohol (Nobel, 2010). *Agave* fits the bill by effective CO_2 sequestration in water deficient environments and producing high sugar contents and combined genetic diversity will enable a better response to global climate change (Garcia‐Moya *et al.*, 2011). Elevated levels of atmospheric CO_2 modify the morphology and anatomy of CAM plants, including *Agave*. The chlorenchyma has been shown to increase in thickness, which might be related to higher CO_2 concentrations deeper within the leaves (Powles *et al.*, 1980), root systems expand and shoot development occurs more rapidly (Nobel, 2010). In *Agave deserti,* cladodes were 11% thicker under a doubled atmospheric CO_2 concentration (Graham and Nobel, 1996; Zhu *et al.*, 1997). *Agave* plants tested showed significant stimulation of biomass accumulation under increasing CO_2 (Table 1.5). Owen & Griffiths (2014) predicted bioethanol yield potential for *Agave* species in Australia, by developing a geospatial model based on the Environmental Productivity Index (EPI) approach. The modelling approach was used to predict crop production on marginal lands under current and future conditions. Simulations for predicted *Agave* productivity under future climate conditions

look promising and could have a beneficial impact on *Agave* production for Kuwait, indicated by the blue colour on simulation (b) in Fig 1.14.

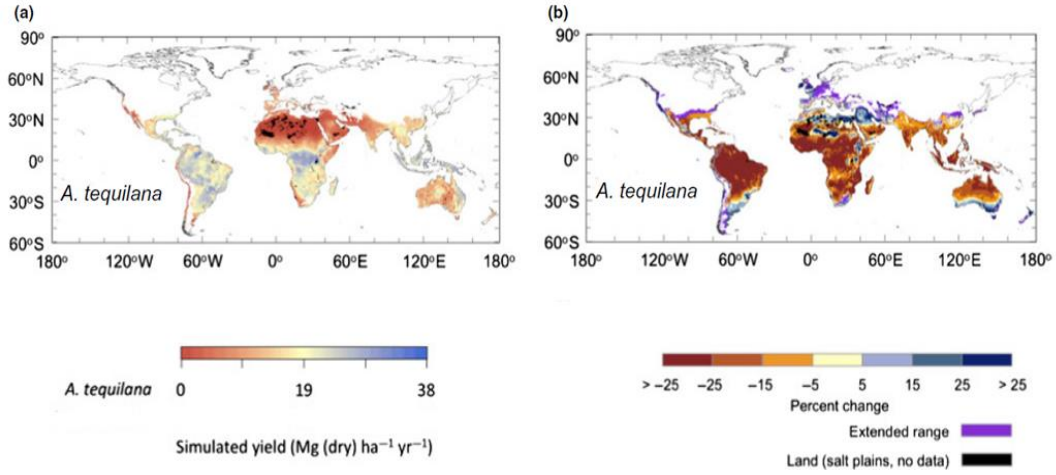


Figure 1.14 Simulations of predicted *Agave tequilana* productivity under current and future climate conditions. (a) Simulations under current climate conditions, geographical distribution of highly productive areas (Environmental Productivity Index (EPI)>0.5) is restricted for *A. tequilana* due to high sensitivity to nocturnal temperature and lower capacity to buffer against low soil water potential capacities. Response of higher saturation point for carbon uptake to photosynthetically active radiation (PAR) for *Agave*= 29 mol m^{-2} d^{-1}, has a negative impact on yields at latitudes >30°S or 30°N. (b) Simulated productivity under future climate conditions in the year 2070. Outside the range of 30°S to 30°N climate change has a beneficial impact on *A. tequilana* productivity. Simulations used environmental inputs averaged over the period 1950-2000. (Yang *et al.*, 2015b)

Table 1. 4 Response of biomass of *Agave* to long term (>1 month) exposure to doubled atmospheric CO_2 concentrations. Adapted from (Ceusters and Borland, 2011). Controls were maintained under ambient atmospheric CO_2 concentrations for the same period.

Species	Biomass (% increase over control)	References
Agave deserti	30-31	(Nobel and Hartsock, 1986; Graham and Nobel, 1996)
Agave salmiana	17	(Nobel, 1996)
Agave vilmoriniana	28	(Idso *et al.*, 1986)

1.5 Physiological ecology of *Agave*

1.5.1 CAM photosynthesis and water use efficiency (WUE)

Agave has the specialised photosynthetic pathway known as Crassulacean acid metabolism (CAM), which was first found in the Crassulaceae family of plants (Keeley and Rundel, 2003). This carbon concentrating mechanism is found in approximately 7% of all vascular plant species (Nobel, 2010), allowing high productivity under constrained water availability (Cushman, 2001). CAM is a photosynthetic pathway where carbon dioxide (CO_2) is fixed as a four carbon acid malate during the night, when the stomata are open. During the day, the malic acid is broken down to release CO_2 which is re-fixed by Rubisco behind closed stomata. The opening of stomata at night, rather than during the day reduces evapotranspiration, because it is cooler and more humid at night. Thus, CAM renders the plant more water efficient which in turn enables CAM plants to adapt to arid conditions (Nobel, 1991).

The temporal separation of carboxylases is what distinguishes CAM from C_3 and C_4 photosynthetic pathways. There are four distinct phases of gas exchange in CAM plants based on stomatal behaviour, modes of CO_2 uptake and fixation, and C_4 acid and carbohydrate accumulation over a course of the diurnal cycle (Osmond, 1978; Winter, 1985; Lüttge, 1987; Griffiths, 1988) as shown in Figure 1.15

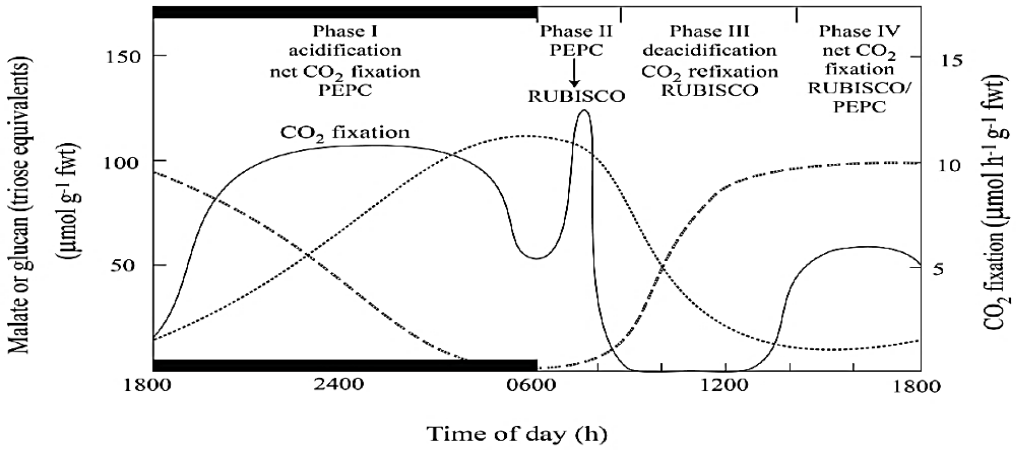


Figure 1.15 Generalised schematic representation of day/night CO_2 fixation (solid line), malic acid (dotted line) and carbohydrate (dashed line) content observed in well watered CAM plants. The dark period is indicated by the black bar (Osmond, 1978; Leegood and Osmond, 1990; Smith and Bryce, 1992).

Phase I During the night, the stomata are open, allowing CO_2 to enter the mesophyll cells, where it is ultimately fixed by the enzyme phospho*enol*pyruvate carboxylase (PEPC). The eventual carboxylation product is the 4-C organic acid malate which accumulates overnight in vacuoles of the cell as malic acid. The PEP required for malate synthesis is provided by the nocturnal breakdown of carbohydrate. Rates of nocturnal CO_2 assimilation are governed by carbohydrate storage reserves (Cushman *et al.*, 2008) as well as vacuolar storage capacity, rather than by stomatal conductance (Winter, 1985; Winter *et al.*, 1985). Phase I results in reduced transpiration and helps to improve water economy which is the fundamental of CAM adaptation (Griffiths, 1989).

Phase II This is a transitional phase between PEPC-mediated and ribulose-1,5-biphosphate carboxylase/oxygenase (RUBISCO)-mediated CO_2 fixation (Silvera *et al.*, 2010). The stomata open during the early hours of the light period. Stomatal conductance declines as internal CO_2 partial pressure gradually increases as a result of the onset of malate breakdown. PEPC is deactivated in the morning by dephosphorylation, which renders the enzyme sensitive to malate inhibition (Winter, 1982; Nimmo *et al.*, 1984)

Phase III The decarboxylation of malic acid occurs over the middle part of the day, producing CO_2 and C_3 carbon backbones for carbohydrate synthesis and C_3 photosynthesis. This is accompanied by stomatal closure. Malate effluxes from the vacuole and is decarboxylated to release CO_2 which enters the chloroplasts and is concentrated around the enzyme Rubisco, thus entering the Calvin Cycle to produce triose-P and ultimately carbohydrate. This CO_2 concentrating mechanism suppresses photorespiration during phase III (Silvera *et al.*, 2010).

Phase IV Is a second transitional phase. Stomata re-open, due to exhaustion of malate and a drop in internal CO_2 concentration. Direct fixation of exogenous CO_2 occurs by the Calvin Cycle via Rubisco for the remainder of the light period (Borland *et al.*, 2009). Phase IV may involve both C_3 and C_4 carboxylation processes if PEPC is re-activated before the dark period commences (Ritz *et al.*, 1986; Griffiths *et al.*, 1990)

The duration of each phase of the CAM cycle varies between species, environmental conditions and the stage of leaf development (Winter *et al.*, 2008).

1.5.2 Physiology of leaf gas exchange in *Agave*

Measurements of photosynthesis and transpiration for *A. americana* were first conducted by Neales *et al.*(1968), Ehrler (1969) and Kirsten (1969). The data showed the nocturnal opening of *Agave* stomata (Neales *et al.*, 1968; Ehrler, 1969; Kristen, 1969), with 75% of daily net CO_2 uptake occurring at night. In this *Agave* species, net CO_2 uptake during phase II (early photoperiod) lasted for less than 1 hour but a significant phase IV was observed (Figure 1.16).

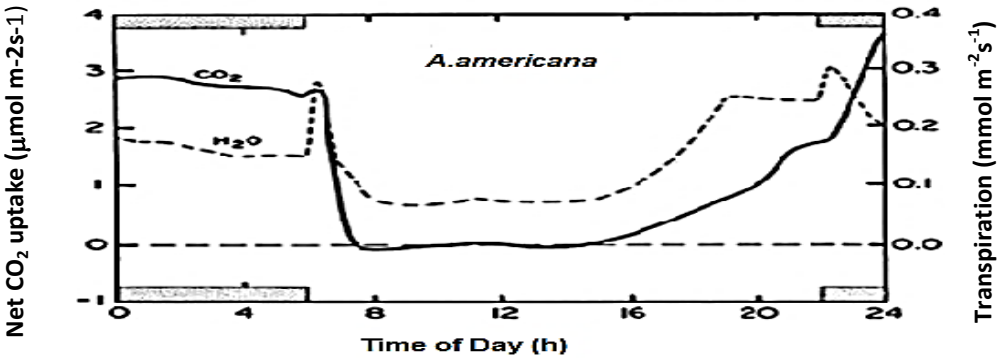


Figure 1.16 Day/night pattern of leaf gas exchange by *A. americana* showing net CO_2 uptake and transpirational water release. The solid bars on the x-axes indicated the periods of darkness(Nobel, 2003)

For many other succulent species of *Agave*, CAM is a ubiquitous trait with generally reduced gas exchange at Phases II and IV (Alejandra *et al.*, 2013). Other *Agave* species with gas exchange patterns comparable to that illustrated in Fig. 1.16 are *A. deserti, A. angustifolia, A. salmiana, A. fourcryodes, A. lurida, A. parryi, A. murpheyi, A. weerii, A. scabra, A. schottii, A. lechuguilla, A. vilmoriniana, A. tequilana, A. shawii and A. utahensis* (Eickmeier and Adams, 1978; Woodhouse *et al.*, 1980; Alejandra *et al.*, 2013).

Hartsock (1976) reported C_3-CAM facultative behaviour for *A. deserti* under well-watered and droughted greenhouse conditions. A complete and

reversible switch from CAM to C_3 diel gas exchange was observed depending on watering regime. Under well-watered conditions, net CO_2 uptake was only observed during the day and no day/night acid fluctuations were observed (Hartsock and Nobel, 1976). However, even constitutive CAM species of *Agave* can show plasticity in the magnitude and duration of CAM phases. In the constitutive CAM *A. tequilana*, photosynthetic plasticity is observed between young and adult plants allowing the modulation of daytime contribution (Phases II and III) and night-time (Phase I) carbon acquisition when exposed to different environmental conditions (Pimienta-Barrios *et al.*, 2001). Both young and adult plants of *A. tequilana* perform some daytime gas exchange (Phase IV) (although the % of day: night-time net CO_2 uptake is generally higher in young plants. Phase IV net CO_2 uptake can be maintained in some *Agave* species during dry spells which is not commonly observed among other CAM plants growing in arid environments (Pimienta-Barrios *et al.*, 2001).

1.6 CAM biochemistry

In considering the biochemical processes of CAM, the day/night metabolic cycle and its underlying biochemistry are best considered within the context of the 4 phases of gas exchange described above (Osmond, 1978). Starting from the end of the photoperiod, the CAM cycle begins at night with Phase I and the metabolic steps are illustrated in Figure 1.17. In the cytosol, oxaloacetate (OAA) is produced by the carboxylation of phosphoenolpyruvate (PEP) with HCO_3^- which is catalysed by the enzyme phosphoenolpyruvate carboxylase (PEPC). HCO_3^- is produced from the action of carbonic anhydrase on CO_2. OAA is quickly converted to malate via the enzyme malate dehydrogenase (MDH) and malate then enters the cell vacuole via malate selective voltage-gated ion channels providing charge balance for tonoplast bound H^+ATPase and or H^+ Pyrophosphatase (H^+-PPiase) (Smith and Bryce, 1992; Bartholomew *et al.*, 1996; Smith *et al.*, 1996; Hafke *et al.*, 2003). The H^+ electrochemical difference established by ATP and PPiase pumps maintains an inside positive potential which drives the influx of $malate^{2-}$ anions across the tonoplast through the vacuolar malate channel (Hafke *et al.*, 2003). Malic acid accumulation and net CO_2 uptake continue for most of the dark period, with concentrations of vacuolar malic acids reaching ~200 mM by dawn (Borland *et*

al., 2009; Escamilla-Treviño, 2012). The activation of PEPC at night occurs via post-translational modification (see Figure 1.17). The phosphorylation of PEPC during the dark is hypothesised to lower internal partial CO_2 pressure inside the leaf, and it is further hypothesised that this action, signals stomatal opening during the dark period thus providing a sustainable supply of CO_2 to carbonic anhydrase and PEPC (Borland *et al.*, 2009).

PEPC is dephosphorylated in the few hours before dawn during phase II, making it ~10 times more sensitive to inhibition by malate. This is a critical step curtailing futile cycling at the start of the photoperiod in CAM plants (Borland *et al.*, 1999). Rubisco activation is mediated via Rubisco activase commencing at the start of the photoperiod. A surge of CO_2 uptake may occur in Phase II where CO_2 is fixed by both PEPC and Rubisco for a brief period.

During the day, malate is exported from the vacuole to the cytosol where it is decarboxylated (Phase III). Malate decarboxylation can occur by several routes and enzymes depending on the CAM species (Dittrich *et al.*, 1973; Dittrich, 1976; Holtum *et al.*, 2005). Decarboxylation can occur by either phosphoenolpyruvate carboxykinase (PEPCK) or cytosolic $NADP^+$- and/or mitochondrial NAD^+-malic enzymes (ME) (Holtum *et al.*, 2005), a feature which is broadly species dependant (Christopher and Holtum, 1996; Christopher and Holtum, 1998). In *Agave*, the activity of PEPCK is reportedly low or not detectable and thus it is believed that malic enzyme(s) are responsible for decarboxylation in the *Agave* genus (Black et al, 1992; (Escamilla-Treviño, 2012). Increasing levels of CO_2 generated by malate decarboxylation in phase III behind close stomata, saturates the carboxylase and supresses oxygenase function of Rubisco, even though internal O_2 levels are also elevated. In well watered CAM plants, stomata may re-open later in the photoperiod (Phase IV) due to exhausted supply of malate and internal CO_2 concentrations drop. Direct fixation of atmospheric CO_2 by Rubisco follows for the remainder of the light period. The magnitude and duration of each phase of the CAM cycle is highly plastic and varies with species, response to the environment and leaf development (Winter *et al.*, 2008).

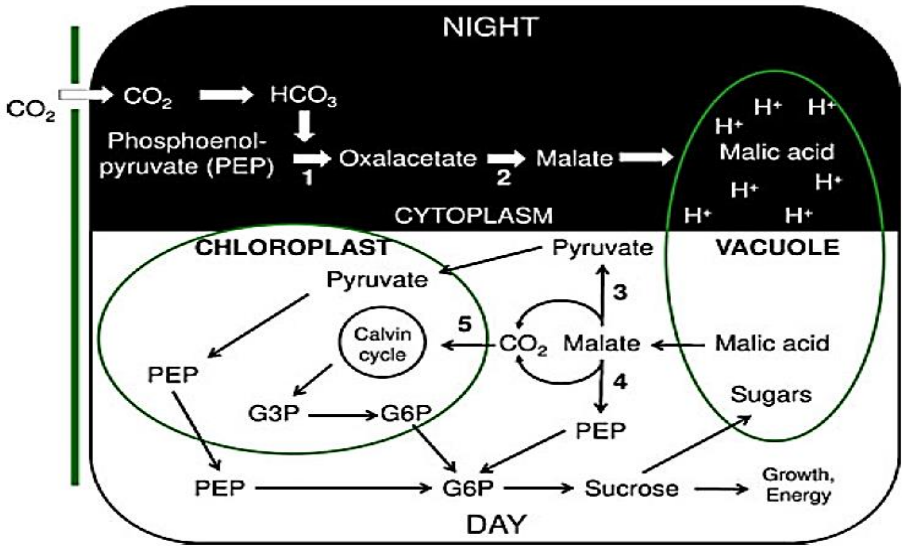


Figure 1.17 The CAM pathway in a mesophyll cell. The green line on the left of the diagram represents leaf epidermis with a gap represents stomatal pore. Black represents the night and white represents during the day. Active enzymes during night are (1) PEPC and (2) malate dehydrogenase. In *Agave*, it is not clear if decarboxylation to pyruvate occurs by the $NADP^+$ malic enzyme and/or NAD^+-malic enzyme (ME) Adopted from (Escamilla-Treviño, 2012)

1.6.1 PEPC regulation in CAM

The CAM form of PEPC needs to be active at night and inactive during the day to avoid competitive carboxylation and futile cycling of organic acids. *In vitro* PEPC activity does not change over the day/night cycle, and instead the enzyme activity is regulated via post-translational modification (Nimmo *et al.*, 1984; Honda *et al.*, 1996). At night, PEPC is activated via phosphorylation by a dedicated PEPC kinase which reduces enzyme sensitivity to inhibition by malate. During the day PEPC is dephosphorylated and inactive and sensitive to malate inhibition (Nimmo *et al.*, 1984; Nimmo *et al.*, 1986). However, studies on different constitutive and facultative CAM species showed that up to 50% of CO_2 uptake over 24 h can occur during Phase II (Borland *et al.*, 1996). Studies in the laboratory and field on *Clusia* genus gave evidence of PEPC activity remaining 4-5 h after dawn as indicated by continued accumulation of organic acids and low values of instantaneous carbon isotope discrimination measured during leaf gas exchange (Borland *et al.*, 1993; Roberts *et al.*, 1997). In contrast, in *Kalanchoe daigremontiana*, PEPC is rapidly phosphorylated within the first hour of the photoperiod (Borland and Griffiths, 1997). The degree of PEPC phosphorylation can modulate carbon gain in response to short term environmental changes which alter the amount and/or partitioning of malate

between vacuole and cytosol (Borland *et al.*, 2000). In leaves of *C.minor* and *K*.*daigremontiana* which were prevented from accumulation of malate overnight in an N_2 atmosphere, subsequent transfer to ambient air at the start of the photoperiod, resulted in an increase PEPC phosphorylation for 2-3 h of the photoperiod, accompanied by an increase in net CO_2 uptake during Phase II, and de-phosphorylation occurred some 3-4 h into the light (Borland and Griffiths, 1997). Dephosphorylation of PEPC is by a type 2A protein phosphatase, showing constant expression throughout the CAM cycle, whereas PEPC kinase transcript and protein abundance fluctuates during the 24 h cycle (Carter *et al.*, 1990; Carter *et al.*, 1991). Thus, PEPC phosphorylation/activation is primarily dependent on the activity of the protein kinase. This kinase is highly specific to PEPC and in CAM plants is a Ca^{2+} independent kinase (*Ppck1*) synthesised *de novo* on a daily basis under circadian control (Carter *et al.*, 1996; Hartwell *et al.*, 1996; Hartwell *et al.*, 1999; Taybi *et al.*, 2000).

1.6.2 Rubisco regulation in CAM

Rubisco catalyses the uptake of CO_2 that is released from malate decarboylation behind closed stomata (Phase III) and is also responsible for the direct uptake of atmospheric CO_2 when stomata open during Phase IV. It is believed that Phase IV uptake of CO_2 by Rubisco determines the growth and productivity of CAM species (Nobel, 1996). A range of regulatory mechanisms controls the response of Rubisco to changes in the environment, and should thus serve to modulate C_3 carboxylation in response to CO_2 fluctuating supply occurring over the daytime phases of CAM. Investigations on *K*. *daigremontiana and C. fluminesis* showed changes in initial and final Rubisco activities over the course of the day (Maxwell *et al.*, 2002). Both species displayed highest Rubisco activity and percentage activation towards the end of the day when de-carboxylation is complete, and stomata re-opened with net CO_2 uptake in evidence. Up-regulation of Rubisco at this time, serves to maintain carboxylation strength and WUE, which might help to compensate for diffusion limitations to CO_2 during Phase IV (Maxwell *et al.*, 1997). Low Rubisco activity measured during Phase II seems to be correlated with extended PEPC activation into the photoperiod which might be expected to more effectively scavenge C (in the form of HCO_3), as well as the binding of endogenous

Rubisco inhibitors such as CA 1P (Borland *et al.*, 2000). Rubisco regulation may underpin the plasticity of daytime gas exchange patterns depending on CAM species, which can range from continuous daytime CO_2 uptake as found in CAM cycling species to CAM-idling where stomata remain closed over 24 h (Borland *et al.*, 2000).

1.6.3 Co-ordination of carboxylation and decarboxylation processes

The CAM pathway does not appear to require any special regulation of Rubisco (compared to C_3 plants), but for the efficient nocturnal accumulation of organic acids and daytime de-acidification, Rubisco must be inactive at night and active during the day. Rubisco forms a substantial proportion of protein present in CAM plants (Von Caemmerer and Farquhar, 1981). Rubisco is activated by carbamylation which is the reversible binding of CO_2 to lysine residue in the catalytic site, followed by binding of Mg^{2+} (Lawlor and Cornic, 2002). This activation is facilitated by the chloroplast stomatal protein, Rubisco activase (Lawlor and Cornic, 2002). Rubisco activase activity is regulated through reduction of the large subunit via ferredoxin-thioredoxin reductase (Zhang and Portis, 1999; Dodd *et al.*, 2002) Rubisco and PEPC activities overlap during Phase II and IV of the CAM cycle, but differ between species (Borland and Griffiths, 1997; Maxwell *et al.*, 2002). Rubisco activation status increases slowly during phases II and III and Phase II may be dominated by PEPC. This is due to the delayed activity of Rubisco activase in CAM plants compared with C_3 plants (Maxwell *et al.*, 1999; Maxwell *et al.*, 2002). Rubisco activity is also sensitive to elevated levels of CO_2 (Drennan and Nobel 2000). Both Rubisco and PEPC are greatly influenced by substrate concentrations (Dodd *et al.*, 2002). The supply of ribulose- 1,5-biphosphate requires a sufficient rate of photosynthetic electron transport to regenerate substrate together with enzymatic demand and therefore it is predicted to be limited when light is minimal during Phase II (Dodd *et al.,* 2002).

1.6.4 Diel carbohydrate partitioning

The operation of CAM requires a considerable day/night turnover of carbohydrate, which is essential for providing substrate (PEP) for nocturnal CO_2 uptake and for the growth and productivity of CAM plants. There is considerable

biochemical diversity in the type of carbohydrate used to fuel CAM and growth in CAM species (Kenyon *et al.*, 1985; Christopher and Holtum, 1996; Christopher and Holtum, 1998). Carbohydrate availability is a key limiting factor for the expression of CAM (Borland and Dodd, 2002; Dodd *et al.*, 2003). During Phase III, 75% of carbohydrate synthesised via gluconeogenesis and re-fixation and processing of CO_2 via C_3 photosynthesis, needs to be retained as reserve for carbon assimilation for the following night (Borland *et al.*, 2000). The remaining carbohydrates and any produced from Phase IV are directed towards growth. Some 8-20 % of leaf dry matter is committed each day/night to carbohydrate turnover (Black *et al.*, 1982; Black *et al.*, 1996; Winter and Smith, 1996). A variety of strategies in CAM plants have been observed for C conservation as carbohydrate during the light, which is divided into two groups. One group of species stores mainly starch and glucans in the chloroplasts (Pucher *et al.*, 1949; Sutton, 1975; Madore, 1992; Paul *et al.*, 1993). *Agave* belong to the second group of species, where vacuolar soluble sugars are the predominant form of carbohydrate accumulated during the day and which support the CAM cycle (Smith *et al.*, 1996). CAM plants are further divided according to the major decarboxylases that release CO_2 for re-fixation during the light. Plants having PEPCK as the major decarboxylase occur in families Asclepiadaceae, Bromeliaceae, Euphorbiaceae and Portulacaceae, and species with ME as the major decarboxylase occur in Aizoaceae, Cactaceae, Crassulaceae and Orchidaceae (Dittrich *et al.*, 1973). It is postulated that the variation in carbohydrate partitioning between different CAM species is a result of two principal factors. The first being constraints on C flow imposed by the CAM cycle and the second as different evolutionary histories resulting in a diversity in carbohydrate biochemistry across CAM species (Christopher and Holtum, 1996). Despite the energetic costs associated with carbohydrate synthesis and turnover for CAM, high productivity is not affected. Important CAM species including pineapple (*A. comosus*) and *Agave* can show productivities rivalling that of sugar cane (Bartholomew and Kadzimin, 1977; Nobel, 1996). Growth and productivity of most CAM plants are maximal when direct daytime fixation of CO_2 via Rubisco (Phase IV) predominates (Borland and Taybi, 2004).

For ME species such as *Agave* that store extra-chloroplastic carbohydrate, PEP is exported from the chloroplast but not in exchange for triose-P as occurs in ME starch storing CAM species but rather Pi from extrachloroplastic hexose polymerization (Figs 1.18 A,C)

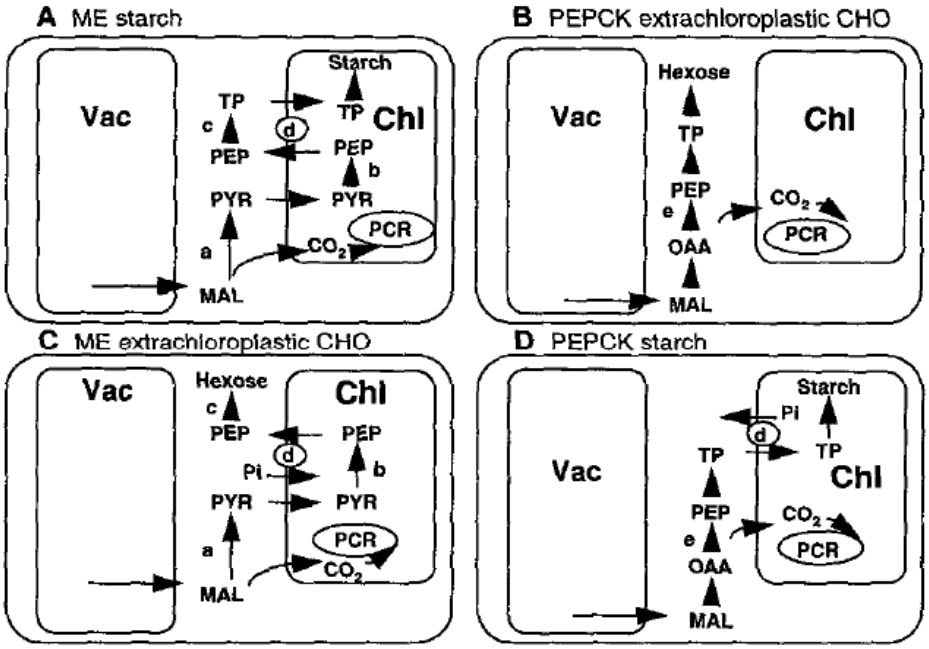


Figure 1.18 Proposed C flow from the four CAM groups: ME starch former (A), PEPCK extrachloroplastic carbohydrate (CHO) former (B), ME extrachloroplastic carbohydrate former (C) and PEPCK starch former (D). Membrane transporters and enzymes indicated: cytoplasmic NADP-ME or mitochondrial NAD-ME (a), pyruvate Pi dikinase (b), enolase and phosphoglyceromutase (c), Pi/triose-P transporter (d), PEPCK (e), Chloroplast;MAL, malate; OAA, oxaloacetic acid; PYR, pyruvate; PCR, photosynthetic C reduction cycle; TP, triose-P; Vac, vacuole (Christopher and Holtum, 1996)

From the 11 CAM species examined by Christopher and Holtum (1996), *Agave. guadalajarana* did not store starch as the major reciprocating carbohydrate. However, the nocturnal depletion of glucose, fructose and sucrose could not account for the C needed for nocturnal PEP regeneration, and a possible use of alternative extra-chloroplastic carbohydrate such as fructans was proposed (Alejandra *et al.*, 2013) Figure 1.18 E). However, the diel fluctuations in sucrose were found to account for more than 83% of carbon needed for nocturnal PEP regeneration in *A. americana*, suggesting differences between *Agave* species in the sorts of carbohydrates used to fuel nocturnal CO_2 uptake (Raveh et al.,

1998). In *Fourcroya humboldiana*, fructans represent the exclusive source of PEP for dark CO_2 fixation (Olivares and Medina, 1990).

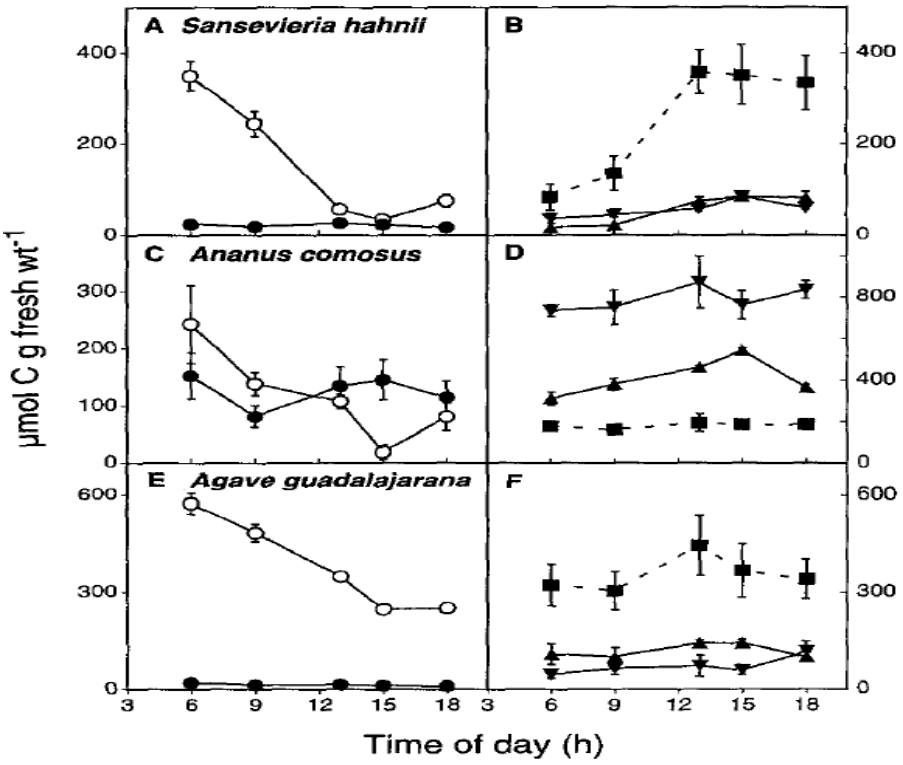


Figure 1.19 Concentration of (A,C,E) malate(⬭) and starch (⬬), and (B,D,F) Glc (▲), Fru (▼) and Suc (■) in the CAM species (A,B) *S. hahnii*, (C,D) *A. comosus*, and (E,F) *Agave guadalajarana*. Dawn was at 5:50 AM and sunset at 6:10 PM. Values are the means ± SE (n=6) (Christopher and Holtum, 1996)

The major decarboxylase was ME for *Agave guardalajarana* shown in Table 1.5

Table 1.5 Maximum extractable activities for decarboxylases PEPCK, NADP-ME, and NAD-ME in crude extracts from 11 CAM species. Values are the means ± SE for (n=3), ND=Not Detectable

Species	**Decarboxylase Activity**		
	PEPCK	**NADP-ME**	**NAD-ME**
A. comosus	247± 52	7 ± 1	3 ± 0.1
P. petropolitana	209 ± 67	21 ± 3	3 ± 1
H. carnosa	105± 29	19 ± 1	2 ± 0.5
S. gigantea	137 ± 43	20 ± 6	1 ±0.5
A. vera	122 ± 25	11 ± 3	5 ± 1
K. tubiflora	ND	5± l	5±2
K. pinnata	ND	25±12	11 ± 3
K. daigremontiana	ND	18 ± **3**	7±5
A.guadalajarana	ND	12 ± 1	11 ± 4
S. hahnii	ND	12 ± 1	4 ± 0.3
V. fragrans	ND	10± 4	7 ± 0.4

1.6.5 Carbohydrate metabolism and sugar allocation in *Agave*

During decarboxylation, in Phase III of CAM, carbohydrate is recovered by gluconeogenesis, ensuring substrate for nocturnal carboxylation and partitioning for growth (Antony and Borland, 2009). As described above, *Agave* species use soluble sugars to provide the substrate (PEP) for dark CO_2 uptake (Black *et al.*, 1996). Thus, carbohydrates that will provide nocturnal substrate for nocturnal reactions in *Agave* are transferred into the vacuole and stored as sucrose, hexose or fructan (Christopher and Holtum, 1996). Vacuolar sugar transporters would seem to play a key role in the diel operation of the CAM cycle in *Agave* (Kenyon *et al.*, 1985; Christopher and Holtum, 1998). The intracellular sugar transport requirements for soluble sugar storing CAM plants are seen in Figure 1.20

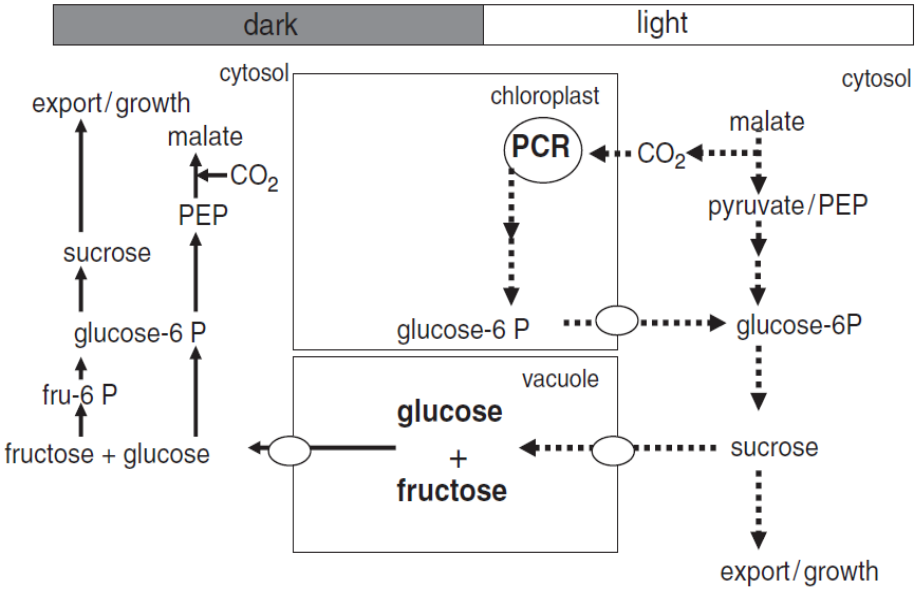


Figure 1.20 Carbon flow and intercellular sugar transport processes for CAM plants using soluble sugars as substrate for nocturnal carboxylation. Dotted lines indicate Day-time fluxes and solid lines are Night-time fluxes. Sugar transporters are represented by the circles located on the chloroplast and vacuole membrane. Adopted from (Antony and Borland, 2009).

Given the central role of soluble sugars in the operation of CAM in *Agave*, genes which encode enzymes involved in carbohydrate metabolism and fructan synthesis may be good candidates for genetic manipulation to enhance fructan accumulation in *agave* for bioenergy production.

1.7 CAM and vacuolar sugar transporters

In CAM plants, the vacuole serves as a storage reservoir for malic acid which accumulates as a consequence of dark CO_2 uptake. In CAM species, an equivalent of 17% of total cell dry mass may cross the tonoplast everyday (Holtum *et al.*, 2005). The three major protein components of the tonoplast are V-ATPases, V-PPases that catalyse the transport of H+ into the vacuole (Marquardt and Lüttge, 1987) and aquaporins (water channels). Other components of the tonoplast are lipids which are likely to play a role in regulating enzyme activity, vesicle trafficking during tonoplast biogenesis, tonoplast protein targeting, signal transduction by membrane lipids and physiochemical properties of the tonoplast (Maeshima, 1992). The tonoplast is composed of several lipids which include phospholipids, free sterols, ceramide monohexoside and digalactosyldiglyceride.

Sugar synthesis represents a main feature of plant physiology which fulfils a number of essential functions that include serving as a general source for metabolic energy and starting points for carboxylate and amino acid synthesis (Heldt and Piechulla, 2004). Sucrose, glucose and fructose are found in high levels in the vacuole (Rees, 1994). In CAM leaves, sucrose import to the vacuole likely occurs by an ATP-independent mechanism due to an existing concentration gradient between the cytosol and vacuolar lumen (Martinoia *et al.*, 1987; McRae *et al.*, 2002) Figure 1.21). Sucrose accumulation is of high importance for photosynthesis (Kaiser and Heber, 1984) and for primary metabolism in storage tissues (Rees, 1994). In CAM plants sugars have an additional key role as providers of phosphoenolpyruvate (PEP), the substrate for nocturnal CO_2 uptake (Antony and Borland, 2009). Typical organic compounds which accumulate in the vacuole are carbohydrates, fructans and carboxylic acids. Malate enters the vacuole either by anion channel specific for malate^{2-} (Hafke *et al.*, 2003) or by a solute carrier (Emmerlich *et al.*, 2003) Figure 1.21).

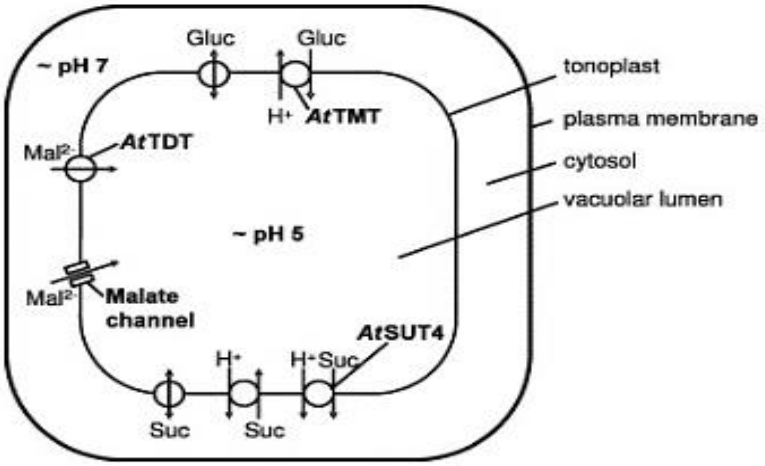


Figure 1.21 Adopted scheme of sugar and malate transport processes across the tonoplast in *Arabidopsis thaliana* vacuoles (Neuhaus, 2007).

Examination of the proteome of vacuolar membranes of *Arabidopsis* cells provided the first evidence on the molecular nature of a vacuolar sucrose carrier (Endler *et al.*, 2006). The first transport proteins involved in the movement of monosaccharides (hexoses) across the tonoplast have been identified which belong to the Tonoplast Monosaccharide Transporter (TMT) group (Wormit *et al.*, 2006). These proteins belong to the monosaccharide transporter (-like) (MST) gene family (Lalonde *et al.*, 2004), and are integral membrane proteins and localized to the tonoplast membranes (Wingenter *et al.*, 2010). *At*TMT were directly identified from *Arabidopsis* with 12 predicted transmembrane α helices and comprised of two units of six connected by central loop varying in length (Lemoine, 2000). The *At*TMT transporters are believed to operate by proton-coupled anti-port mechanism, allowing active transport and accumulation of hexoses (glucose and fructose) in the vacuole especially when induced by cold, drought or salinity. These stimuli promote sugar accumulation in *Arabidopsis* (Wormit *et al.*, 2006). To date, the transporters responsible for sucrose and hexose transfer across the tonoplast membrane have not been identified in *Agave*. It seems likely that such transporters would also be important for regulating fructan content and turnover in *Agave*.

1.8 Project aims and hypotheses tested

A central aim of the thesis was to start to identify traits for the improvement of *Agave* species for biomass production on arid lands. One objective was to examine if the capacity for CAM, and fructan accumulation are linked traits across different CAM species (Chapters 2, 4). The thesis also examined the biochemical basis for differences in CAM activity between *Agave* species (Chapter 3) and set out to identify tonoplast sugar transporters that might regulate CAM and/or sugar accumulation in *Agave* (Chapter 5).

Several hypotheses were tested:

In Chapter 2:

H_1: High leaf succulence is associated with increased magnitude of CAM in *Agave* as manifested as higher nocturnal net CO_2 uptake and nocturnal accumulation of titratable acids,

H_2: Fructan content is positively linked to CAM activity and succulence and is the substrate for nocturnal CO_2 fixation

H_3: Different leaf portions (i.e. leaf tip versus leaf base) in *Agave* play distinct physiological roles in terms of CAM activity and fructan accumulation,

In chapter 3, the aim was to test 4 hypotheses relating to succulence and the biochemical capacity for C3 and C4 carboxylation in *Agave*.

H_1: Abundance of PEPC will vary between species in relation to leaf succulence and age and will vary along the leaf, in line with differences in CAM activity.

H_2: Abundance of Rubisco and Rubisco activase will vary between species in relation to leaf succulence and age and will vary along the leaf but with an inverse relationship to CAM activity,

H_3: In the more succulent Agave species, drought will have less impact on the abundance of PEPC, Rubisco and Rubisco activase compared to the less succulent species

H: The abundance of Rubisco activase will vary over the diel cycle, particularly in leaves of the more succulent species of Agave.

In Chapter 4 screening of inter-specific variation across *Agave* in traits associated with the operation of CAM and fructan accumulation was conducted with tested hypotheses:

H_1: leaf succulence is associated with increased magnitude of CAM across 14 Agave species which will be manifested in nocturnal accumulation of titratable acidities.

H_2: Fructan content is linked with the potential for CAM and leaf succulence across *Agave* species.

H_3: Sucrose rather than fructan is the substrate for nocturnal CO_2 uptake across different species of *Agave*

H_4: Carbohydrate composition influences leaf osmotic pressure in *Agave*

H_5: Specific leaf area is inversely related to the magnitude of CAM in *Agave*.

In Chapter 5, the central aim was to develop a method that could be used to identify candidate vacuolar sugar transporters in *Agave*. A method described for isolating tonoplasts from pineapple was tested for *Agave* leaves. This was followed by a proteomics approach which was used to analyse the purified tonoplast membrane. This involved fractionation of the proteins by SDS-PAGE and analysis by LC-MS/MS, to identify candidate vacuolar sugar transporter proteins which are hypothesized to play a key regulatory role in determining sugar turnover for CAM and fructan accumulation.

Chapter 2

Finding CAM-A-LOT. Is the capacity for CAM in *Agave* related to leaf succulence and fructan accumulation?

2.1 Introduction

Drought is one of the prime abiotic stresses limiting crop production. Agave plants are known to be well adapted and grow naturally in dry, arid conditions, and only require 20% of water for cultivation, when compared to calculated values of crop water demand for the most water efficient C_3 and C_4 crops(Borland *et al.*, 2009). This makes *Agave* good candidates for exploitation on marginal or uncultivated land for bioenergy. *Agave* plants have high cellulose and sugar contents, along with high biomass yield. The high water-use efficiency of *Agave* is due to its crassulacean acid metabolism (CAM), which is adopted by approximately 6 % of plant species as an adaptation to water deficit in terrestrial and epiphytic habitats (Winter and Smith, 1996). Water use-efficiency (WUE) refers to the ratio of CO_2 fixed to water lost. WUE varies according to different environmental conditions such as partial pressure of water vapour in the atmosphere and leaf age, averaging 4-10 mmol CO_2 (mol H_2O)$^{-1}$ for mature CAM leaves over a 24 hour period (Szarek and Ting, 1975; Le Houerou, 1984). WUE is a crucial determinant of success for plants in regions with modest annual rainfall and, in general CAM plants have a greater WUE than do C_3 and C_4 plants (Nobel, 1991).

Leaf succulence is one of the key morphological correlates of the capacity for CAM (Winter *et al.*, 1983; Borland *et al.*, 1998; Griffiths *et al.*, 2008) A survey conducted on *Kalanchoe* (*Crassulaceae*), by (Kluge *et al.*, 1993) found that succulence was positively correlated with the contribution from CAM activity to total carbon gain. Large cell size and succulence are pre-requisites for CAM photosynthesis (Griffiths, 1989; Borland *et al.*, 2000). The large cell size is due to large vacuoles that are important for overnight malic acid storage and which also act as water reservoirs (Osmond *et al.*, 1999; Borland *et al.*, 2000). Such water storage and high WUE associated with CAM can extend periods of net CO_2 uptake under conditions of drought that would be limiting and even potentially devastating for C_3 and C_4 plants (Nobel, 1991).

Agave species are hexose utilizing CAM plants (Black *et al.*, 1996), balancing acidity with water soluble hexoses, and potentially using hexoses as substrates for PEP synthesis. *Agave* also accumulates fructans in the leaves

and their main function is storage (Lewis, 1984). Fructans are water soluble fructose polymers with one glucose moiety per molecule (Sanchez, 2009). The fructans are synthesized in the vacuole by fructosyl transferase enzyme using imported sucrose as a substrate (Valluru and Van den Ende, 2008), and are generally stored in the stems and the leaf bases. In *Agave*, fructans are the major source of ethanol and are also important vacuolar sinks for photo assimilate in mature leaves (Borland *et al.*, 2009). Fructans can also act as osmo-protectants and membrane stabilizers during drought and other abiotic stressors (Wang and Nobel, 1998). This is accomplished by inserting at least part of the polysaccharide into the lipid head group region of the membrane, preventing leakage when water is removed during drought (Livingston Iii *et al.*, 2009). Advantages to the plant in accumulating fructan rather than starch in the leaves include: i) fructan's high water solubility and thus potential use as an osmoticum, ii) fructan resistance to crystallization of membrane at sub-zero temperatures, and iii) continued operation of the fructan synthesis pathway at low temperatures (Vijn and Smeekens, 1999). Fructans also have the potential to drive the CAM cycle by providing the substrate (PEP) for the synthesis of malic acid at night. Fructose can potentially be hydrolyzed from fructan via the enzyme fructosyl transferase and used for PEP synthesis (Black *et al.*, 1996). During the light period, fructans may be re-synthesized from carbon compounds produced by decarboxylation of malate (Marys and Izaguirre-Mayoral, 1995).

To date, most research on *Agave* has revolved around A. *tequilana* due to its economic importance in the tequila production industry. In this species, the pina's, which are swollen stem bases, contain high levels of fructans (Davis *et al.*, 2011b). Production cost of *Agave* per year in Mexico is less when compared to sugarcane production (Sanchez, 2009) and *Agave* produces more ethanol per hectare even with low biomass production due to its high fructan concentration. *Agave* shows economic and environmental advantages over other bioethanol producing crops. It is sustainable because it is an environmentally friendly crop in many ways such as its high water use efficiency; it is a non-food crop and doesn't compete with food crops over fertile land. *Agave* also restrains soil erosion and desertification and can enable carbon sequestration on marginal, degraded land (Borland et al, 2009). Thus, *Agave*

has the potential of producing energy without impacting food security and the environment, plus it is economically sustainable.

A central aim of this chapter was to start to identify traits for the improvement of *Agave* species for biomass production on arid lands, by first examining if the capacity for CAM, and fructan accumulation are linked traits. To address this question, three species of *Agave* that vary in succulence were compared under different water regimes. Measurements were made of leaf gas exchange and titratable acidities as markers of CAM and of soluble sugar and fructan content using high performance liquid chromatography (HPLC).

The experiments specifically addressed 3 hypotheses:

H_1: High leaf succulence is associated with increased magnitude of CAM in Agave as manifested as higher nocturnal net CO_2 uptake and nocturnal accumulation of titratable acids,

H_2: Fructan content is positively linked to CAM activity and succulence and is the substrate for nocturnal CO_2 fixation

H_3: Different leaf portions in *Agave* play distinct physiological roles in terms of CAM activity and fructan accumulation.

With regard to this final hypothesis, it was predicted that the highest CAM activity will be found in the tip region whilst most fructan accumulation will occur in the base of the leaf. These predictions will indicate if CAM activity and fructan accumulation are subject to contrasting anatomical and physiological control processes.

2.2 Materials & Methods

2.2.1 Plant material, watering regimes and sampling strategy

The *Agave* species under investigation were *Agave americana* (most succulent = 3.15 Kg m^{-2}) (Figure 2.1.A), *A. angustifolia* (succulence= 2.54 kg m^{-2}) (Figure 2.1 B) and *A. attenuata* (least succulent = 0.91 Kg m^{-2}) (Figure.2.1 C). All plants were maintained under controlled conditions of a 12 hour photoperiod and day/night temperatures of 28/22^{O}C. Soil was made up in 127 mm pots containing a mixture of 1 part sharp sand (J. Arthur Bower's, UK),4 parts John Innes No. 3 (JI no. 3),1 part gravel. Plants were exposed to two watering regimes, namely 70% field capacity (F.C.) and 20% F.C. In order to impose the different water regimes, plants were first droughted for approximately two weeks and plant, plus soil and pot was weighed. This represented 0% F.C (A). The plants were then re-watered for several days until water was freely draining from the bottom of the pot and weighed again. This weight this represented 100% F.C (B) .The following equations were used to calculate how much water had to be added to the plants to achieve 20% & 70% F.C.by calculating what the weight of plant, plus soil and water would be at 20 or 70% F.C

For 70% F.C:

Weight of plant, soil and water = A + ((B-A/100) x 70)) [2.1]

For 20% F.C:

Weight of plant, soil and water = A + ((B-A/100) x 20) [2.2]

For gas exchange, titratable acidity and carbohydrate measurements, unless indicated otherwise, all measurements were made on leaf No.4 (mature) counting from the centre of the rosette. For acidity and carbohydrate measurements, leaf discs with an area of (2.36 cm^2) were collected from different leaf portions (tip, middle, base), both at dawn and dusk periods, snap frozen in liquid N_2 and stored at -80^{o}C.

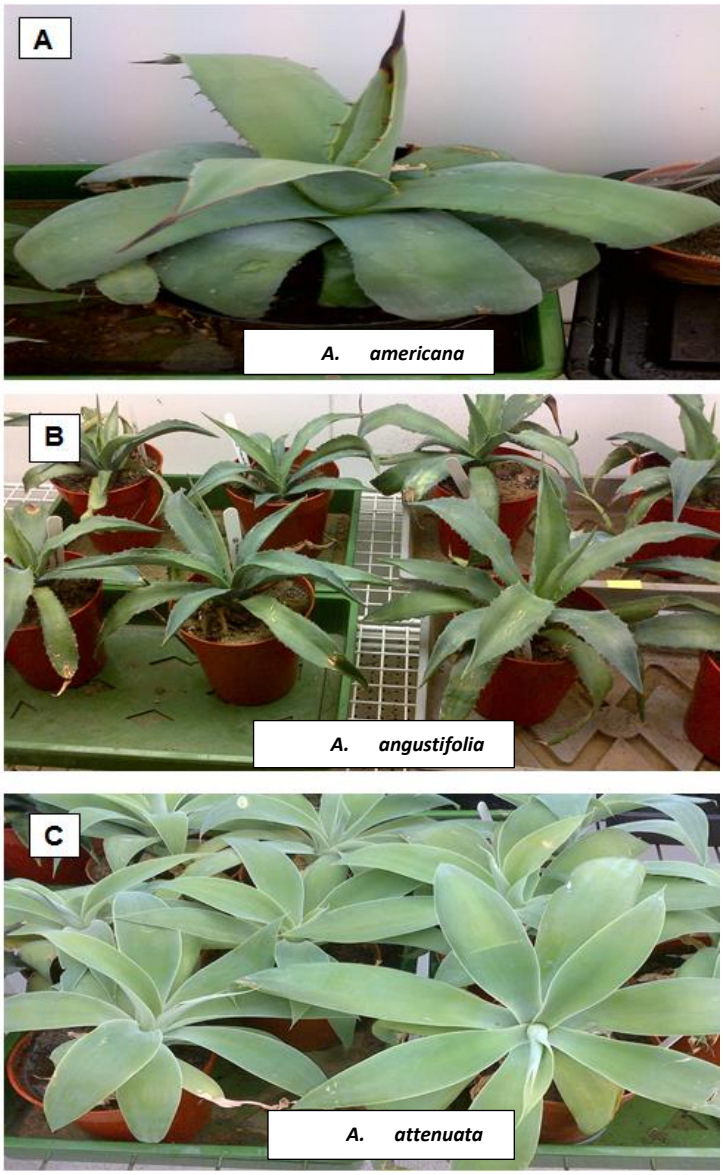


Figure 2.1 *Agave* species varying in leaf succulence (A) *A. americana*, (B) *A. angustifolia* and (C) *A. attenuata*.

2.2.2 Leaf gas exchange profiles and instantaneous water use efficiency (WUE)

Net CO_2 uptake was measured using a Walz CMS-400 Compact Mini Cuvette system (Heinz Walz, Effeltrich, Germany) with BINOS-100 infrared analyser (IRGA). This provided a direct, non-destructive method of measuring instantaneous and daily carbon gain. Direct CO_2 measurements identify the relative contribution of the four phases of the CAM cycle to total carbon gain. Fully expanded mature leaves (leaf No.4) were maintained in a cuvette for 24-48 hours (Figure 2.2), with 4 biological replicates taken for each *Agave* species.

Conditions of light and temperature in the cuvette tracked those in the growth room. Data for net CO_2 uptake and evapo-transpiration were recorded every 15 minutes, using an open gas exchange system. Net CO_2 uptake was determined by difference in CO_2 mole fractions between gas entering and leaving the cuvette (equation 2.3). This approach follows the work of Von Caemmerer and Farquhar (1981).The gas flow was maintained between 400 and 500 ml min^{-1} avoiding water condensation inside the cuvette. Data were analysed using DIAGAS software based on the area of the leaf inside the cuvette.

Figure 2.2 Leaf gas exchange measurements made by clamping cuvette on fully expanded *Agave* leaves.

Agave leaves were maintained in the cuvette for up to 48 hours in order to obtain a reproducible 24 h pattern of leaf gas exchange. Data were logged every 15 minutes and differential zero point measurements taken every 10 data collection periods.

$$A = \frac{\mathrm{Um\ (Ce - Co)}}{\mathrm{S}} \qquad [2.3]$$

Where

A= net rate of CO_2 uptake per leaf area (μmol $m^{-2}s^{-1}$)

Um= molar flow rate (mol s^{-1})

S= leaf area (m^{-2})

C_e-C_O= difference in me fraction between CO_2 entering and leaving the cuvette

C_e and C_O are equivalent to the reference gas and the measuring of CO_2 mole fractions, respectively, so C_e-C_O is CO_2 ppm differential between reference and measuring gas flows. The molar gas flow U_m is calculated from the volumetric flow rate (Uv; m^3 s^{-1}), and that one mole of an ideal gas volume equals 0.0224 m^3 at 273.15 K and 101.3kPa (equation 2.4) (Holum, 1994)

$$um = \frac{\text{uvx273.15xp}}{\text{0.0224x101.3xT}} \qquad [2.4]$$

Where p is atmospheric pressure (kPa) and T is temperature (K).

Water Use Efficiency was calculated over a 24 h light/dark cycle by:

$$\text{WUE mmol } CO_2 \text{ per mol } H_2O = \frac{\text{Amount of } CO_2 \text{ fixed by photosynthesis}}{\text{Amount of water lost by transpiration}} \qquad [2.5]$$

Total leaf area was calculated by scanning and analysing via *Image J* software (Appendix A gives details of use of Image J for leaf area measurements).

2.2.3 Titratable Acidity

Titratable acidity analysis was used as a marker for CAM expression along the leaves of three species of *Agave* varying in succulence, under two watering regimes (20% and 70% field capacity) as assessed by differences in acidity measured at dawn and dusk.

Samples were collected at dawn and dusk for *Agave* species varying in succulence, and different leaf portions (tip, middle, base) were sampled, 4

biological replicates each. Samples were wrapped in foil, snap frozen in liquid N_2 and stored at -80°C until analysis.

About 200 mg of frozen leaf tissue (weight recorded using a Sartorius balance) was ground in liquid nitrogen using a pestle and mortar. Tissue was heated in 5ml 80% methanol at 80°C for 40 minutes. Exactly 1ml extract was then diluted with 2ml of distilled water and titrated against 0.005M NaOH to neutrality, using 3 drops of phenolphthalein as an indicator. The number of moles (Z) of H^+ in 5ml extract was calculated using the following equations:

$$Z \text{ (moles } H^+) = \text{NaOH titre} \times 0.005/1000 \times 5 \qquad [2.6]$$

$$Z/\text{fwt} = \text{moles } H^+ \, g^{-1}\text{fwt (fresh weigh basis)} \qquad [2.7]$$

$$Z \times 10000/\text{area of 4 discs in } cm^2 \text{ (moles } H^+ \, m^{-2}\text{) (Area basis).} \qquad [2.8]$$

2.2.4 High Performance Liquid Chromatography (HPLC)

The amounts of sucrose, fructose, glucose, inositol and sorbitol present at dawn and dusk in samples taken from the 3 *Agave* cultivars (3 biological replicates of each) were determined using high-pressure liquid chromatography (HPLC). The methanol extract was desalted via ion exchange using columns of Dowex AG50W X4 – 200 (Sigma-Aldrich,USA) and Amberlite IRA – 67 (Sigma-Aldrich, USA) in series. To prepare the ion exchange columns, exactly 30 g each of Dowex and Amberlite were used. Dowex was washed with 95% ethanol with one change over 30 minutes to remove the color and then rinsed with several changes of de-ionized water. Amberlite was washed with 4 to 5 volumes of 1M NaOH for 30 minutes and rinsed with de-ionized water to neutrality. Then the columns were prepared by placing a thin layer of glass wool at the bottom of a 2.5 ml plastic syringe and carefully layered with 0.5 cm^3 of Amberlite then 0.5 cm^3 of Dowex on top. The columns were then washed with high-grade water multiple times before adding the extract to the top of the column. Exactly 200 µl of the extract were passed through the column. To completely collect the

desalted extract, the column was washed with 3 ml of high-grade water. Exactly 20 µl of eluent was injected into an HPLC via a Rheodyne valve onto a Carbopac PA-100 column (Dionex, Sunnyvale, California, USA). Approximately, 100 µl of sample was inserted into an analysis vial so as to ensure optimal immersion of the auto-sampler syringe. Sample components were eluted from the column isocratically using 100 mM NaOH (de-gassed by helium gas) flowing at 1 ml/min for 8 min at room temperature. The chromatographic profile was recorded using pulsed amperometric detection with an ED40 electrochemical detector (Dionex, Sunnyvale, California, USA). Elution profiles were analysed using the Chromeleon software package (Thermo Fisher Scientific Inc., MA, USA). Daily reference traces were obtained for glucose, fructose and sucrose by injecting calibration standards with concentrations of 20 ppm. for each sugar (Adams *et al.*, 1992). Standards were run after every ten samples. Total fructan quantification was analyzed using the Subtraction Method (Liu et al, 2011), involving two steps of HPLC analysis. First, levels of free glucose, fructose and sucrose were measured. Second, total glucose and fructose were measured after hydrolysis of fructans was performed by adding 150mM concentrated HCL and incubating samples at 80^{o}C for 90 min.

2.2.4.1 HPLC analysis of sugars

An eluent of 50% NaOH (7.7ml) was added to 1 liter of nano-pure water, and was left standing over night. Standards of glucose, fructose and sucrose were run through HPLC (20 ppm) to calibrate. Samples were then injected and analyzed.

Calculation of sugar contents:

Grams of sugar in 20 µl injection = ppm x 20/1,000,000 = **Y** [2.9]

Y x150 (amount of sugar in 3 ml washed through column) = **Z** [2.10]

Amount of sugar in starting extract = **Z** x 5 = **P** (took 200 µl of 1 ml of extract to pass through column)

[2.11]

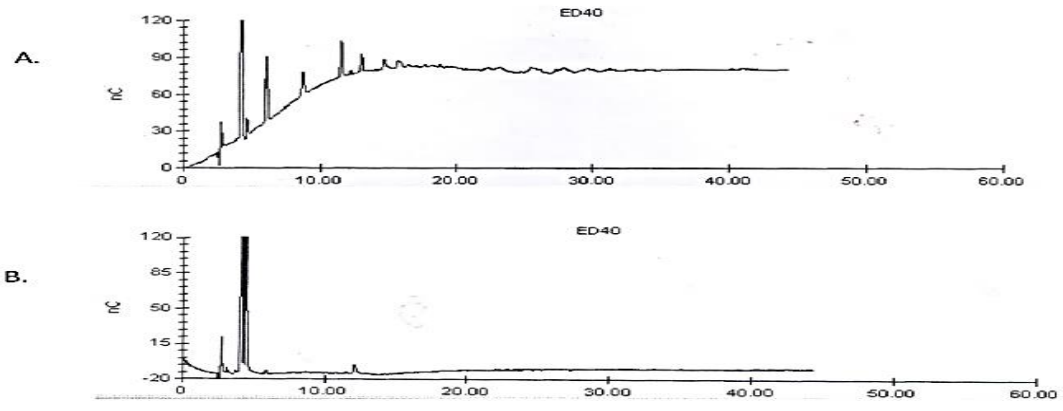
Moles of sugar g fwt = **P**/180/fwt discs [2.12]

Moles of sugar per m^2 area = **P**/180 x 10,000/area of discs in cm^2 [2.13]

2.2.4.2 HPLC of fructan oligosaccharides

For further fructan analysis, 2 ml of the 3ml extract collected from ion exchange wash was placed in a clean tube and dried down overnight. It was later taken up in 200μl nano-pure water and vortexed thoroughly. A preparation of 3 eluents was made. The first was eluent B: 90 mM NaOH (7.2 ml of 50% NaOH made up to 1 L with nano-pure water). The second eluent was eluent C: 350 mM sodium acetate in 90 mM NaOH (28.7 g NaAcetate, 800 ml nano-pure water, 7.2 ml 50% NaOH). It was made up to a volume of 1 L with nano-pure water. The final eluent was eluent D: 1 M NaOH (80 ml of 50% NaOH made up to 1 L with nano-pure water). All eluents were left to stand overnight. The HPLC was set to run an acetate gradient from 20 to 350 mM for around 40 min, followed by 10 min of 1 M NaOH to regenerate column and 20 min equilibrium of 20 mM sodium acetate in 90 mM NaOH. Standards of ketose, neoketose and kestopentaose (25 ppm) were run through the HPLC to calibrate. A few targeted samples from *Agave* were analyzed which had a running time over 70 minutes.

Total leaf fructans were analysed using the Subtraction Method, (Liu *et al.*, 2011) involving two steps of HPLC analysis. First, levels of free glucose, fructose and sucrose were measured. Second, total glucose and fructose were measured after hydrolysis of fructans was performed by adding 150mM and incubating samples at 80^{o}C for 90 min (Figure 2.3).



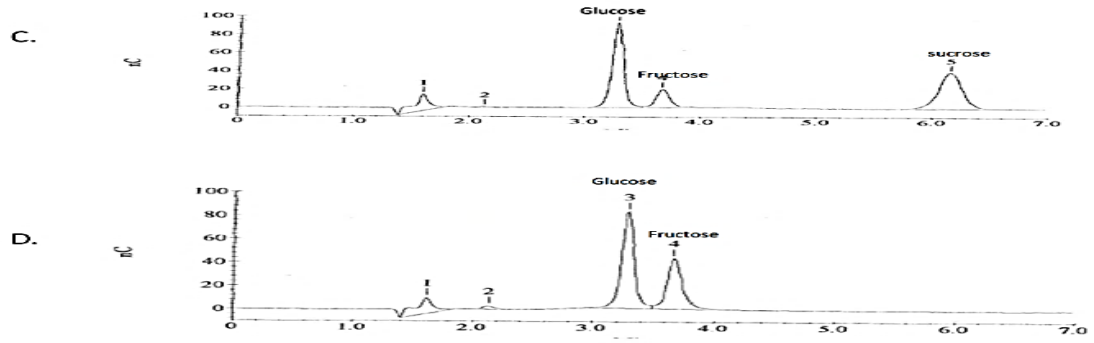


Figure 2.3 Sodium Acetate Gradient HPLC profile of water-soluble carbohydrates extracted from *Agave attenuata*, leaf base at dusk. (A) Before hydrolysis showing the presence of high molecular weight fructans (B) after acid hydrolysis with 150 mM HCl showing accumulation of fructose residues. Sodium hydroxide isocratic HPLC (100 mM NaOH) in (C, D) *Agave attenuata*, leaf base at dusk. (C) Before hydrolysis (showing glucose, fructose and sucrose); (D) after acid hydrolysis with 150 mM HCl, (showing glucose & fructose).

2.2.5 Statistical Analysis

All data presented are the mean values expressed from four replicates ± standard error (S.E.) in each group. Where appropriate, data were analyzed using SPSS (IBM SPSS Statistics 21 64Bit) and graphs were produced using Microsoft Office Excel 2010. Normal distribution was tested using Normality test ($P > 0.005$) and significant differences between mean values were verified using a post hoc Least Significant Difference test (LSD) ($P < 0.05$) following one-way ANOVA.

2.3 Results

Three different *Agave* species (*A. attenuata, A. americana, A. angustifolia*) varying in leaf succulence were compared under two watering regimes (70% and 20% field capacity). Net CO_2 assimilation and titratable acidity measured the magnitude of CAM against the degree of leaf succulence. Soluble sugars and fructans in *Agave* were quantified using phenol/sulphuric acid method (Dubois *et al.*, 1956) and profiled using High Performance Liquid Chromatography (HPLC).

2.3.1 Gas exchange profiles & water use efficiency (WUE)

Net CO_2 uptake was measured for each species over 24 h. Each gas exchange curve is representative of that obtained for 4 biological replicates. The most succulent *A. americana* (3.15 kg m^{-2}) achieved the highest nocturnal net CO_2 uptake under both watering regimes, (see Table 3.1). The proportion of net dark CO_2 uptake to day-time uptake increased under drought conditions in all 3 cultivars (Figure 2.4). Under well watered conditions i.e. 70% F.C, for both *A. americana* and *A. angustifolia*, highest rates of dark net CO_2 uptake were noted at the start of the night (beginning of Phase I) with A. *angustifolia* briefly exceeding A. *americana*, before declining over the rest of the night There was no phase II in either *A. angustifolia* or *A. americana*. During Phase III (behind closed stomata), no net CO_2 uptake was observed, but net CO_2 uptake commenced again later in the photoperiod in Phase IV. The least succulent species *A. attenuata*, seemed predominantly C_3 under well watered conditions with most net CO_2 uptake occurring during the day under 70% F.C. Under drought conditions, most net CO_2 uptake occurred in Phase I for all three species, and rates of net dark CO_2 were enhanced under the droughted conditions for all 3 species. The 20% F.C treatment resulted in a reduction of Phase IV for both *A. angustifolia* and *A. attenuata* but had no effect on the most succulent species *A. americana.* A little Phase II was present for both *A. angustifola* and *A. attenuata* with a slight surge of net CO_2 uptake at the start of the photoperiod under 20% F.C.

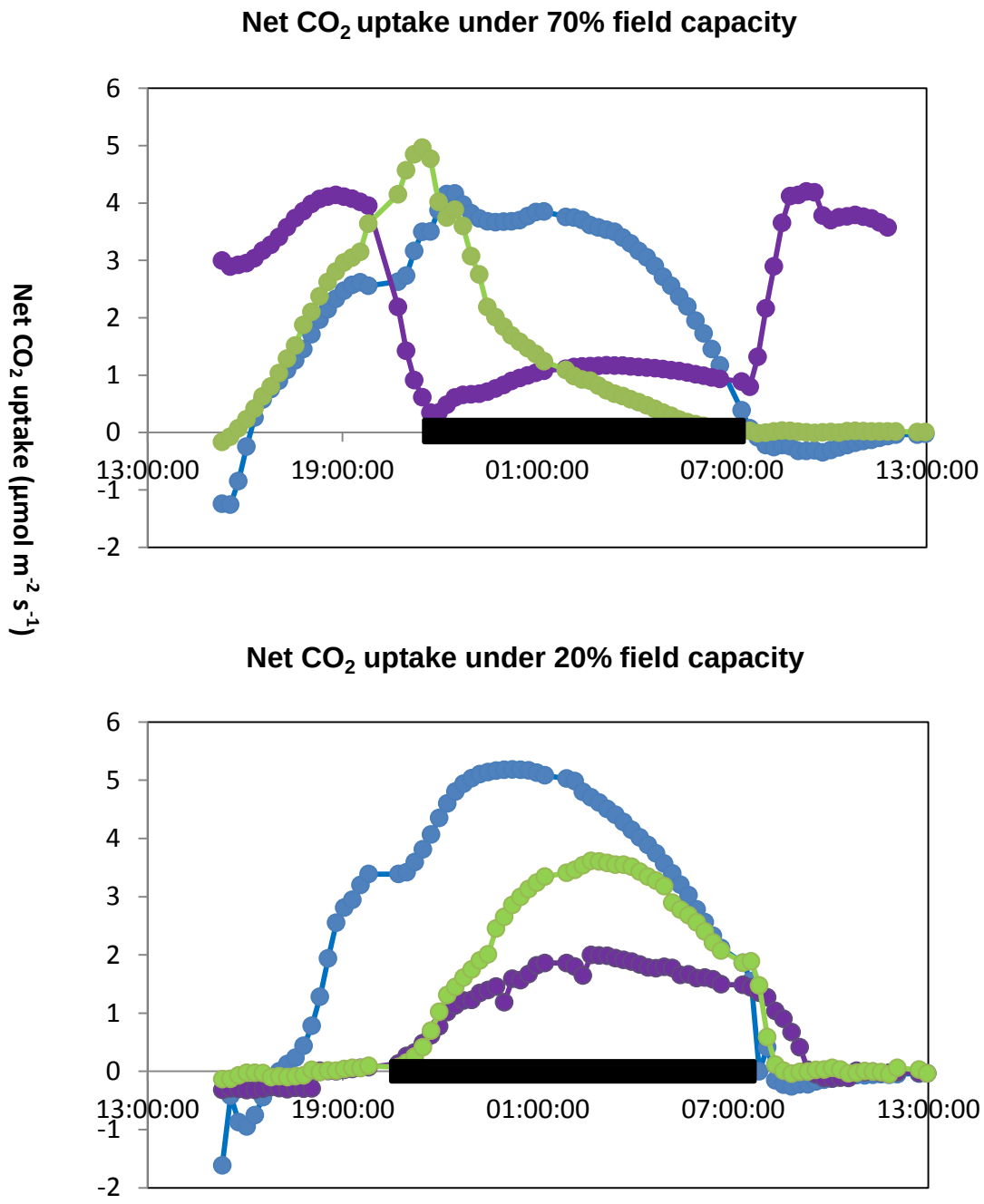


Figure 2.4 Net CO_2 assimilation by *Agave americana* (succulence = 3.15 kg m^{-2}), *Agave angustifolia* (succulence = 2.54 kg m^{-2}) and *Agave attenuata* (succulence = 0.91 kg m^{-2}) over a 24-h light/dark period under 20 and 70% field capacity. The black bar on the x-axis represents the dark period.

The most succulent species *A. americana* had the highest WUE which showed a positive relationship to the magnitude of nocturnal CO_2 uptake. Table 2.1

The data suggest that higher leaf succulence serves to buffer water availability, maximizing nocturnal net CO_2 uptake even under conditions of drought.

Table 2.1 Water-use efficiency and nocturnal CO_2 uptake of 3 investigated A*gave* species varying in succulence *

Agave cultivars	Field Capacity	Water-use efficiency (mmol CO_2 mol $^{-1}$ H_2O)	Nocturnal CO_2 uptake (mmol CO_2 m^{-2})
A. attenuata	20%	3.2	20.81
	70%	3.8	14
A. angustifolia	20%	7.6	88.49
	70%	6	57.55
A. americana	20%	8.14	148.4
	70%	9.0	111.69

*Water-use efficiency (mmol CO_2 mmol $^{-1}$ H_2O) of *Agave attenuata, Agave angustifolia and Agave americana* under 20% and 70% field capacity.

2.3.2 Titratable Acidities

Titratable acidity analysis identified nocturnal acid accumulation as a marker for CAM expression along the leaves of three species of *Agave* varying in succulence, under two water regimes (20% and 70% field capacity) as assessed by differences in acidity measured at dawn and dusk (Figure 2.5).

The magnitude of CAM (i.e. the difference in acidity measured at dawn-acidity measured at dusk) showed a gradient in CAM expression along the leaf decreasing from tip to base of the leaf, and was highest in the most succulent cultivar (*A. americana*) when expressed on a leaf fresh weight basis under well watered conditions (70% field capacity), (*p=0.013*).

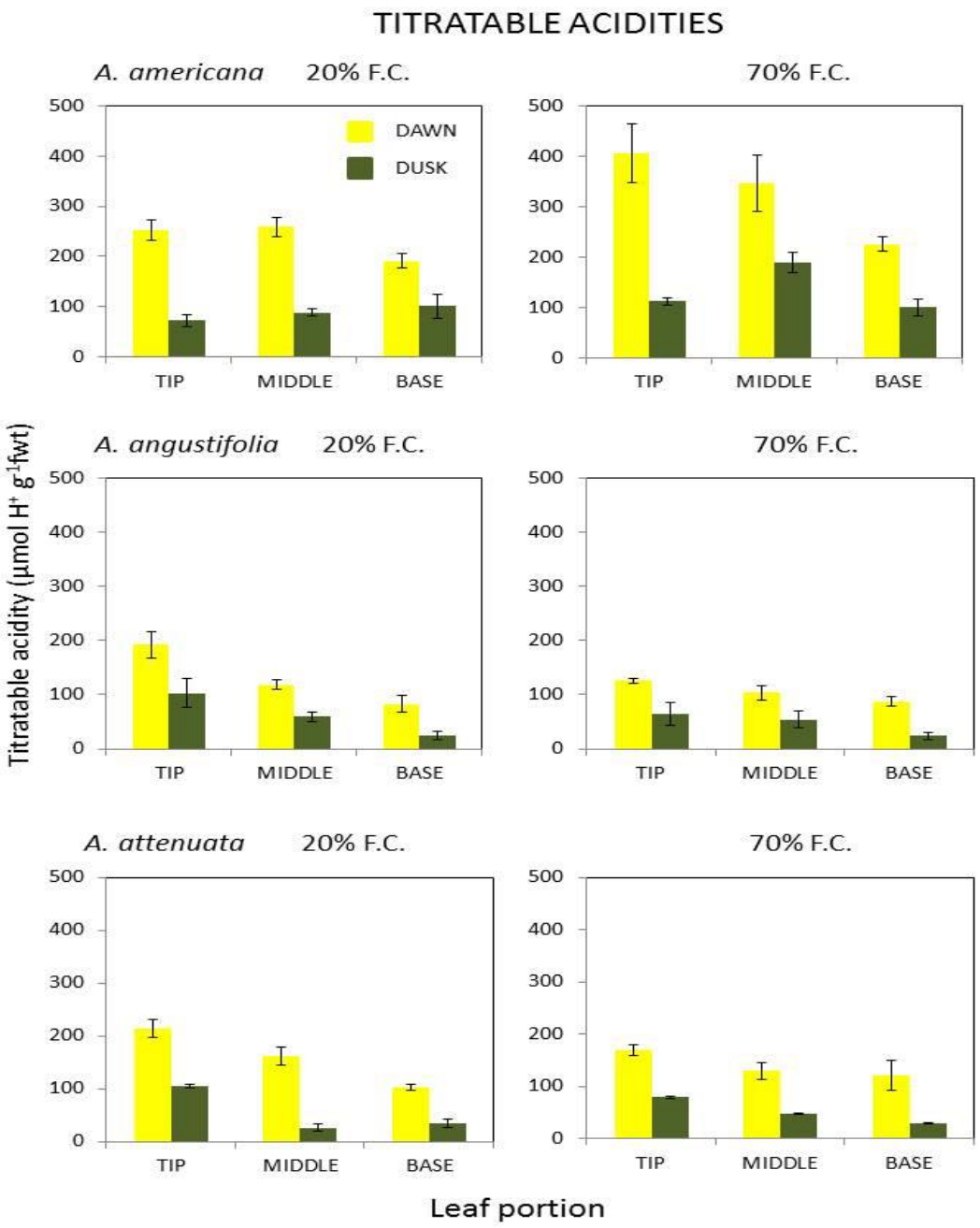


Figure 2.5 Day/night changes in titratable acidity along the leaves of three *Agave* cultivars varying in succulence under 20% and 70% field capacity, expressed on a fresh weight basis (µmol H^+ g^{-1} fwt) for dawn and dusk periods(n = 4 ± standard errors).

The dawn-dusk acidities were calculated and expressed as malate (2 H^+ = 1 malate; Table 2.2). The response to drought in terms of nocturnal malate accumulation differed between species and portion of the leaf. Drought (20 % F.C.) stimulated malate accumulation in the leaf tip and mid-leaf sections in both *A. attenuata* and *A. angustifolia*. However, drought stimulated nocturnal malate accumulation was only evident in the middle section of leaves of *A.*

americana (Table 2.2). In general, drought had little impact on nocturnal malate accumulation in the leaf bases of any of the *Agave* species under investigation (Table 2.2).

2.3.3 Fructan accumulation

Fructan content generally increased from the tip to the base of the leaf and was higher in the two most succulent *Agave* species (i.e. *A. americana* and *A. angustifolia*) when expressed on a leaf fresh weight basis. There was no significant impact of watering regime on fructan content or day/night turnover. Only *A. attenuata* showed significant day/night turnover of fructans and this was most evident in the tip and middle portions of the leaves (Figure 2.6)

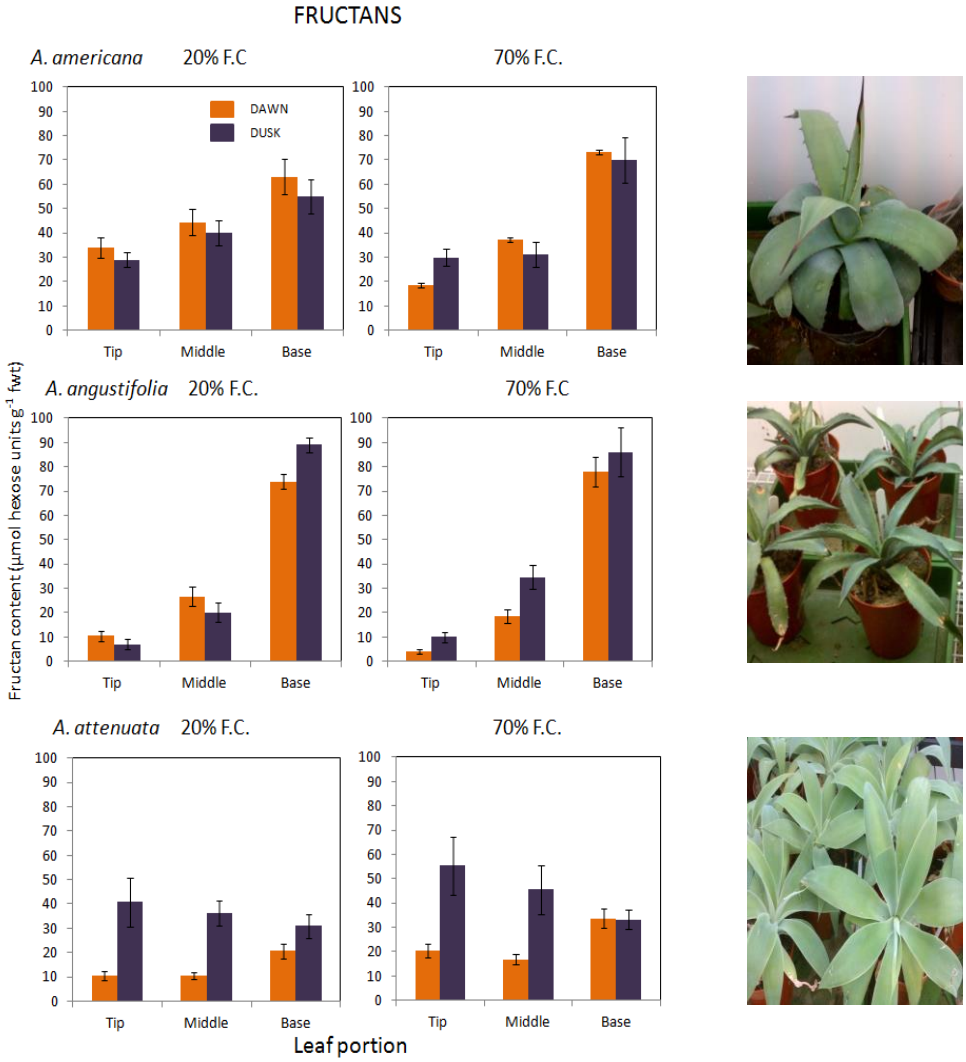


Figure 2.6 Fructan content along the leaves of three *Agave* cultivars, under 20% and 70% field capacity with samples taken at dawn and dusk and expressed on a leaf fresh weight basis (µmol hexose units g^{-1} fwt). (n = 4 ± standard errors).

Sucrose was present in greater abundance compared to fructose and glucose. Sucrose content and diel turnover decreased from tip to base (Figure 2.7) .There was no significant effect of watering regime on sucrose content. All three species of *Agave* followed the same trend of sucrose decreasing from tip to base of the leaf with higher levels at dusk compared to dawn.

In contrast to sucrose, the patterns observed for glucose (Figure 2.8) and fructose (Figure 2.9) content tended to increase from the tip to the base of the leaf, except for *A. attenuata* under well watered conditions (70% F.C). This pattern of higher glucose and fructose contents towards the base of the leaf was similar to the pattern observed for fructan (Figure. 2.6). The glucose content of leaves was generally higher than that of fructose. In general there was little effect of watering regime on glucose or fructose contents.

The potential amounts of phosphoenolpyruvate (PEP) that could be generated from nocturnal depletion of different sugar fractions from different leaf portions for the 3 *Agave* species, maintained under contrasting water regimes is displayed in Table 2.2. This was compared with measured nocturnal malate accumulation with the assumption that 1 mole PEP gives rise to 1 mole malate. From this data it appears that sucrose was the major sugar for nocturnal acid production in all 3 *Agave* species. Only in *A. americana* did it seem that nocturnal breakdown of fructans might be required to generate PEP in the tip and middle portions of the leaf. The two other *Agave* species had an excess of soluble sugar breakdown at night which could more than account for the PEP needed for malic acid accumulation.

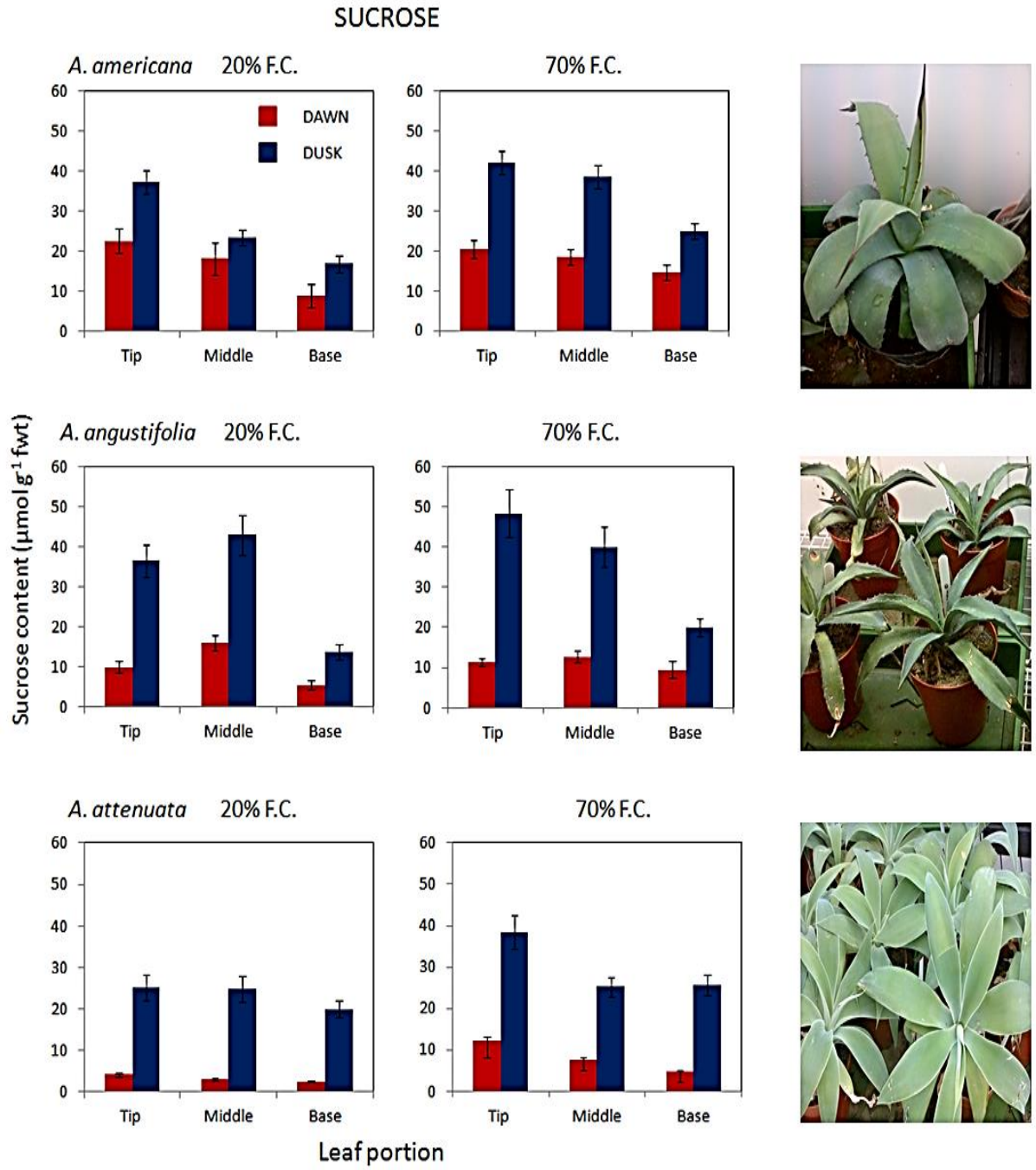


Figure 2.7 Day/night changes in sucrose content in three *Agave* species expressed on a fresh weight basis (µg g^{-1} fwt) under 20% and 70% field capacity and, measured at different positions of the leaf. (n = 4 ± standard errors).

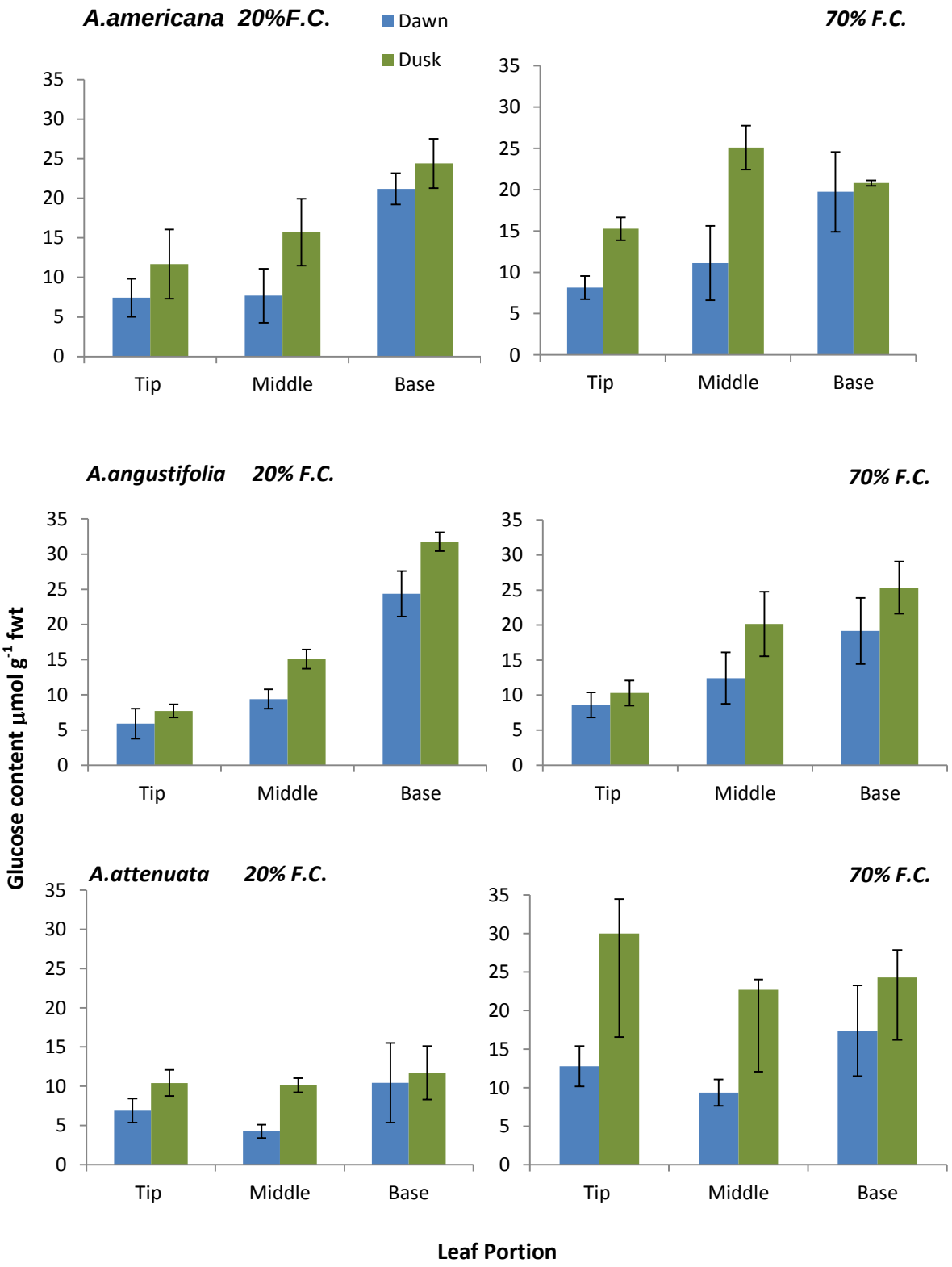


Figure 2.8 Glucose content along the leaves of *A. americana, A. angustifolia & A. attenuata* expressed on fresh weight basis (µg g^{-1} fwt) under 20% and 70% field capacity with samples taken at dawn and dusk.(n = 4 ± standard errors).

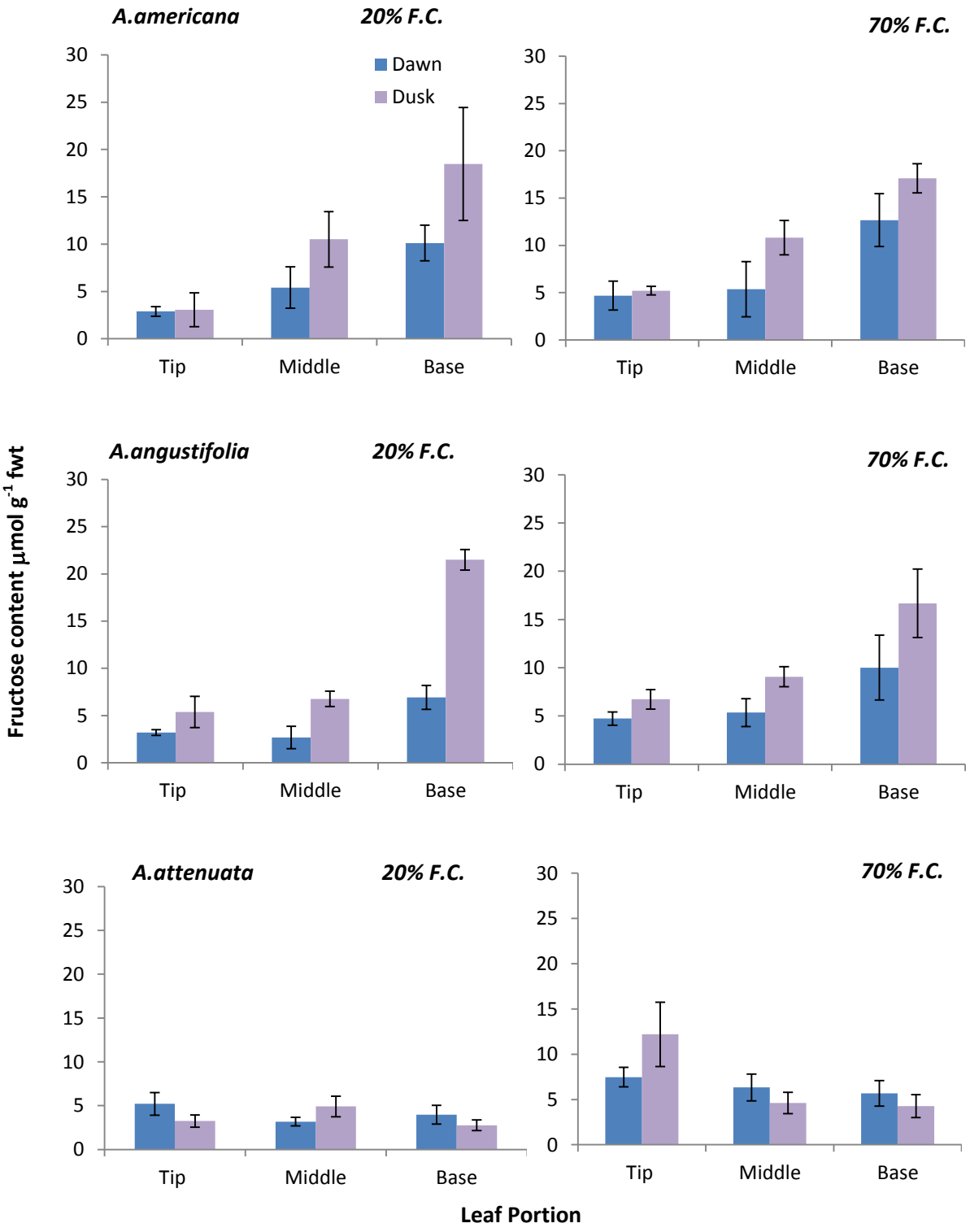


Figure 2.9 Fructose content along the leaves of *A. americana, A. angustifolia & A. attenuata* expressed on fresh weight basis (µg g^{-1} fwt) under 20% and 70% field capacity with samples taken at dawn and dusk. (n = 4 ± standard errors for error bars indicated).

Table 2.2 A summary of nocturnal malate accumulation (estimated from titratable acidity measured at dawn and dusk) and the potential amounts of phosphoenolpyruvate (PEP) that could be generated from the nocturnal depletion of different sugar fractions from different leaf portions in three species of *Agave* maintained under 20% or 70% field capacity (+) Indicates PEP shortfall, (-) excess of sugars.

	Leaf portion	μ mol g^{-1} fwt				
		Δ **malate**	Δ PEP **SUCROSE**	Δ PEP **GLUCOSE**	Δ PEP **FRUCTOSE**	**PEP shortfall**
A. americana 20% F.C	TIP	90.04	58.8	8.5	0.38	+22.36
	MID	84.99	20.8	16.1	10.2	+37.89
	BASE	45.03	31.6	6.42	16.69	-9.68
A. americana 70% F.C	TIP	187.55	86.4	14.22	1.04	+85.89
	MID	68.12	80	27.98	10.9	-50.76
	BASE	62.79	41.6	2.2	8.86	+10.13
A. angustifolia 20% F.C.	TIP	44.48	106.4	3.6	4.34	-69.86
	MID	29.1	108	11.4	8.16	-98.46
	BASE	29.5	33.2	14.8	29.2	-47.7
A. angustifolia 70% F.C.	TIP	30.71	148	3.42	3.98	-124.69
	MID	24.55	109.2	15.47	7.41	-107.53
	BASE	31.89	41.6	12.39	13.32	-35.42
A. attenuata 20% F.C.	TIP	54.9	83.6	7.04	3.92	-39.66
	MID	67.9	86.8	11.83	3.46	-34.19
	BASE	34.21	69.6	2.548	2.9	-40.838
A. attenuata 70% F.C.	TIP	27.75	104.8	34.46	9.48	-120.99
	MID	51.22	40.8	26.66	3.42	-19.66
	BASE	59.35	83.6	13.87	2.82	-40.94

NOTE: Δ malate = (mean dawn TA – mean dusk TA)/2 [2.14]

Δ PEP FRUCTAN = (mean dusk fructan – mean dawn fructan) x 2 [2.15]

Δ PEP SUCROSE = (mean dusk sucrose – mean dawn sucrose) x 4 [2.16]

Δ PEP GLUCOSE = (mean dusk glucose – mean dawn glucose) x 2 [2.17]

Δ PEP FRUCTOSE = (mean dusk fructose – mean dawn fructose) x 2 [2.18]

Glucose and fructose chemical formulae $C_6H_{12}O_6$, so each mol of glc or fru can generate 2 moles PEP. Sucrose chemical formula $C_{12}H_{22}O_{11}$, so one mole suc can generate 4 moles PEP.

2.4 Discussion

The aim of the study was to test 3 hypotheses related to succulence, the magnitude of CAM and fructan accumulation in three species of *Agave.*

2.4.1 Leaf succulence determines CAM expression under contrasting water regimes

Certain species of *Agave* display impressive rates of biomass production (Simpson *et al.*, 2011a), which might be associated with several anatomical and physiological adaptations that ensure continued growth and survival under water limiting conditions, with the expression of CAM photosynthesis being the most important character. As predicted, the data presented in this chapter showed that the magnitude of CAM increased with succulence, being the highest in *A. americana*, followed by *A. angustifolia* and *A. attenuata.* The higher CAM activity in *A. americana* was manifested in a higher ΔH^+, and higher rates of nocturnal net CO_2 uptake. High vacuolar capacities maximize the amount of CO_2 that can be taken up by PEPC, converted to malate and stored in the vacuole during phases I and II, enhancing photosynthetic carbon gain of CAM species (Osmond *et al.*, 1999). The findings that the magnitude of nocturnal CO_2 fixation tends to be greater in thicker leaved, more succulent *Agave* species has been reported for other CAM species inhabiting arid regions (Teeri *et al.*, 1981; Winter *et al.*, 1983). The tight cell packing which accompanies increased leaf succulence seems to enhance CAM efficiency by reducing CO_2 leakage in phase III but restricts access of CO_2 during C_3-mediated phase IV by reducing internal CO_2 conductance (g_i) (Maxwell *et al.*, 1997; Borland *et al.*, 2000; Nelson and Sage, 2008). However, reduced g_i may be essential to CAM function by limiting efflux of CO_2 released from malate decarboxylation during phase III therefore promoting overall carbon economy (Nelson *et al.*, 2005) and could be one of the selection pressured influencing CAM evolution (Griffiths, 1989). Reduced g_i does not appear to limit atmospheric CO_2 uptake in phase I because vacuolar capacity and PEP availability are probably the main controls over night time CO_2 acquisition (Maxwell *et al.*, 1997; Osmond *et al.*, 1999; Borland *et al.*, 2000).

In *Agave*, the large, generally succulent leaf rosettes, also serve to buffer abrupt and longer term changes in water availability, helping to maximize nocturnal CO_2 uptake and extend the duration of atmospheric CO_2 demand

beyond the night period. Also, shallow root systems, which are typical for *Agave* species, allow rapid uptake of sudden precipitation. Reports in the literature have shown that 24 h CAM activity in *Agave* seems to operate with reduced Phases II and IV. However, as in other CAM plants, drought seems to influence 24 h patterns of CO_2 uptake in *Agave* plants (Nobel and Hartsock, 1978; Nobel, 1985; Nobel *et al.*, 1998). A higher fraction of daytime CO_2 uptake was lost compared to night time CO_2 uptake (Nobel, 1985) in *A. fourcryodes* exposed to 11 days of drought which exhibited a reduction of 99% in net daytime CO_2 uptake and 76% in night time CO_2 uptake (Nobel, 1985). For the work described in this chapter, the proportion of net dark CO_2 uptake to day-time uptake increased under drought conditions in all 3 *Agave* species. A certain level of photosynthetic plasticity was observed in the 3 *Agave* species examined, allowing them to modulate the contribution of daytime (Phase II and III) and night-time (Phase I) carbon acquisition when faced with different environmental factors. Under well watered conditions, Phase II was reduced for the 2 most succulent species, and the least succulent *A. attenuata* showed that net CO_2 uptake was dominated by day-time, C_3 fixation under well watered conditions. Some A*gave* species such as *A. deserti* are able to change from CAM to C_3 as manifested in daytime CO_2 uptake and no day/night acid fluctuations, under well watered conditions (Hartsock and Nobel, 1976). When facing different environmental conditions, photosynthetic plasticity has been observed in young and adult plants of *A. tequilana* which can adjust carbon gain during daytime (Phase II and III) and nighttime (Phase I) (Pimienta-Barrios *et al.*, 2001). Even though most CO_2 uptake occurs at night (Phase I) (Nobel *et al.*, 1998) , it has been observed in young and adults of *A. tequilana* that at least some Phase IV CO_2 uptake can be maintained during the driest months of the year. This phenomenon is not common amongst other CAM plants growing in arid environments (Pimienta-Barrios *et al.*, 2001). Leaf succulence seems to determine how plastic CAM expression can be. Dodd *et al*., (2002) revealed that thinner leaved *Kalanchoë* species (i.e. *K. pinnata)* were highly plastic in photosynthetic expression and displayed more day-time CO_2 uptake compared to thicker leaved, more succulent species (i.e.*K. daigremoniana*) (Dodd *et al.*, 2002), which seem to have diffusional constraints to CO_2 uptake (Maxwell *et al.*, 1997) which makes them more bound to nocturnal CO_2 fixation for 24 h C supply. In the data presented in this chapter, the least succulent *Agave* species

(A. *attenuata*) displayed similar behavior to the thin leaved *K. pinnata* showing high plasticity in photosynthetic expression under 20% and 70% F.C.

Internal water supply is crucial to ensure high photosynthetic performance in plants growing in water-limited habitats. Studies on *Agave* species have demonstrated that leaf succulence is the key for allowing substantial net CO_2 uptake even when soil water content is low (Pimienta-Barrios *et al.*, 2001). However, young leaves of *A. tequilana* which are less succulent than mature leaves, and therefore have lower internal water storage, were able to exhibit almost matching photosynthetic assimilation rates during both dry and wet seasons (Pimienta-Barrios *et al.*, 2001). This could be due to continuous water movement from the medullar hydrenchyma to the marginal chlorenchyma during the dry season, allowing the occurrence of relatively high levels of CO_2 assimilation year-round, even in young leaves (Pimienta-Barrios *et al.*, 2001). The large storage parenchyma does not participate directly in the CAM cycle but is vital in the recharge of the chlorenchyma and maintenance of overall tissue water status (Smith *et al.*, 1987; Yakir *et al.*, 1994; Borland *et al.*, 2000). *Agaves* face many challenges living in arid environments such as high rates of evaporation, so having internal water storage tissues are more appropriate than an external water reservoir such as found in tank bromeliads (Alejandra *et al.*, 2013)

2.4.2 Flexibility of carbohydrate source pools to sustain dark CO_2 uptake in *Agave*

Carbohydrate turnover is an essential component determining the magnitude of CAM (Borland and Dodd, 2002). There is a large biochemical commitment of between 8 to 20% of total cell dry matter into the diel cycle (Black *et al.*, 1996). A distinguishing feature of *Agave* is the production of fructans, which are polymers of B-fructofuranosyl residues synthesized from sucrose and stored in vacuoles of the parenchyma of leaves and stems. Fructan content and metabolism are closely related to frost and drought tolerance (Pontis, 1989; Coninck *et al.*, 2007; Valluru and Van den Ende, 2008). The data presented in this chapter indicated that the most succulent *Agave* species under investigation, *A. americana* accumulated larger amounts of fructans than the less succulent species. Thus CAM activity and fructan

accumulation appear to be linked traits. In a study on *A. americana* (Raveh *et al.*, 1998), evidence was provided that fructans are not generally broken down during the dark period to provide PEP as a substrate for nocturnal CO_2 fixation. In the present study, there was no appreciable day/night turnover of fructan in the two most succulent species, but nocturnal fructan depletion was noted in the tip and middle leaf portions of *A. attenuata.* The nocturnal depletion of fructan was also implied in a study on *A. guadalajarana*, in which there was insufficient glucose, fructose or sucrose breakdown at night to account for the required PEP production/malate accumulation (Christopher and Holtum, 1996). However, a survey of *A. humboldiana*, showed an inverse relation between fructans and malic acid (Olivares and Medina, 1990). Together, the findings described above suggest that there may be genotypic variation across *Agave* in the source of carbohydrate used to provide PEP for nocturnal CO_2 uptake.

It has been suggested elsewhere that *Agave* utilizes soluble hexose sugars as their carbohydrate reservoir, which are stored in the vacuole (Black *et al.*, 1996). Other studies have observed diel fluctuations in leaf sucrose which could account for more than 83% of carbon needed for PEP regeneration in *A. americana* (Raveh *et al.*, 1998). This finding is in general agreement with results of this chapter. Thus, nocturnal sucrose depletion decreased from tip to base, in line with the decrease in nocturnal accumulation of titratable acids. Sucrose was the major sugar used for nocturnal acid production in *Agave*. In the bromeliad *Aechmea maya*, sucrose became the major source of carbohydrate for nocturnal carboxylation as drought progressed (Ceusters *et al.*, 2009). Sucrose was the major reserve carbohydrate in the 3 species tested in this chapter, providing substrate for nocturnal PEP production. In contrast, fructose and glucose are the major sugars used for nocturnal acid production in *A. comosus* (Carnal and Black, 1989) and *Clusia minor* (Popp *et al.*, 1987). Stoichiometric analyses of sugar breakdown and PEP requirements for CAM indicated that of the 3 *Agave* species studied in this chapter, only *A. americana* showed a shortfall in sucrose for PEP, implying that some nocturnal fructan depletion may be required in this species to provide PEP. Flexibility of major carbohydrate source used for the sustainability of dark CO_2 uptake is crucial for energy demands and carbon acquisition for environments with limited precipitation.

2.4.3 Different parts of the *Agave* leaf show contrasting physiological roles in terms of CAM and fructan accumulation

Agaves are rosette plants with new leaves produced in the center of the rosette. Variations in the magnitude of CAM differed along the leaf of 3 *Agave* species varying in succulence. At the leaf level, nocturnal changes in titratable acidity increased with distance from the leaf base, and the highest CAM activity was found at the tip. This data is consistent with that as shown in *Fourcroya humboldtiana* (Olivares and Medina, 1990), and in *Guzmania monostachia*, with a significant rise in the levels of nocturnal accumulation of titratable acidity in the apical region (tip) (Freschi *et al.*, 2010). Within the plant, the base is shaded by the blades of upper leaves, therefore, a CAM gradient may be expected from the base to the tip (Olivares and Medina, 1990). In contrast, most fructan accumulation occurred in the base of the leaf. This might compromise CAM and malate storage in leaf base if sugars are preferentially directed towards the storage of fructans. High vacuolar capacities maximize the amount of CO_2 that is taken up by PEPC, converted to malate and stored in the vacuole during phases I, and II, enhancing carbon gain (Osmond *et al.*, 1999).The results presented here showing contrasting expression of CAM and fructan accumulation along the leaf indicate that CAM and fructan accumulation are subject to contrasting anatomical and physiological control processes.

2.5 Conclusions

As shown in this study and elsewhere (Kluge *et al.*, 1993; Kluge and Brulfert, 1996; Kluge *et al.*, 2001; Griffiths *et al.*, 2008), high leaf succulence is associated with increased magnitude of CAM, manifested as higher ΔH^+ and nocturnal CO_2 uptake. Fructan accumulation also increased with leaf succulence in *Agave*. Sucrose provided most, if not all of the substrate required for dark CO_2 uptake. Lower water availability enhanced the proportion of dark CO_2 uptake but did not influence fructan accumulation. At the leaf level, highest CAM activity was found in the tip region whilst most fructan accumulation occurred in the base of the leaf. These results indicate that CAM and fructan accumulation are subject to contrasting anatomical and physiological control processes.

It is not clear if increased vacuolar capacity for malate accumulation and CAM activity is accompanied by increased investment in PEPC protein (Winter *et al.*, 1982; Borland *et al.*, 1998). Further work is needed to understand the biochemical capacity of C_3 and C_4 carboxylation in *Agave* in order to examine if this is related to succulence. This question will be considered in Chapter 3.

Chapter 3

Is leaf succulence related to the biochemical capacity of C_3 and C_4 carboxylation in *Agave*?

3.1 Introduction

Agave is a succulent genus known to be well adapted and grow naturally in dry, arid conditions. In general, *Agave* requires only 20% of water for cultivation, when compared to calculated values of crop water demand for the most water efficient C3 and C4 crops (Borland *et al.*, 2009). The high water-use efficiency of *Agave* is due to its crassulacean acid metabolism (CAM), which is adopted by approximately 6 % of plant species as an adaptation to water deficit in terrestrial and epiphytic plants (Winter and Smith, 1996). Putting it at the simplest level, CAM is a photosynthetic system in which the C_3 (Rubisco) and C_4 (PEPC) carboxylases occur in a common cell with temporal separation of enzyme activity (Dodd *et al.*, 2002). Leaf succulence is one of the key morphological correlates of the capacity for CAM (Winter *et al.*, 1983; Borland and Griffiths, 1989; Griffiths *et al.*, 2008). Surveys on the genus *Kalanchoë* (Crassulaceae), found that succulence is positively correlated with the contribution from CAM activity to total carbon gain (Kluge *et al.*, 1993; Kluge *et al.*, 2001). Other studies have reported that succulence and the magnitude of CAM display a positive relationship in a taxonomically diverse range of CAM lineages (Sage, 2002; Nelson *et al.*, 2005). Large cell size and succulence are pre-requisites for CAM photosynthesis (Borland *et al.*, 2000), due to the requirement for large vacuoles that are important for overnight malic acid storage and which also act as water reservoirs (Osmond *et al.*, 1999); (Borland *et al.*, 2000). However, relatively few studies have considered the implications of this morphology on the biochemical properties of CAM (Griffiths *et al.*, 2008). For example, it is not known if increased leaf succulence is accompanied by increased abundance of the C4 (PEPC) as well as the C3 (Rubisco) carboxylases.

During the night, the stomata open in CAM plants, allowing CO_2 to enter the mesophyll cells of the leaf, and be fixed as organic acid by the enzyme phosphoenolpyruvate carboxylase (PEPC). The CAM form of PEPC needs to be active at night but inactive during the day in order to avoid futile cycling of organic acids which would result in the hydrolysis of ATP. The day/night regulation of PEPC is also important for avoiding competitive carboxylation with Rubisco which is active during the day. The day/night regulation of PEPC is accomplished through reversible phosphorylation catalysed by PEPC kinase which is exclusively regulated at the level of transcript abundance (Hartwell *et*

al., 1999; Taybi *et al.*, 2000). Phosphorylation renders PEPC insensitive to malate inhibition, thus PEPC can be active at night (Nimmo *et al.*, 1984; Nimmo *et al.*, 1986; Grams *et al.*, 1997). The product of PEPC-mediated carboxylation is malate which accumulates in vacuoles of the cell, during phase I of the CAM cycle. PEPC regulation by reversible phosphorylation restricts C_4 mediated CO_2 uptake to Phase I and early Phase II, thus curtailing futile cycling of CO_2 during the day during carboxylation dominated by Rubisco (Dodd *et al.*, 2002). Phase II is a transitional phase between dominating PEPC-mediated and Rubisco-mediated CO_2 fixation (Griffiths *et al.*, 1990) when stomata open during the early hours of the light period. A peak of CO_2 fixation is often noted during this phase due to both fixation of CO_2 by PEPC and direct assimilation via Rubisco (Acevedo *et al.*, 1983; Lüttge, 1986; Maxwell *et al.*, 1998). The decarboxylation of malate (Phase III), occurs during daytime when stomata are closed. Malate exits the vacuole passively following a downhill gradient (Lüttge and Nobel, 1984). CO_2 is released and concentrated around the enzyme Ribulose-1,5-biphosphate carboxylase oxygenase (RuBisCo) and thus entering the Calvin Cycle to ultimately produce carbohydarte. Rubisco is activated by the enzyme Rubisco activase, which functions to promote and maintain the catalytic activity of Rubisco (Lawlor and Cornic, 2002) Stomata re-open during Phase IV, due to exhaustion of malate and internal CO_2 concentrations drop. Direct fixation of atmospheric CO_2 is via Rubisco, for the remainder of the light period (Borland *et al.*, 2009). The duration of each phase of the CAM cycle varies between species, response to the environment and leaf development (Winter *et al.*, 2008).

Leaf succulence also influences the phases of CAM, as illustrated in Chapter 2. In *Agave*, the more succulent species fixed CO_2 predominantly at night (Phase I) while the least succulent species (*A. attenuata*) fixed CO_2 during Phases I, II and IV. High degrees of leaf succulence reduce intercellular airspace (IAS) between mesophyll cells and a reduction to length of mesophyll cell length exposed to intercellular air space (L_{mes}/area; (Smith and Heuer, 1981; Maxwell *et al.*, 1997; Nelson *et al.*, 2005; Nelson and Sage, 2008). These traits reduce internal CO_2 conductance (Borland *et al.*, 2011) which can provide higher photosynthetic efficiency to CAM plants that rely heavily on dark CO_2 uptake (Phase I), with 70% of carbon gained at night. These plants are known as strong CAM plants, and leaf $\delta^{13}C$ value of *Agave* species are typically in the

strong CAM range. The close cell packing in succulent leaves minimizes loss of C previously fixed during the day (Griffiths, 1992). It has also been proposed previously (Bartholomew and Kadzimin, 1977; Winter *et al.*, 1985; Borland *et al.*, 1994) that atmospheric CO_2 fixed directly by Rubisco at the end of the day (Phase IV) contributes substantial carbon for growth in high yielding CAM species. Research has indicated that increased succulence (dense cell packing) reduces CO_2 availability for Rubisco during Phase IV (Maxwell et al, 1997). It might be postulated that succulent CAM species compensate for this by either investing in more Rubisco protein or by activating Rubisco more effectively during Phase IV (i.e. via increased abundance of Rubisco activase) in order to maximise draw down and uptake of CO_2 across the leaf.

Succulence in *Agave* would appear to represent a key trait for enhancing CAM activity by providing a high vacuolar storage capacity for malic acid, maximizing nocturnal PEPC capacity and potentially extending its activation for several hours in the day. Extending Phase II is beneficial for carbon gain by delaying the onset of Phase III decarboxylation until the warmest, brightest time of day (Borland *et al.*, 1996). This could improve the efficiency of Rubisco refixation of CO_2 and minimize the net efflux of CO_2 during Phase III, which also maximizes carbon gain in mature *Agave tequilana* (Borland *et al.*, 2011)

The aim of this chapter was to establish if the level of leaf succulence influences the investment in C3 and C4 carboxylases in *Agave*. It was hypothesized that the more succulent species of *Agave* will have higher PEPC protein abundance. In terms of Rubisco abundance, two scenarios were postulated; 1) there is an inverse relationship between PEPC and Rubisco protein abundance or 2) the more succulent species have higher abundance of Rubisco and/or Rubisco activase in order to maximise CO_2 uptake and draw-down across the densely packed cells of the leaf.

Four hypotheses relating to succulence and the biochemical capacity for C3 and C4 carboxylation in Agave were tested.

H_1: Abundance of PEPC will vary between species in relation to leaf succulence and age and will vary along the leaf, in line with differences in CAM activity.

H2: Abundance of Rubisco and Rubisco activase will vary between species in relation to leaf succulence and age and will vary along the leaf but with an inverse relationship to CAM activity,

H3: In the more succulent *Agave* species, drought will have less impact on the abundance of PEPC, Rubisco and Rubisco activase compared to the less succulent species

H4: The abundance of Rubisco activase will vary over the diel cycle, particularly in leaves of the more succulent species of *Agave*.

Measurements of 24 h changes in titratable acidity and soluble sugar content were made to assess the magnitude of CAM expression in two species that varied in succulence, namely *A. americana* and *A. attenuata*. Abundances of PEPC, Rubisco and Rubisco activase were compared between species, between leaf ages and between base and tip of the leaf. The impact of drought on the abundance of PEPC, Rubisco, RA as well as leaf growth was also examined. Finally, an interrogation of transcriptome and proteome databases for *A. americana* database was conducted to examine 24 h changes in transcript and protein abundances for PEPC and Rubisco activase in mature (succulent, full CAM) and young (less succulent, low CAM) leaves of *A. americana*.

3.2 Materials & Methods

3.2.1 Plant Material

The *Agave* species under investigation were *A. americana* (most succulent species, mature leaf succulence= 3.15 Kg m^{-2}) *A. attenuata* (less succulent species, mature leaf succulence = 0.91 Kg m^{-2}). All plants were maintained under controlled conditions of a 12 hour photoperiod and day/night temperatures of 28/22^{O}C. Soil was made up in 127 mm pots containing a mixture of 2 parts sand (East Riding Horticulture Ltd, UK), 8 parts John Innes No. 3 (JI no. 3), 2 parts grit and 0.5 mg Osmocote. Plants were watered twice a week. For leaf samples that were collected for westerns (tip vs. base), plants were maintained under a 16 hour photoperiod. Leaf samples were collected over a 24 hour period. Plants were exposed to well watered conditions (70%

F.C.) and drought conditions (20% F.C.). See section 2.2.1 for calculations of 20% and 70% F.C.

3.2.2 Titratable Acidity

Titratable acidity analysis was used as a marker for CAM expression. Measurements of leaf titratable acidity were made using samples taken over a 24 hour cycle. Samples were collected every four hours for the two *Agave* species varying in succulence, and for different leaf ages (unfolded, young, mature), 3 biological replicates for each. Samples were wrapped in foil, snap frozen in liquid Nitrogen and stored at -80^{o}C until analysis. About 200 mg of frozen leaf tissue (weight recorded using a Sartorius balance) was ground in liquid nitrogen using a pestle and mortar. Tissue was heated in 5ml 80% methanol at 80°C for 40 minutes. Exactly 1ml extract was then diluted with 2 ml of distilled water and titrated against 0.005M NaOH to neutrality, using 3 drops of phenolphthalein as an indicator. The number of moles of H^+ in extracts were calculated using equations described in section 2.2.3.

3.2.3 Soluble Sugar Analysis

Soluble sugar analysis was determined using a colorimetric method (Dubois *et al.*, 1956), using the same methanol extracts used for titratable acidity measurements. Simple sugars give an orange yellow precipitate when treated with phenol and concentrated sulfuric acid. The volume of methanol extract analyzed must fall within the linear range of glucose calibration. Exactly 20 μl of plant extract (*A. americana*) was added to 480 μl H_2O and 0.5 ml 5% phenol and then 2.5 ml concentrated sulphuric acid was added. For *A. attenuata*, 30 μl of plant extract was added to 470 μl H_2O and 0.5 ml 5% phenol and then 2.5 ml concentrated sulphuric acid was added. Samples were mixed with a glass rod and left to cool for 15 minutes. Readings were taken at 483 nm using a spectrophotometer (GENESYS 10 VIS, UK) and compared with glucose standards of known concentration from 0 to 150 μg. Results were expressed as mmol glucose equivalent per unit leaf area or as μmol glucose equivalent per g fresh weight.

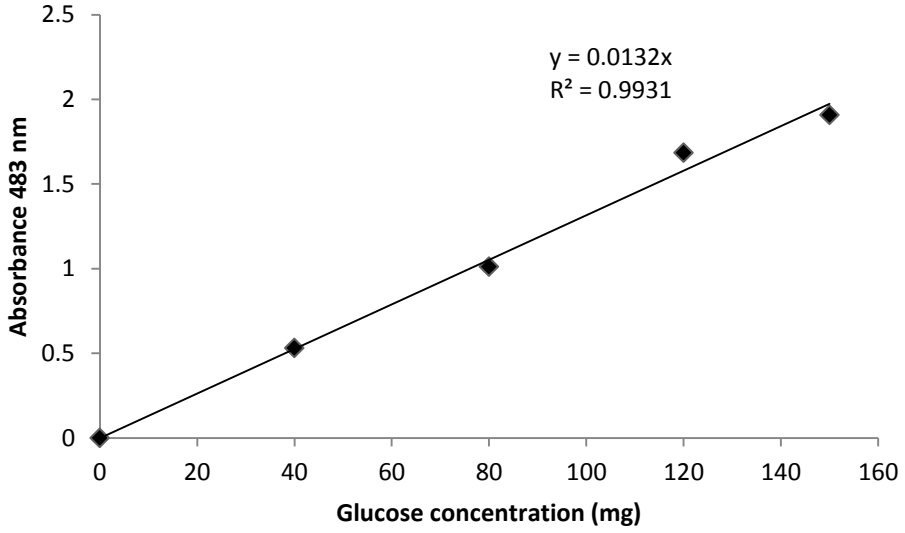


Figure 3.1 Example of linear calibration curve for determining leaf total soluble sugar content

3.2.4 Western blotting for Rubisco, PEPC and Rubisco activase

3.2.4.1 Sample preparation

Samples were prepared from frozen leaf discs that had been harvested every 4 hours over a 24 hour cycle, snap frozen in liquid nitrogen then stored at -80^{0}C until analysis. Samples were ground to a powder by adding liquid nitrogen in a pestle and mortar. Each sample was weighed to 250 mg and placed in a 2 ml Eppendorf tube, adding 280 µl of chilled extraction buffer (300 mM Tris pH 8.3, 100 mM NaCl and 2% PEG (Polyethylene Glycol 20,000) for 70% F.C samples. For the 20% F.C samples 280 µl of chilled extraction buffer (1M Tris pH 8.3, 100 mM NaCl and 2% PEG 20,000) was used. Also, 50 µl DTT (100 mM), 10 µl phenylmethylsulfonyl fluoride(PMSF)10 mM, 40 µl E-64, 40 µl Leupeptin, 40 µl protease inhibitor cocktail (Sigma-Aldrich) and 40 µl EDTA (16 mM), were added. Samples were left for 1 minute on ice then were mixed by inversion and shaking. Samples were then centrifuged for 10 minutes at 13,000 rpm at 4 ^{O}C using a microcentrifuge (Eppendorf 5417R). The supernatant was removed and added to a fresh eppendorf tube and centrifuged for 5 minutes at 13,000 rpm. Once again the supernatant was collected and 10% (v/v) glycerol was added. Samples were snap frozen in liquid nitrogen and stored at -80 ^{O}C.

3.2.4.2 Protein estimation

Protein contents of plant extracts were determined by a colorimetric assay as described by Bradford (1976). This is a protein determination method which involves protein binding to Coomassie Brilliant Blue G-250 (Bradford Reagent), causing a shift in absorption maximum of the dye from 465 to 595 nm, which is monitored (Bradford, 1976). Samples were analysed with a spectrophotometer to determine their absorbance at 595 nm. Bradford reagent was prepared by dissolving 100 mg of Coomassie Brilliant Blue (Sigma-Aldrich) in 50 ml 95% ethanol (v/v) and orthophosphoric acid (v/v), adjusting the volume to 1 litre with distilled water, and storing the solution in a brown bottle, and shaken before use. In each cuvette, 100 μl of water was added to 20 μl of extracted sample. Finally a volume of 4 ml of Bradford reagent was added to each cuvette. Samples were analysed after 15 minute incubation. Samples were compared with a standard curve using bovine serum albumin (BSA), ranging from 0-140 μg protein per ml for all experiments. The blank was made up of 100 μl of deionised water and 4 ml of Bradford reagent.

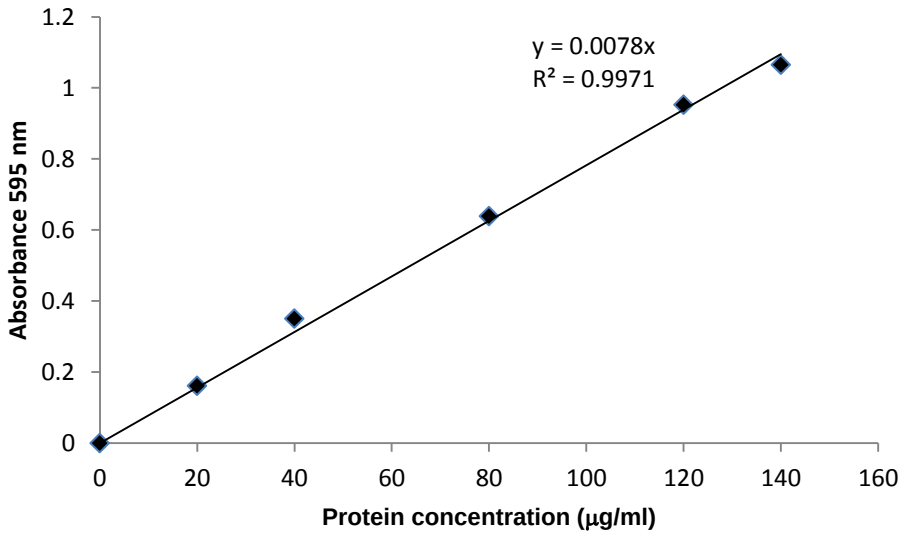


Figure 3.2 Example of linear calibration curve for Bradford method of determining total soluble proteins

3.2.4.3 Discontinuous SDS-PAGE gel preparation for protein separation

Proteins were separated by molecular mass, using polyacrylamide gel electrophoresis SDS-PAGE (Laemmli, 1970) with vertical Mini-Protean II ^{TM} gel system (Bio-Rad Laboratories Ltd, Hemel Hempstead, Hertfordshire). The multiphasic system employs a separating gel in which samples are fractionated, and a lower percentage stacking gel added above it. In the stacking gel, sample components are stacked into thin, sharp zones prior to separation. One large and one small glass plate were used per gel (0.75 mm thick, 7 cm long and 8 cm wide), and cleaned with acetone and further rinsed with distilled water using transfer pipette. Plates were blot dried with blue roll. Both glass plates and spacers were assembled in the clamp. The clamp was tightened in the casting stand, and placed on the casting stand using the grey rubber strips to seal the bottom of assembly. The glass plate was marked 1 cm below level of well, indicating the level of separating (resolving) gel.

First, the separating gel mixture was prepared in a glass beaker, using quantities set out in Table 3.1. As soon as it was prepared, the gel mix was transferred into the glass plates within the gel apparatus using a pipette. The gel was overlaid with 200 μl of 1X buffer (taken from LWGB pH 8.8 and diluted with distilled water). The gel was placed in the cold room overnight slowing down the process of polymerisation. After the separating gel had set, indicated by the formation of a clear line between buffer and gel, the buffer was removed and washed twice with distilled water. The stacking gel was prepared as indicated in Table 2.1 and immediately poured on the top of the separating gel. A Teflon comb was inserted to create loading wells and the stacking gel was allowed to set for 30 minutes in the cold room. The comb was removed and gels were immersed in an electrophoresis tank filled with reservoir buffer (25mM Tris-HCL pH 8.3, 200 mM glycine and 1% (w/v) SDS).

Table 3.1 Components of separating & stacking gel for SDS-PAGE electrophoresis of proteins

	30% acrylamide solution (ml)	Deionised water (ml)	LWGB buffer (ml)	UPGB buffer (ml)	Ammonium persulphate (μl)	TEMED (μl)
12% Separating gel	7.25	6.25	4.50	-	100	20
4% Stacking gel	1.2	5.6	-	2.25	54	1

Note: Quantities given are sufficient for 4 gels

Ammonium persulfate was made up fresh before use and added immediately before gel casting

LWGB buffer (1.5 M Tris-HCL pH 8.0, 0.4% (w/v) SDS

UPGB buffer (0.5 M Tris-HCL pH 6.8, 0.4% (w/v) SDS

TEMED: N,N,N',N'-tetra-methyl-ethylenediamine

3.2.4.4 Protein loading, separation and visualization

Samples were mixed with 1X SDS-PAGE loading buffer (62.5 mM Tris-HCL pH6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) 2-mercaptoethanol, 0.0025% (w/v) bromophenol blue). Prior to loading, samples were heated in boiling water for 10 minutes, denaturing the proteins. Samples were centrifuged (prevents smearing) and immediately placed on ice. Equal amounts of protein extract (15 μl) were loaded into each well. A pre-stained protein molecular marker was loaded in the first lane with size ranging from 10-170 kDa (Fermentas, UK). Samples were run at 75 V until they reached the top of resolving gel, then run at 150 V until the pre-stained standard and samples reached the end of the gel. Running of the gel took place in the cold room to improve resolution.

Identical gels were run simultaneously; one was used as a protein gel, i.e. confirming that equal amounts of protein are loaded for each sample. The other gel was used for western blotting. Gels were removed from apparatus and placed in fixative solution (80% (v/v) methanol and 14% (v/v) glacial acetic acid) for 2-3 minutes then the fixative solution was returned to its original bottle. Coomassie Blue ® stain solution (12 ml Coomassie Blue ® G-250 (Biorad, USA)

and 3 ml of methanol) was added to the gel which was stained overnight on a rocking shaker. The gel was then de-stained in 30% methanol and 10% glacial acetic acid. An image of the gel was captured by digital camera.

3.2.4.5 Western blotting

The remaining SDS gel was immersed in blot transfer buffer. Six sheets of blotting paper and one piece of Immobilin-P membrane (Whatman®, PROTRAN BA 85, pore size 0.45 μM) were cut to the same size as the gel and dipped in blot transfer buffer for a few minutes. A sandwich that was made up of three pieces of blotting paper, the membrane, the gel and three pieces of blotting paper on top was placed over the anode plate of the blot transfer apparatus. Removing air bubbles was done by using glass test tube over the assembled sandwich which was covered with the cathode plate of the transfer apparatus. Proteins were transferred to the membrane using a Trans-Blot® SD semi-dry transfer cell (ATTO, Japan). The transfer was conducted at 15 V with a maximum current setting of 0.2 A per gel for 120 minutes. To confirm successful transfer of proteins from gel to membrane, the membrane was stained with 0.1% Ponceau-S stain in 5% acetic acid (Sigma-Aldrich, USA) for 10 minutes. Membranes were washed with Tris-buffered saline solution (TBS, 20 mM Tris-HCL, pH 7.3, 137 mM NaCl, 0.38% (v/v) 1 N HCL) then stored in TBS overnight. Next day, the membrane was blocked with 5% skimmed milk in 1X TBS for 1 hour. Membranes were incubated with primary antibody (Rubisco, PEPC or Rubisco activase) in 5% skimmed milk in 1x TBS at the concentration 1:3000 for 1 hour on rocking shaker. After incubation in the primary antibody, the membrane was washed twice with 1x TBST (0.1% (v/v) Tween 20 in TBS) for 10 minutes and then washed in TBS for 10 minutes. Secondary antibody (15 µl of goat anti-rabbit IgG; Sigma-Aldrich, USA in 15 ml skimmed milk solution) was added for one hour. Membrane was washed three times with TBST. Proteins were visualized by enhanced chemi-luminescence (ECL). The membrane was soaked for 30 seconds per side in 3 ml ECL1 and ECL2 reagents (GE Health Suppliers, UK) mixed immediately then wrapped in cling film and placed in a film cassette. The film (Kodak Biomax-XAR) was placed on the membrane under darkness in a film cassette for 30 seconds to 5 minutes. Film was developed using Kodak developer and fixer reagents.

3.2.4.6 Interrogation of transcriptomic and protein databases

An *A. americana* transcriptome database (see Appendix B) was interrogated by first obtaining the sequence of the *Arabidopsis* ortholog of the gene of interest (using NCBI) and then blasting this sequence against the *A. americana* transcriptome database using BioEdit. The abundance (RPKM) of the *A. americana* transcripts which showed the best matches (assessed via log e value) were then plotted against time to reveal day/night patterns of abundance. The *A. americana* transcript identifiers (i.e. Aam 356801) were then used to search the *A. americana* proteome database (see Appendix B), and protein abundance was also plotted against time over the day/night cycle.

3.2.5 Plant growth under contrasting water availability

Both *A. americana* and *A. attenuata* were exposed to two contrasting water regimes (70% & 20% F.C), for a period of 6 months. Leaf number was recorded every two weeks. Each treatment had four replicates. See section 2.2.1 for calculations of 20% and 70% F.C.

3.2.6 Statistical Analysis

All data presented are the mean values of three replicates. Values are expressed as means of three replicates ± standard error (S.E.) in each group. Where appropriate, data were analyzed using SPSS (IBM SPSS Statistics 21 64Bit) and graphs were produced using Microsoft Office Excel 2010. Normal distribution was tested using a normality test ($P > 0.005$) and significant differences between mean values were verified using a post hoc Least Significant Difference test (LSD) ($P < 0.05$) following one-way ANOVA.

3.3 Results

3.3.1.1 The effect of leaf succulence and leaf age on CAM expression

Two different *Agave* species (*A. attenuata, A. americana,*) varying in leaf succulence were compared over a 24 hour cycle. Titratable acidity measured the magnitude of CAM in both species and in different leaf ages (unfolded, young and mature).

The results showed a difference in titratable acidity between the beginning and end of photoperiod, indicating an overnight accumulation of acidity which is a

diagnostic feature of CAM. Data was expressed both on an area basis (mmol m^{-2}) Fig 3.3 (A& B), and fresh weight basis (μmol H^+ g^{-1} fwt;Fig 3.4 C&D).

The magnitude of CAM increased with leaf age from young to mature, and was significantly higher *(P=0.020)* in the most succulent species (*A. americana*) when expressed on an area basis. However, on a fresh weight basis, CAM was higher in mature leaves of A. *attenuata* compared to *A. americana* (*p=0.020*). The magnitude of CAM increased with leaf age in *A. attenuata (p=0.024)*, whereas, there was no significant difference in CAM activity with leaf age in *A. americana* when expressed on a fresh weight basis *(p=0.057)*.

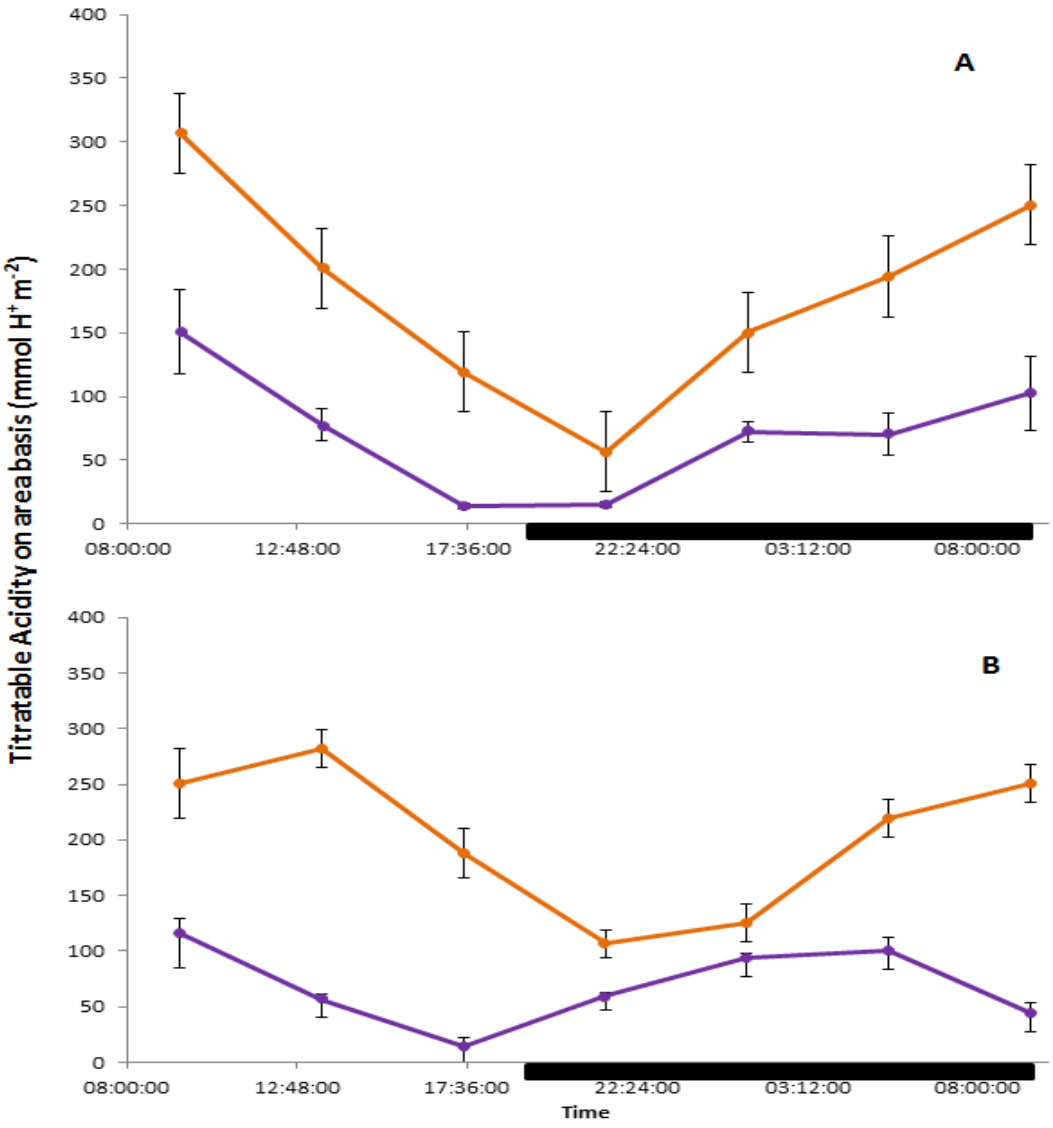


Figure 3.3 Time course kinetics over 24 hours for acid accumulation in leaves of *A. americana& A. attenuata* expressed on area basis (mmol H^+ m^{-2}). Fig 3.3 A. represents mature leaves of *A. americana* and *A. attenuata* and Fig 3.3 B is for young leaves. The black bar indicates the dark period. (n = 3 ± standard error).

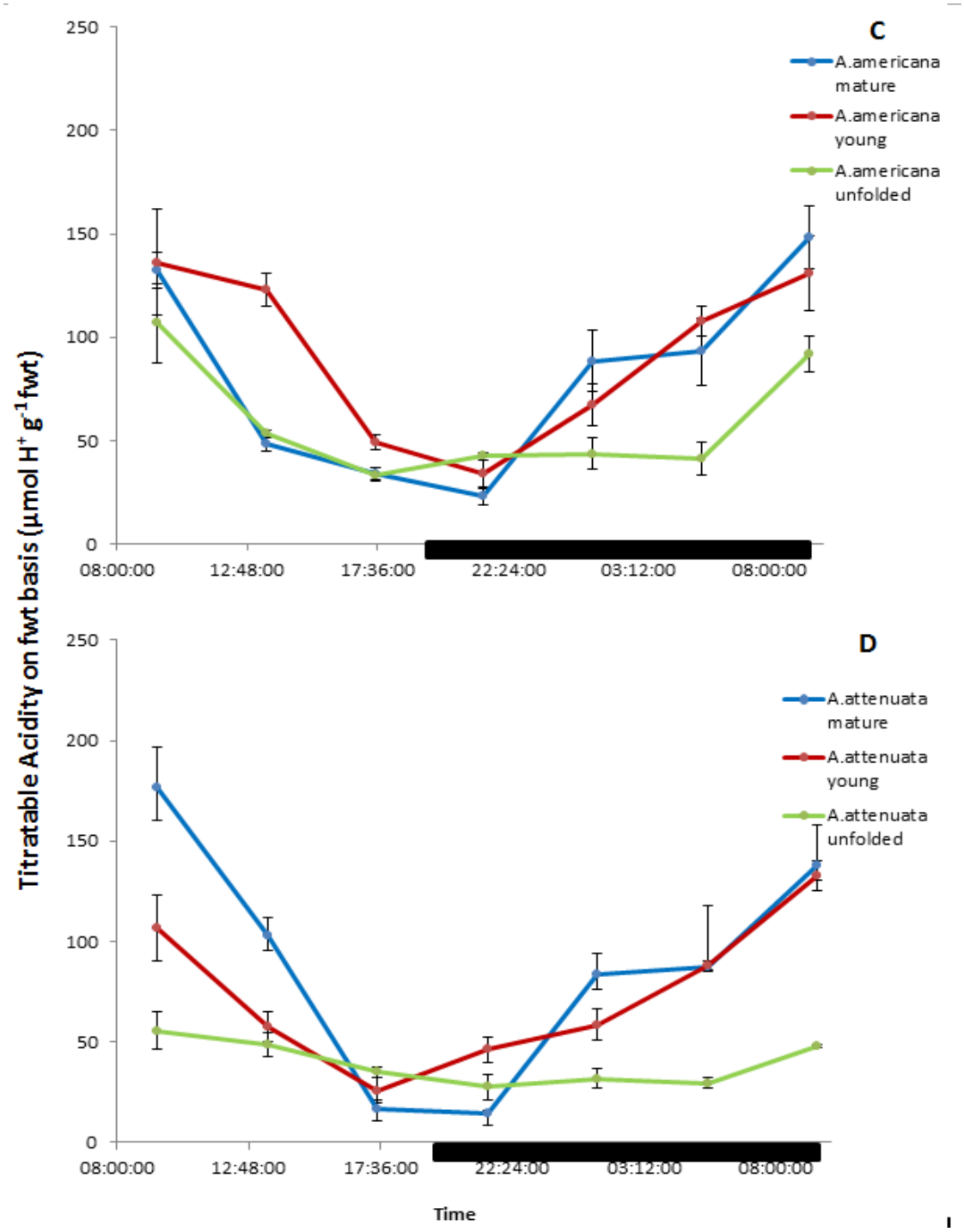


Figure 3. 4 Time course kinetics over 24 hours for acid accumulation in leaves of *A. americana & A. attenuata* expressed on fresh weight basis (µmol H^+ g^{-1} fwt) (Fig 3.4 C and D) Fig 3.4 C, represents *A. americana* (mature young unfolded) leaves. Fig 3.4 D are *A. attenuata*. Black bar on x-axis indicates dark period. (n = 3 ± standard error).

3.3.1.2 The effect of leaf succulence and leaf age on leaf soluble sugars

The most succulent species, *A. americana* had the highest amount of soluble sugars in both mature and young leaves on an area basis *(p=0.001 and p= 0.000)* respectively (Fig 3.5 A and B). On a fresh weight basis, *A. attenuata* contained more soluble sugars than *A. americana* and soluble sugars increased with leaf age in *A. attenuata,* significantly between mature and young leaves *(p=0.001)* and unfolded leaves, *(p=0.003,* Fig 2.6 D*)*. In contrast, there was no significant difference in soluble sugar content with leaf age in *A. americana* leaves (Fig 3.6C, (mature and young leaves *p=0.156*, Young and unfolded leaves *p=0.748*).

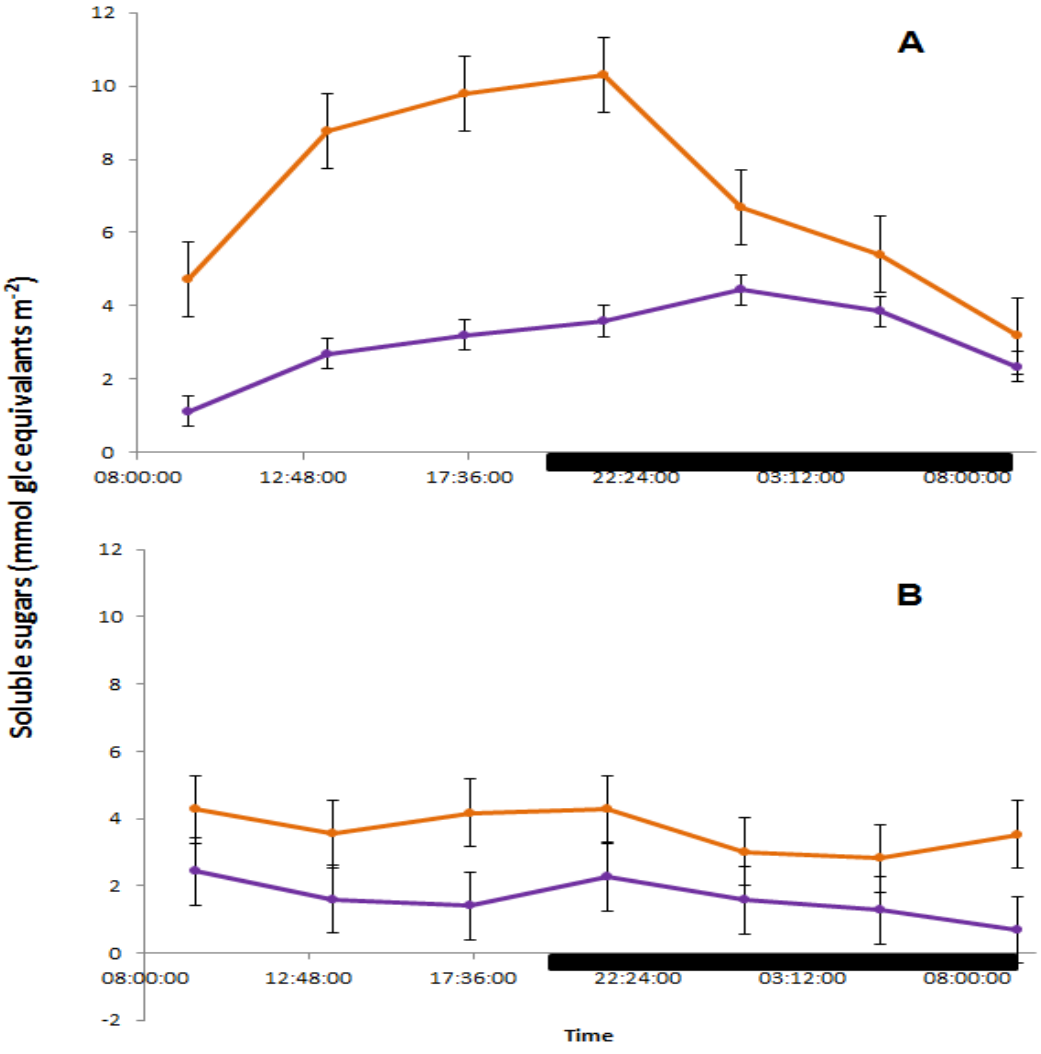


Figure 3.5 Time course kinetics for soluble sugar accumulation and depletion for *A. americana & A. attenuata* expressed on area basis (g m^{-2}) (Fig 3.5 A and B) Fig 3.5 A. Represents mature leaves of *A. americana* and *A. attenuata*

and Fig 3.5 B is for young leaves. Black bar on x-axis indicates dark period. (n = 3 ± standard errors).

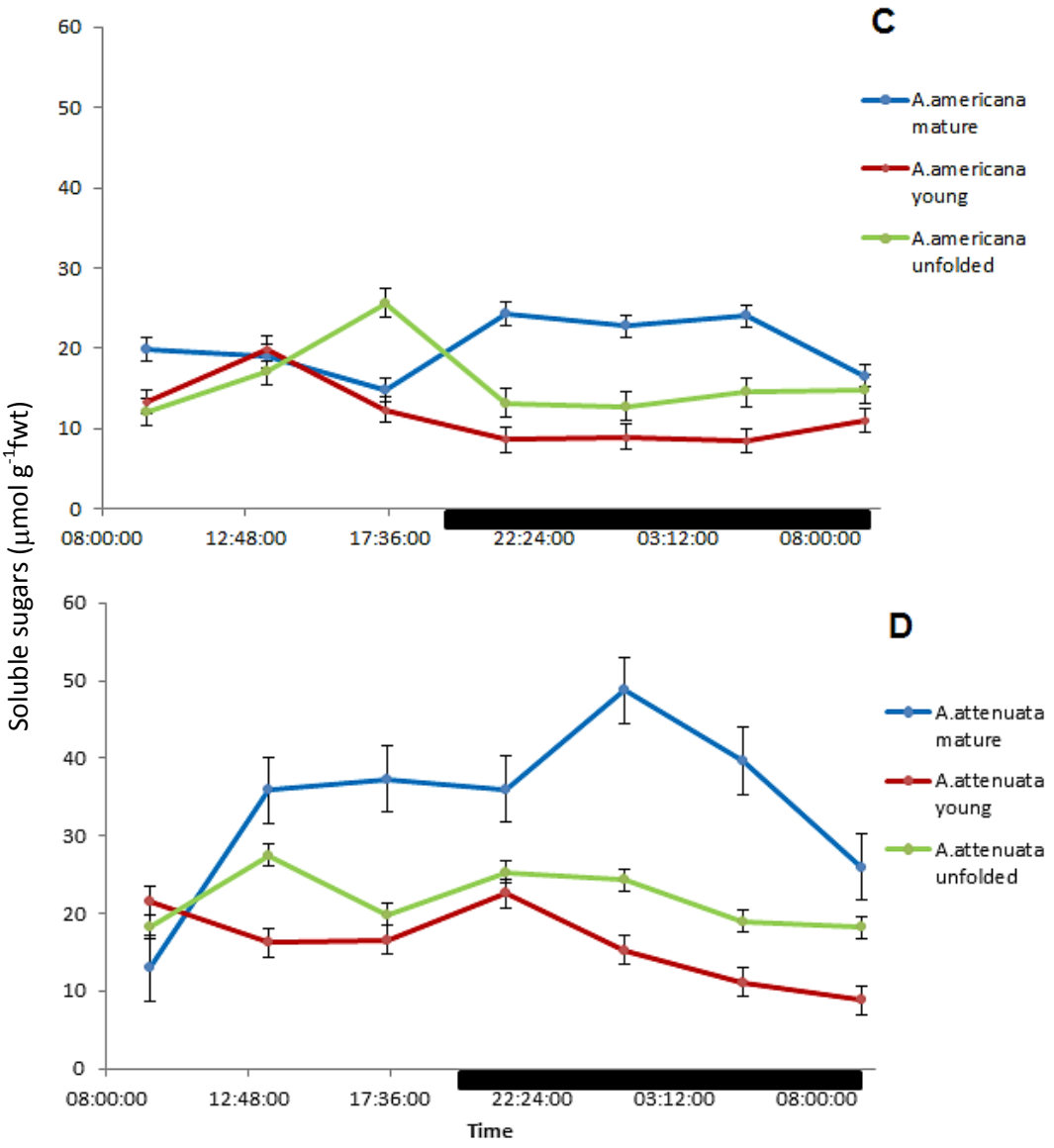


Figure 3.6 Time course kinetics for soluble sugar accumulation and depletion for *A.americana & A.attenuata* expressed on fresh weight basis (µg g^{-1} fwt) .Fig 3.6 C, Samples collected from *A.americana* (mature young unfolded leaves. Fig 3.6 D, are for *A.attenuata.* Black bar on x-axis indicates dark period. (n = 3 ± standard errors for error bars indicated).

3.3.2.1The effect of leaf succulence and leaf age on abundance of PEPC, Rubisco and Rubisco activase

The impact of leaf succulence and age on protein abundance of the key photosynthetic enzymes PEPC, Rubisco and the Rubisco activase was investigated using Western blotting (Figure 3.7).

In general, the abundance of PEPC protein was higher in leaves of *A. americana* compared to *A. attenuata* (Fig. 3.7). In contrast Rubisco protein abundance was higher in leaves of *A. attenuata*. Rubisco activase abundance was comparable in the two *Agave* species. In terms of leaf age, the abundance of PEPC was the highest in mature leaves of both species of *Agave*, complimenting titratable acidity findings (Fig 3.7, Lanes 1&4). Both Rubisco and Rubisco activase were abundant in mature (Lane 1&4) and young (Lane 2&5) leaves of both species, but Rubisco activase protein was below the limits of detection in unfolded (Lanes 3&6) leaves of either species.

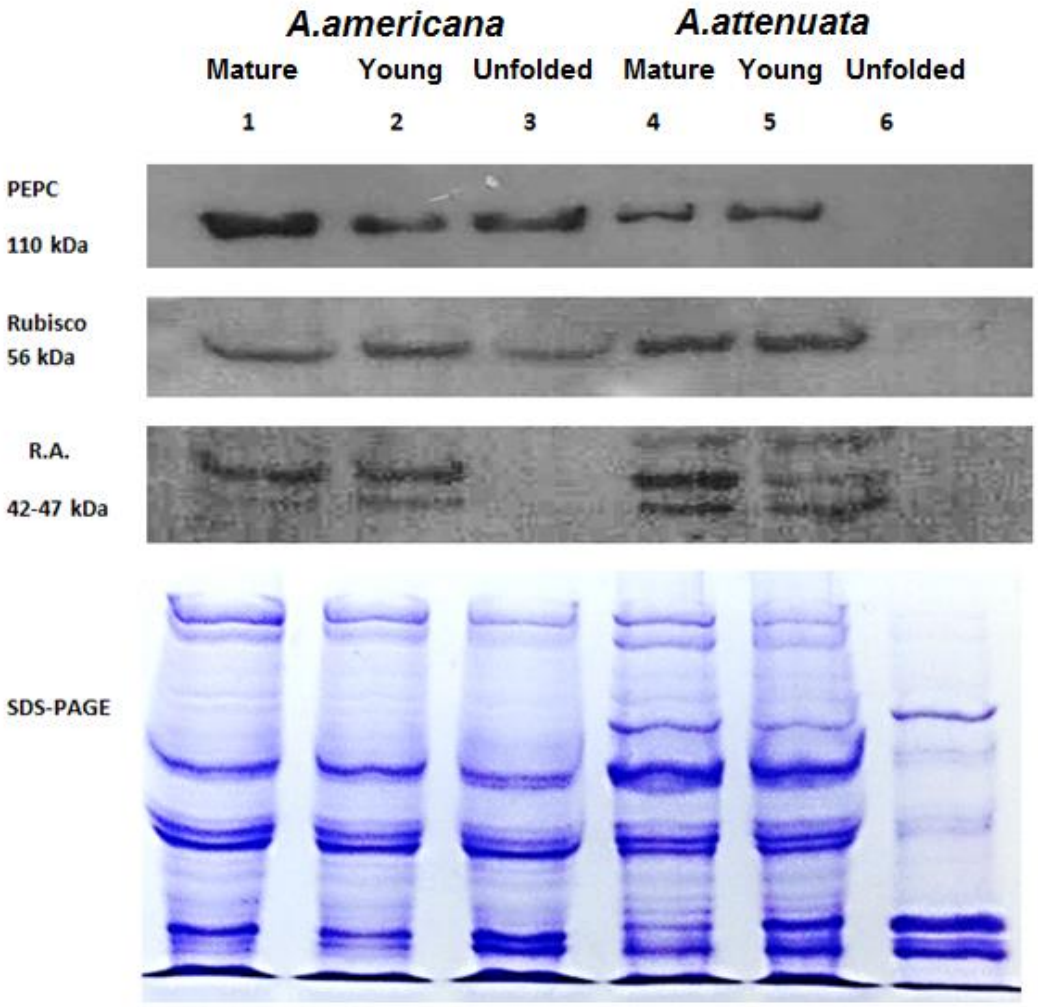


Figure 3.7 Western blots showing the relative abundance of PEPC, Rubisco and Rubisco activase (R.A) proteins in different leaf ages from Mature (Lane 1&4), Young (Lane 2&5) and Unfolded (Lane 3&6) leaves of *A. americana (Lanes 1,2,3) and A. attenuata (Lanes 4,5,6).*Additional SDS-PAGE gel shows the loading of protein.

3.3.2.2 The effect of leaf position and watering regimes on PEPC and Rubisco abundances.

The impact of different watering regimes (20% & 70% F.C) and leaf position (tip vs. base) on protein abundance of PEPC and Rubisco was investigated using Western blotting (Figure 3.8).

For both *Agave* species, Rubisco protein abundance was intensified in the tip portion of the leaf under both watering regimes (Fig.8). The picture for PEPC abundance in leaf tip versus leaf base however was less clear. For *A. americana*, there was more PEPC in the tip compared to the base under

droughted conditions, but under watered conditions (70% FC) this pattern was reversed with more PEPC in the leaf base. *A. attenuata* showed a different response with more PEPC in the leaf base under drought conditions (20% FC) but more PEPC in the tip under watered conditions. Thus, there was no close association with the magnitude of CAM (Chapter 2) in leaf tip and leaf base and PEPC abundance.

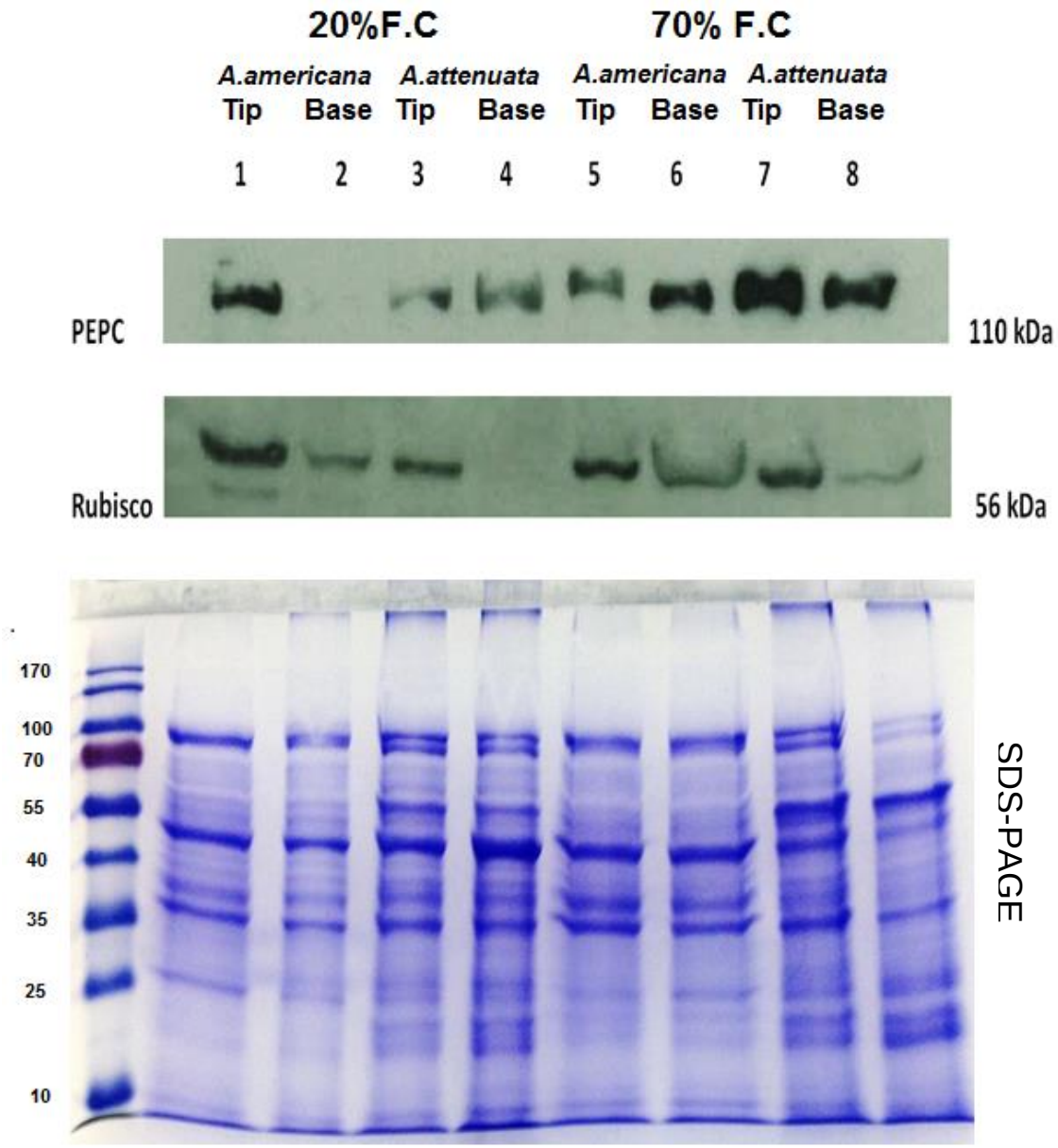


Figure 3.8 Western blots showing the relative abundances of Rubisco and PEPC proteins in leaf tissue of *A. americana and A. atteunata* under two water regimes (20% and 70% F.C). Lanes 1&2 are *A. americana* tip and base of leaf respectively under 20% F.C. Lanes 3&4 are *A. attenuata* tip and base of leaf under 20% F.C. lanes 5 & 6 are *A. americana* tip then Base of leaf under 70% F.C (i.e. well-watered conditions). Lanes 7&8 are *A. attenuata* under well watered conditions (70% F.C.). Additional SDS-PAGE gel shows loading of protein.

3.3.2.3 Diel time course of Rubisco activase abundance in Agave species varying in succulence under different water regimes

The impact of different water regimes (20% & 70% F.C) on the diel protein abundance of Rubisco activase was investigated in mature leaves of *A. americana* and *A. attenuata* using Western blotting over a 24 h period (Figure 3.9) & (Figure 3.10).

In the most succulent species, *A. americana*, Rubisco activase abundance was highest at night under well watered conditions (70% F.C). For droughted plants of *A. americana*, the overall abundance of Rubisco activase increased, with highest abundance observed at the end of the day, through the night and the start of the day. Lowest abundance was observed in the middle of the day under both watering regimes (Fig 3.9).

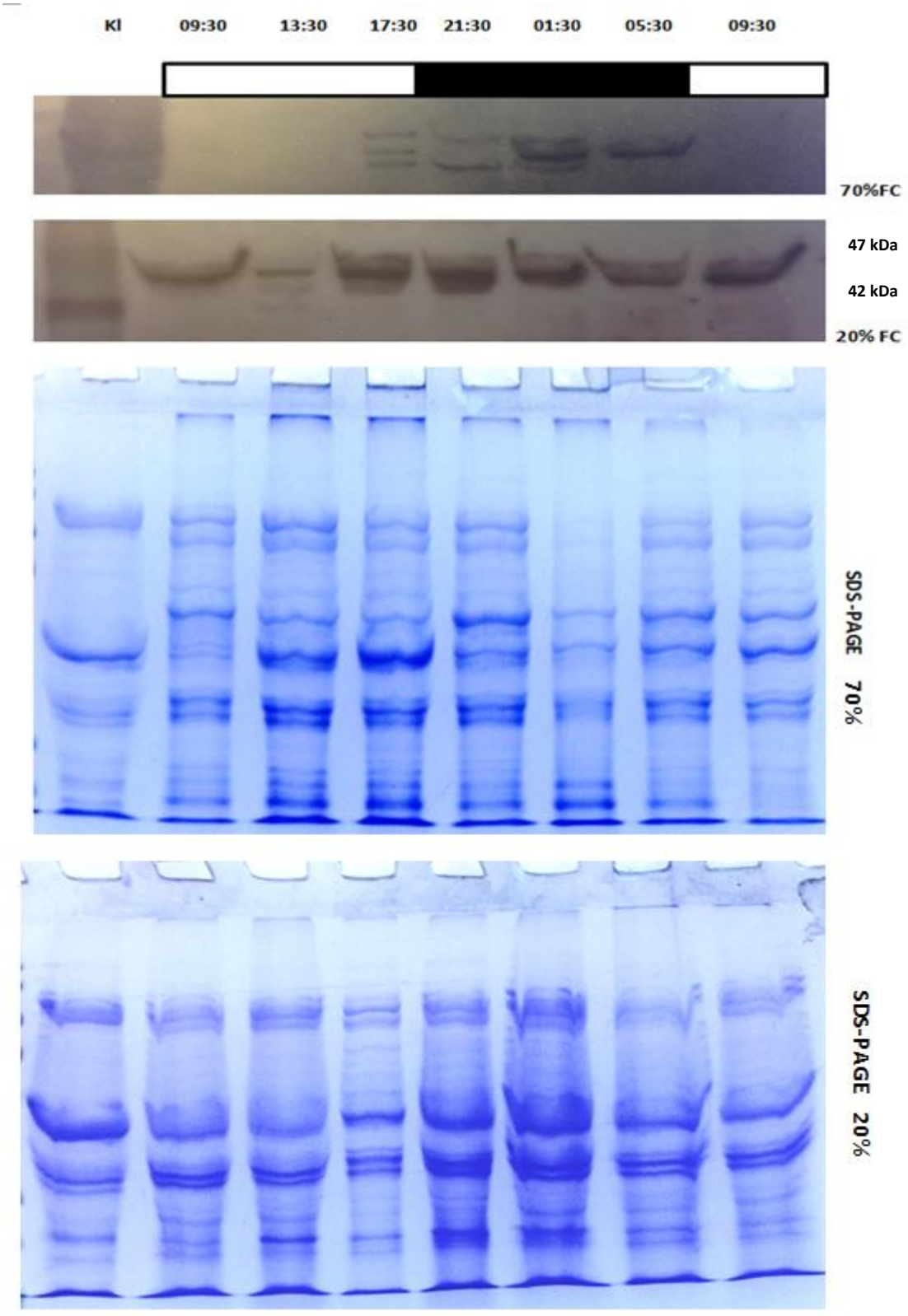


Figure 3.9 Western blots showing Rubisco activase abundance over a diel CAM cycle (24 h) of tips of *A. americana* under well watered (70% F.C) and drought conditions (20% F.C). Black bar indicates dark period (Phase I). *Kalanchoe* (KL) was used as a control. Additional SDS-PAGE gels show loading of protein for 70% & 20% F.C.

For the less succulent species, A. *attenuata*, the diel pattern of Rubisco abundance was less marked compared to that observed in A. *Americana*. Multiple bands were more obvious in this species, suggesting the existence of different isoforms of Rubisco activase. In contrast to A. *Americana*, drought led to a general decrease in the abundance of Rubisco activase in *A. attenuata* (Fig 3.10).

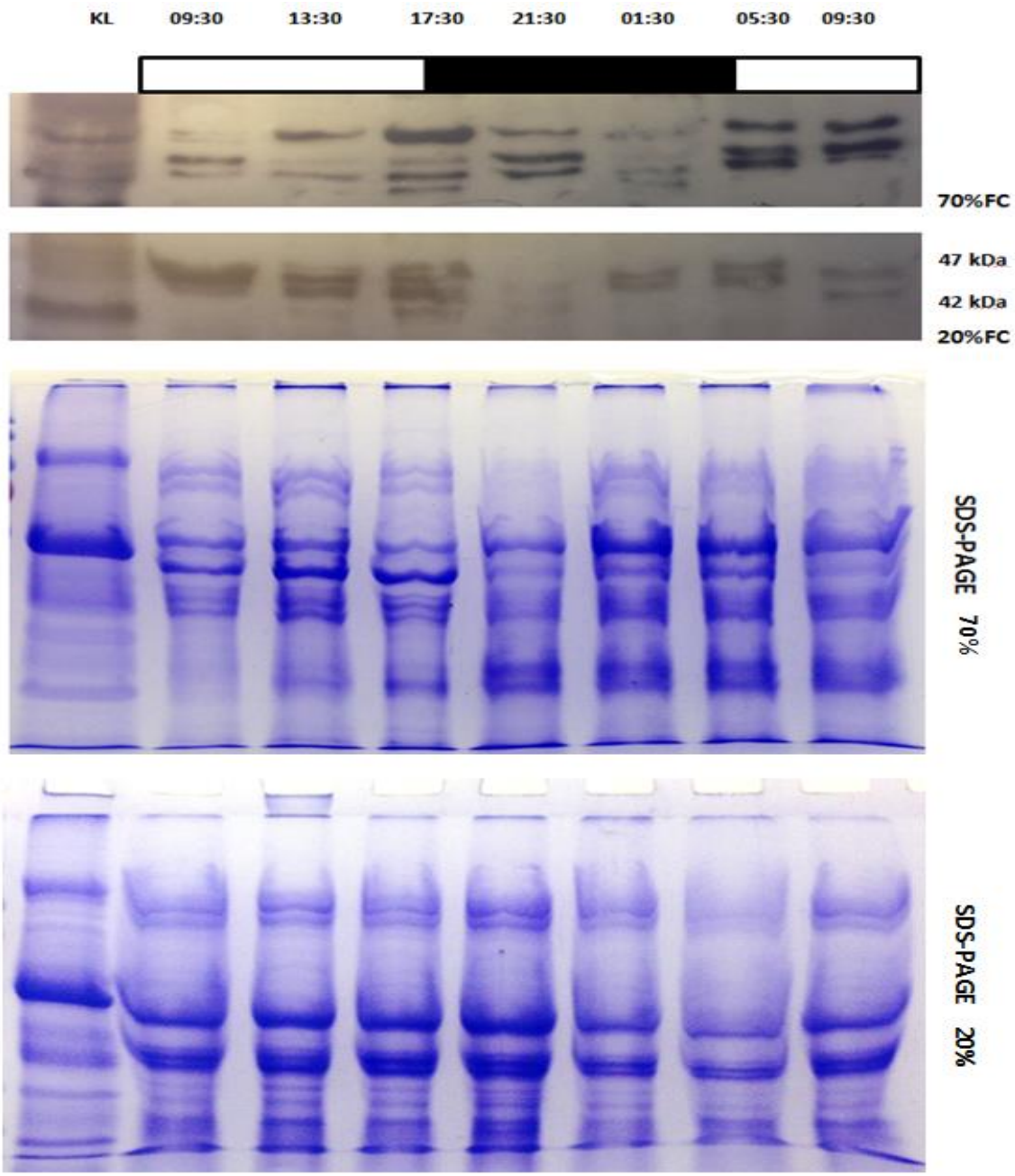


Figure 3. 10 Western blots showing Rubisco activase abundance over a diurnal CAM cycle (24 h) of tips of *A. attenuata* under well watered (70% F.C) and drought conditions (20% F.C). Black bar indicates dark period (Phase I). *Kalenechoe* was used as a control. Additional SDS-PAGE gel shows equal loading of protein gels for both watering regimes (70% &20 F.C)

3.3.4 Interrogation of transcriptome and protein databases related to PEPC and Rubisco activase in *A. americana*

The western blotting data described above for *A. americana* were compared with a transcript and protein database for *A. americana* (Biosciences Research group at the Oakridge National Laboratory in Tennessee). These data bases contain information relating to transcript and protein abundances from mature leaves of *A. americana marginata* sampled at 4 hour intervals over a 24 light/dark cycle. The transcript data base contains information pertaining to global transcript abundances in young, C3 leaves and other plant tissues such as meristem, stem, root and rhizome. Data mining of the transcript and protein data bases was conducted to illustrate transcript and protein abundance for PEPC (Fig 3.11) and Rubisco activase (Fig 3.12) over a 24 h time course. Some 11 transcript sequences were found to correspond to PEPC. Transcript sequence (*Aam080248*) showed the highest abundance in mature leaves and peaked at 6pm, 12am and 3pm in the diel cycle. The transcript also peaked at 6pm and 12 am in the young C3 leaves and in meristem tissue. Sequence (*Aam080248*) also had the highest protein abundance in mature leaves and peaked at 9am. Transcript abundance of (*Aam080248*) in roots, rhizome and stems was much lower than that in young leaves and particularly mature leaves. Thus, this protein may well have a CAM-specific function.

For Rubisco activase, some 10 transcript sequences were found to correspond to Rubisco activase. Transcript sequence (*Aam041100*) showed the highest abundance in mature leaves and peaked at 6am in the diel cycle. The transcript also peaked at 6pm in young C3 leaves and meristem tissue. Sequence (*Aam041100*) had the highest protein abundance in mature leaves with the highest peak at 3am in the diel cycle. Transcript abundance of (*Aam041100*) was lower in roots, rhizome and stem tissue and was much lower than that in young leaves and mature leaves, as would be expected for a protein involved in photosynthetic metabolism.

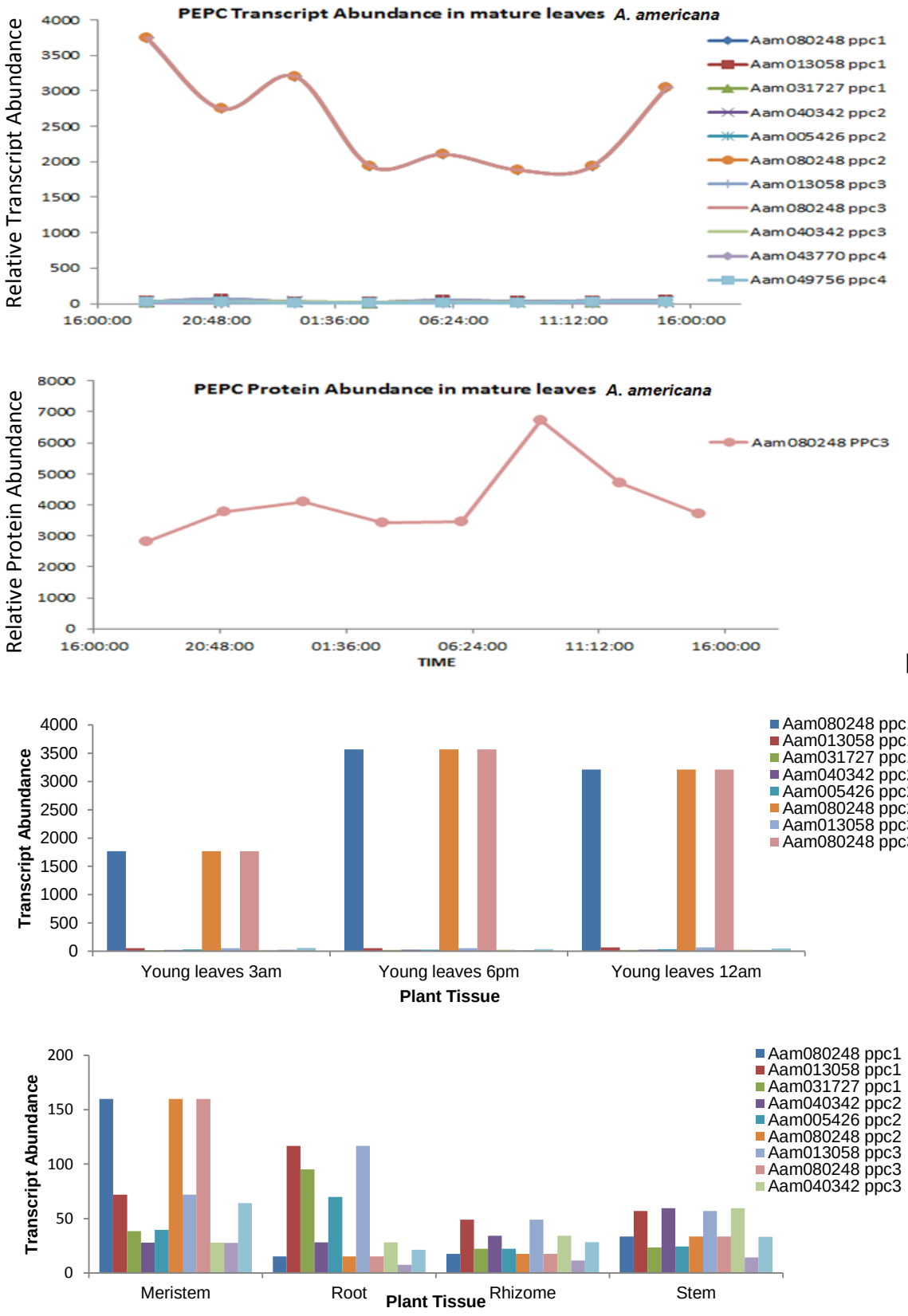


Figure 3.11 Time course kinetics of transcript and protein abundances of PEPC in mature leaves of *A. americana.* The most abundant transcript was Aam 080248. Also shown are transcript abundances for different tissues and C3 young leaves at 3 time points.

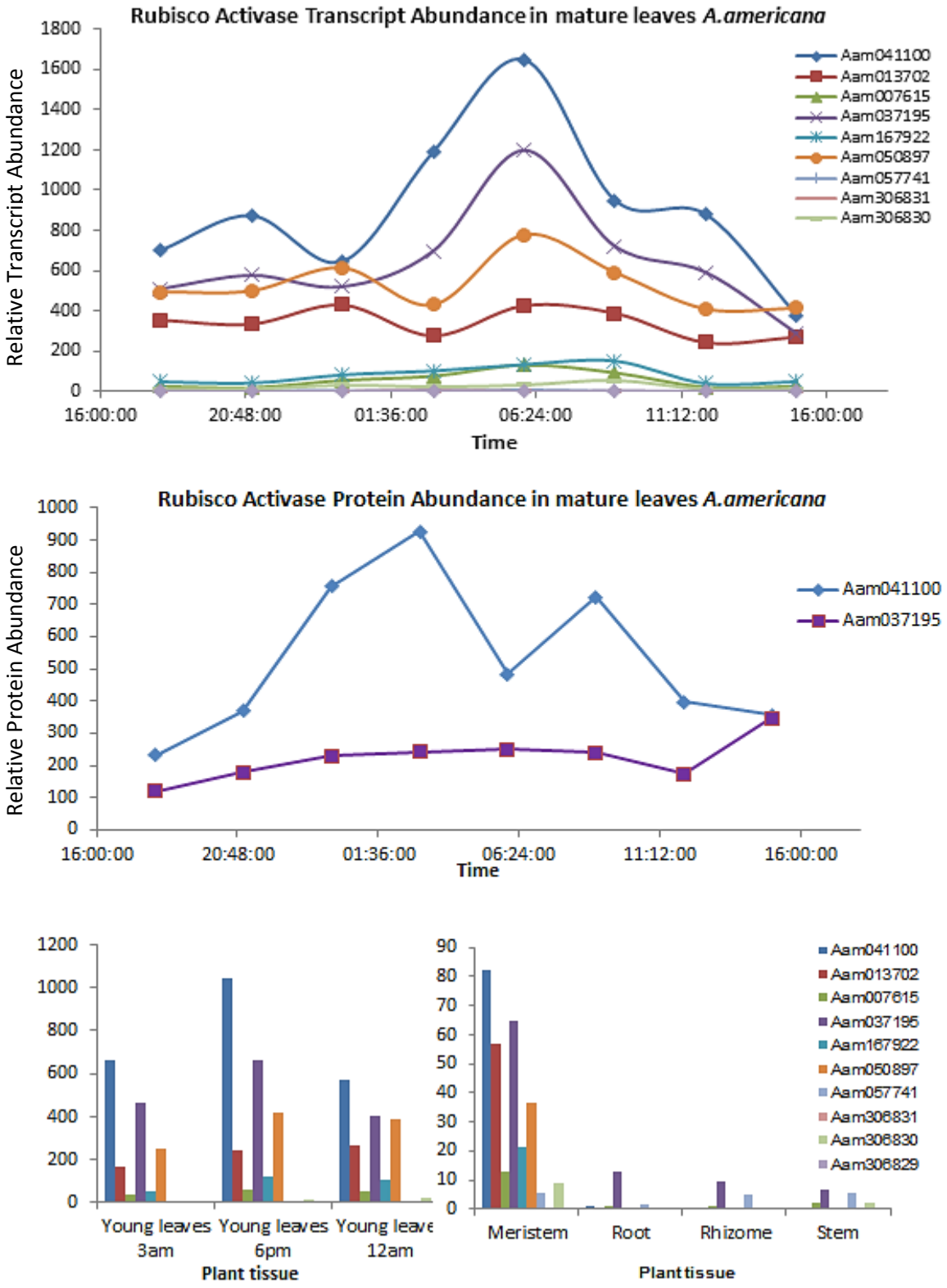


Figure 3.12 Time course kinetics of transcript and protein abundances for Rubisco Activase in mature, young and different tissues of *A. americana.* In mature leaves the transcript sequence Aam 041100 showed the highest abundance. The same sequence showed the highest protein abundance in mature leaves. Also shown are transcript abundances of different Rubisco activase sequences in different tissues and C3 young leaves at 3 time points.

3.3.5 Plant growth under contrasting water availability

Plant growth (as indicated by the number of expanded leaves) in *A. americana* occurred at similar rates under the contrasting water regimes ($p= 0.001$) and was not affected by drought conditions $p=0.766$ (Fig 3.13A). Droughted plants had fewer leaves than watered at the start of the monitoring period since these plants had previously been droughted before starting to monitor growth. This was due to shortage of plant availability. After 12 weeks, ~ 3 new leaves had been produced in *A. americana* under each watering regime.

Drought had a significant effect on the growth of *A. attenuata* ($p= 0.005$) (Fig 3.13B). Again, the droughted plants started off with fewer leaves than well watered since they had been previously droughted. After 12 weeks, ~ 3 new leaves had been produced in the watered (70 % FC) plants of *A. attenuata* and ~ 2 new leaves produced in the droughted (20 % FC) plants.

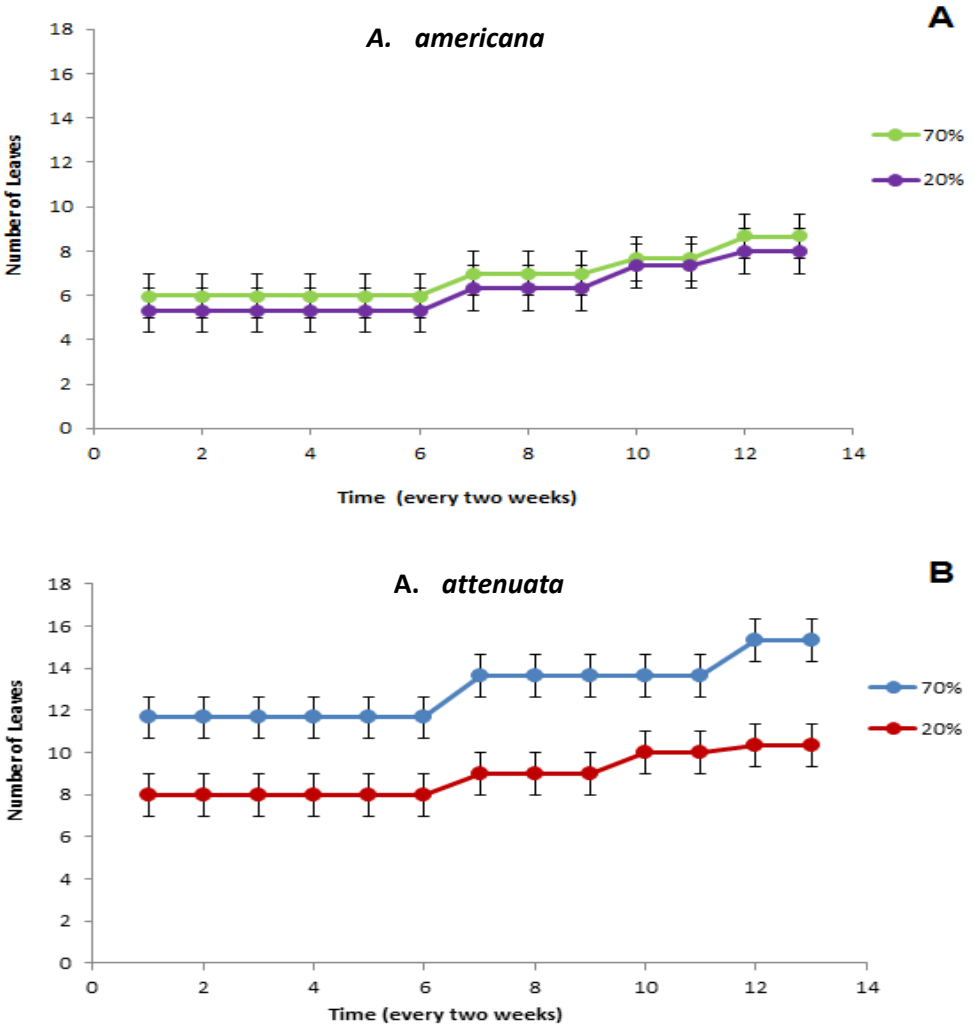


Figure 3.13 Plant growths for both *Agave* varying in succulence over a 6 month period, under contrasting water regimes. Fig 3.13 A represents *A. americana* growing under 70% F.C. and 20% F.C. Fig 2.13B is for *A .attenuata* indicates growth under well watered conditions (70% F.C) and growing under 20% F.C

3.4 Discussion

The aim of the study was to test several hypotheses related to succulence and the biochemical capacity for C_3 and C_4 carboxylation in two species of *Agave*.

3.4.1 Effect of leaf succulence, leaf age and leaf position on CAM activity and PEPC abundance

Two different *Agave* species (*A. attenuata, A. americana,*) varying in leaf succulence were compared. In both species, the magnitude of CAM increased with leaf age from young to mature, and was highest in the most succulent species *A. americana* when expressed on a leaf area basis. That the older and more succulent leaves of *Agave* are more committed to CAM compared to the younger, thinner leaves was similar to findings for the CAM dicot *Kalanchoë* reported by Griffiths et al (2008). In the CAM monocot *Fourcroya humboldiana*, nocturnal changes in titratable acidity were also dependent on leaf age with this parameter increasing from the younger to more mature leaves (counting from rosette centre). Previous studies on *Agave tequilana* (Pimienta-Barrios *et al.*, 2001; Pimienta-Barrios *et al.*, 2006) showed that the magnitude of daily C gain and plasticity in deployment of C_3 and C_4 carboxylation was dependent on plant age. That study also found that maximum rates of instantaneous net CO_2 uptake in mature plants were 40% higher than those in young *A. tequilana*.

When CAM activity in *Agave* was expressed on a leaf fresh weight basis, the least succulent species, *A. attenuata* showed higher CAM than the more succulent *A. americana*. This finding illustrates the importance of the units used to express CAM activity. If CAM activity is expressed as the amount of nocturnal CO_2 uptake, this is usually expressed on a leaf area basis, thus we see a direct relationship between succulence and CAM. However, given the increased density (weight) of the more succulent leaves, when CAM activity is expressed as acid accumulated on a fresh weight basis, the positive relationship between succulence and CAM is lost. A similar trend was noted for soluble sugar content. The most succulent species, *A. americana* had the highest soluble sugar content for both mature and young leaves when this was expressed on an area basis ($p=0.001$ and $p= 0.000$) respectively. However, an a fresh weight basis, soluble sugar content was highest in *A. attenuata,* and increased significantly between mature and young leaves ($p=0.001$) and unfolded leaves, ($p=0.003$,

Fig.2.C) in this species. This data reinforces the importance of the units used to express CAM activity and sugar content in *Agave*.

As predicted, the abundance of PEPC protein was the highest in mature leaves of both species of *Agave*, complimenting titratable acidity findings. Furthermore, on a protein basis, the more succulent species *A. americana* showed a greater investment in PEPC protein compared to the thinner leafed *A. attenuata*. Succulence in *Agave* provides a high vacuolar storage capacity for malic acid. This potential for high CAM activity is complimented by increased investment of leaf protein into PEPC in the more succulent *Agave* species. Increased succulence and PEPC protein abundance also offers the potential to extend C4 carboxylation for several hours at the start of the day (Osmond *et al.*, 1999; Borland *et al.*, 2000). Extending Phase II is beneficial for carbon gain by delaying the onset of Phase III decarboxylation until the warmest, brightest time of day (Borland *et al.*, 1996). That succulence can also buffer against water deficits and maintain growth under water limited conditions was also supported by the growth data collected for the two *Agave* species under watered and droughted conditions. The more succulent *A. americana* produced more leaves under drought compared to the thinner leafed *A. attenuata*. Thus, the more succulent *Agave* species has the potential to outperform the less succulent species under field conditions.

Previous studies of the rosette forming CAM species *F. humboldiana* showed that CAM activity varied with distance from the base of the leaf (Olivares and Medina, 1990). The rosette leaf arrangement creates a variable light intensity environment resulting from inclination of leaf angle, which decreases with age and increasing its exposure to light (Olivares and Medina, 1990). Within a leaf, the formation of a longitudinal light gradient occurs, the base portion is shaded by upper leaf blades and the tip receives more light. Thus, a net acidification gradient that occurs from the leaf base to the tip might be predicted. Popp *et al*. (2003) showed an increase of organic acid (malate, citrate) concentrations from the basal portion to the tip of leaves of *Ananas comosus*. Similar results were reported for *Agave* in Chapter 2 of this thesis. However, there was no clear relationship between the magnitude of CAM along the leaf and PEPC abundance in the leaf tip and base in either of the two *Agave* species investigated here. This finding suggests that the increasing gradient of

CAM activity from base to leaf tip (as shown in Chapter 2) was regulated by something other than C4 carboxylase activity. Given that light intensity is generally higher at the tip than the leaf base, as described above, the higher CAM activity at the leaf base may have been influenced by the abundance of sugars that are used as substrates for nocturnal carboxylation (see also Chapter 2).

3.4.2 Effect of leaf succulence, leaf portion and leaf age on Rubisco and Rubisco activase abundances

In contrast to PEPC, Rubisco protein was intensified in the leaf tip of both species, indicating that light intensity regulates Rubsico abundance but not PEPC abundance in *Agave*. It was postulated that diffusional resistance to CO_2 in thick leafed CAM plants like *Agave*, might be compensated for by an increased investment in Rubsico protein. This could enhance photosynthetic carbon gain, overcoming anatomical constraints imposed by low intercellular air space (IAS) to CO_2 diffusion (Maxwell *et al.*, 1997); (Nelson *et al.*, 2005). However, this hypothesis was not supported by the data presented here. The leaf tips which were the thinnest part of the leaf had the greatest Rubisco abundance. Moreover, the thinner leafed *A. attenuata* invested more of its leaf protein into Rubisco compared to the more succulent *A. americana*. Both Rubisco and Rubisco activase were abundant in mature and young leaves of both species, but Rubisco activase was below the limits of detection in unfolded leaves of either species. Unfolded leaves have lower chlorophyll content (data not shown) and are probably photosynthetically limited compared to the expanded leaves which may have influenced Rubisco activase content. Rubisco activase is required to promote and maintain the catalytic activity of Rubisco within the leaf and could be important for overcoming diffusion limitations of CO_2 across the leaf (Griffiths *et al.*, 2008), thereby optimising CO_2 draw-down and uptake. However, given that there was no clear difference between the two *Agave* species in overall abundance of Rubisco activase the hypothesis that diffusional resistance to CO_2 in more succulent leaves might be compensated for by having more Rubisco activase was not supported.

In high yielding CAM species such as *Agave*, it is proposed that atmospheric CO_2 fixed directly by Rubisco in Phase IV, contributes a substantial proportion of C skeletons required for growth (Bartholomew and Kadzimin, 1977; Winter,

1985; Borland *et al.*, 1994) This aids in selecting appropriate *Agave* cultivars which are appropriate for marginal lands with contrasting rainfall patterns and fluctuating temperatures (Borland *et al*., 2011).

3.4.3 Impact of succulence and contrasting water regimes on Rubisco activase abundance over a diel cycle

The diel (i.e. 24 h) abundance of Rubisco activase was compared for two species of *Agave* contrasting in succulence and under different water regimes. Previous studies have shown that the regulation of Rubisco activation may be modified by environmental conditions including drought stress (Griffiths *et al.*, 2008). It was hypothesised that the abundance of Rubisco activase will vary over the diel cycle, particularly in leaves of the more succulent *A. americana*. The idea was that as internal [CO_2] declines towards the end of phase III, Rubisco will face diffusional limitation of CO_2 thus Rubisco activase abundance will increase to enhance the activation of Rubisco (Maxwell et al, 1999; Griffiths *et al*, 2008). A clear diel pattern of Rubisco activase abundance was noted for A. *americana*, particularly under well watered conditions. The lowest abundance of Rubisco activase was noted during the middle of the day, which is consistent with the idea of compensating for diffusional resistance to CO_2 (Griffiths et al, 2008). Studies on C3 plants have shown that increasing levels of CO_2 within the leaf tend to down-regulate the effectiveness of Rubisco activase (Cockburn W, 1979); (Spalding MH, 1979). Since internal [CO2] in a strong CAM species like A. *americana* will be highest in the middle of the day (phase III), this could explain the lower abundance of Rubsico activase in the middle of the day. Also, interactions with high temperatures at midday tend to reduce the effectiveness of Rubisco activase in some C3 plants (Crafts-Brandner and Salvucci, 2000)

The diel change in protein abundance of Rubisco activase in *A. americana* reported in this thesis was supported by independent studies of the *A. americana* proteome (Plant Systems Biology Group, Oak Ridge National Lab) which also indicated a peak in protein abundance at night. Transcript abundance for Rubisco activase in *A. americana* however peaked at the start of the day so there was no clear correlation between transcript and protein abundances. Such findings might indicate that Rubisco activase is not just regulated at the level of transcription but is subject to additional layers of control. Alternative splicing of Rubisco activase has for example been reported for some

C3 plants (Zhang and Portis, 1999). Alternative splicing could give rise to more than one isoform of Rubisco activase. Several bands were noted for this protein in the western blots, particularly for *A. attenuata*. Overall, the data indicate that regulation of Rubisco activase abundance differed between the two *Agave* species. The physiological significance of this is unclear but could be related to differences in leaf succulence and the relative magnitudes of C3 and C4 carboxylation in the two species.

3.5 Conclusions

Results presented in this chapter, confirmed that the expression of CAM is dependent on leaf succulence and leaf age. Succulence also influenced the abundance of PEPC. Thus, the optimal anatomy for nocturnal malic acid accumulation is accompanied by high PEPC abundance in leaves with higher vacuolar storage capacity. In contrast, the abundances of Rubisco and Rubisco activase showed an inverse relationship to succulence and CAM activity. Thus, in the less succulent *Agave* species which fixes a greater proportion of CO_2 during the day, investment in the C3 carboxylating system was enhanced compared to the more succulent, strong CAM species. Differences between species in the regulation/activation of Rubisco were also apparent. Ultimately, a systems level of understanding the metabolic pathway of CAM will be required for exploiting and maximizing the potential yield of CAM species for biofuel production in marginal ecosystems.

Chapter 4

Inter-specific variation across *Agave* in traits associated with the operation of CAM and fructan accumulation

4.1 Introduction

Agave is a succulent genus of some 200-300 species within the monocot family Agavaceae (Davis *et al.*, 2011a; Escamilla-Treviño, 2012), which inhabit and thrive in arid and semi-arid lands. *Agaves* are perennial xerophytes, with sizes ranging from several centimetres up to 4m in height and with large flowering stalks that range from 2m up to 12m that appear after 5 to 15 years of growth (Valenzuela-Zapata, 1985; Gentry, 2004). The leaves are arranged in a rosette often with a terminal spine and sometimes with spiny margins. The mesophyll contains elongated water storage cells, and stomata are sunken at the base of hypostomatal cavities (Blunden *et al.*, 1973). Analyses on almost all species of the genus *Agave* has shown the presence of crassulacean acid metabolism (CAM) as a carbon concentrating mechanism, and it is assumed that the genus as a whole uses CAM for the majority of net CO_2 uptake (Davis *et al.*, 2011a). The most common commercial uses for *Agave* are for fibres and beverages. The Food and Agriculture Organization (FAO) of the United Nations (FAO, 2010) has estimated that over I Mha of land is used for the cultivation of *Agave* for sisal fibres. In the 1990's,Mexico cultivated 70,000 ha of *Agave tequilana* for the production of alcoholic beverages and 20,000 ha of *A. fourcryodes* for fibre production (Nobel, 1994).The predominant *Agave* species grown for fibre in Brazil and Eastern Africa is *A. sisalana* (Davis *et al.*, 2011a).

Drought is one of the prime abiotic stresses limiting crop production. *Agave* are known to be well adapted and grow naturally in dry, arid conditions, and only require 20% of water for cultivation, when compared to calculated values of crop water demand for the most water efficient C3 and C4 crops (Borland *et al.*, 2009). Optimum growth can be achieved with annual rainfall from 102-127 cm and relatively high production of some *Agave* species has been found in regions with only 25-38 cm of annual rainfall (Kirby, 1963). In order for *Agave* to survive in regions with frequent drought, they must be efficient in their use of water and capable of surviving between rainfall events. *Agave*s are able to achieve this due to the operation of CAM as well as a number of other attributes. These attributes include hydraulic isolation (Davis *et al.*, 2011a) where roots shrink to prevent dehydration, thick cuticles and closed sunken stomata which prevent water loss to the atmosphere and maintain high plant water potential,

which also limits cavitation of roots during prolonged droughts (Linton and Nobel, 1999). Such features make *Agave* good candidates for exploitation on marginal or uncultivated land for bioenergy.

Agave plants have high cellulose and sugar contents, along with high biomass yield. More importantly, the operation of CAM in Agave confers high water-use efficiency. Leaf succulence is one of the key morphological correlates of the capacity for CAM (Winter *et al.*, 1983; Borland *et al.*, 1998; Griffiths *et al.*, 2008). Previous findings conducted on *Kalanchoe* (Crassulaceae), found that succulence is positively correlated with the contribution from CAM activity to total carbon gain (Kluge *et al.*, 1993; Kluge *et al.*, 2001) Large cell size and succulence are pre-requisites for CAM photosynthesis (Griffiths, 1989; Borland *et al.*, 2000), due to their large vacuoles that are important in overnight malic acid storage and which also act as water reservoirs (Osmond et al, 1999; Borland et al., 2000). Data presented in Chapter 2 also showed a positive relationship between succulence and CAM in 3 species of *Agave*.

Agave species are reported to be hexose utilizing CAM plants (Black et al, 1996), balancing acidity with water soluble hexoses, and for nocturnal PEP synthesis and so the vacuole has an additional role as a reservoir for storage carbohydrates to support the diel turnover of organic acids. A study on *Ananas comosus* (Borland and Griffiths, 1989) displayed the osmotic implication of using soluble sugars in the vacuole as sources for PEP. Close stoichiometry between organic acid accumulation and osmotic pressure ($\Delta\pi$) was observed in *A. comosus* with a balance between hexose depletion and malate and citrate accumulation. In the CAM species *Fourcroya humboldtiana*, the relatively high osmotic pressures are probably the result of the accumulation of osmotically active soluble carbohydrates such as fructans (Olivares and Medina, 1990)

Another typical feature of *Agave* is the production of fructans, which are polymers of B-fructofruranosyl residues synthesized from sucrose (Valluru and Van den Ende, 2008). The main function of fructans is storage of excess fixed carbon (Lewis, 1984) and fructans are accumulated in vacuoles of succulent parenchyma cells of leaf bases and stems (pina). Fructans are easily degradable by thermal or enzymatic treatments to yield the ethanol for tequila production (Narváez-Zapata and Sánchez-Teyer, 2010). *Agave* leaves are

usually discarded back in the field after pina harvest but could be employed for biofuel production (Simpson *et al.*, 2011a). Fructans are the major source of ethanol and are important vacuolar sinks for photoassimilate in mature leaves of *Agave deserti* (Borland *et al.*, 2009). The high soluble carbohydrates reserves of *Agave* plants and low lignin require less energy for conversion to fuel and may therefore result in higher quality feedstock (Smith, 2008; Borland *et al.*, 2009). Fructans contribute to plant development and metabolism which includes osmoregulation, cryoprotection and drought tolerance (French, 1989; Ritsema and Smeekens, 2003). There are advantages of accumulating fructan over starch as a protectant in abiotic stress; these includei) fructan's high water solubility, ii) fructan resistance to crystallization of membrane at subzero temperatures, and iii) normal function of fructan synthesis pathway at low temperatures (Vijn and Smeekens, 1999). The degree of polymerization differs with the growing stage of the plant (Lopez *et al.*, 2003; Simpson *et al.*, 2011b). Also, fructans are not as highly polymerized as glucans i.e starch, which maybe of significance for the osmotic pressure of CAM cells (Olivares and Medina, 1990).

Most of the research conducted on *Agave* has centered on A. *tequilana* due to its economic importance in the tequila production industry. However, there are other species of *Agave* that display higher biomass yields compared to *A. tequilana*. These include A. *mapisaga* and A. *salmiana* and A. *fourcroydes Lem* has been reported to possess high fructan content making it a promising plant for biofuel feedstock (Borland *et al.*, 2009; Somerville *et al.*, 2010).

The aim of this Chapter was to identify other species of *Agave* that could be exploited as sources of biofuel from semi-arid marginal lands. Some 14 different species of *Agave* that showed varying levels of succulence were compared. Species were evaluated for traits that included: the capacity for CAM, fructan content, carbohydrate composition, osmotic pressure and the relationship with succulence. Specific leaf areas were also measured. Leaf thickness plays an important role in the strategy for resource use (Vile *et al.*, 2005). For this reason, specific leaf area (SLA) may be used as a tool to screen different cultivars for productivity, and is a good indicator of leaf thickness and tissue density (Vile *et al.*, 2005). The experiments described in this chapter specifically addressed the following hypotheses:

H_1: leaf succulence is associated with increased magnitude of CAM across 14 *Agave* species that is manifested in nocturnal accumulation of titratable acidities.

H_2: Fructan content is linked with the potential for CAM and leaf succulence across *Agave*.

H_3: Sucrose rather than fructan is the substrate for nocturnal CO_2 uptake across different species of *Agave*

H_4: Carbohydrate composition influences leaf osmotic pressure in *Agave*

H_5: Specific leaf area is inversely related to the magnitude of CAM in *Agave*.

4.2 Materials & Methods

4.2.1 Plant Material

The *Agave* species chosen for this work were based on the degree of leaf succulence. Species included were: *A. deserti, A. parry truncula, A. univitata compacta, A. filementosa, A. americana (big blue), A. americana (Gainesvilla), A. americana (marginata), A. salmiana ferox, A. bractiose, A. desmetiana, A. ghiesbreghti, A. decipiens, A. ellemetiana and A. weberi.*

All species were analysed for CAM expression by titratable acidity measurements of leaf samples taken at dawn and dusk, under well watered conditions. Samples were collected from Biosciences Research group at the Oakridge National Laboratory in Tennessee. Plants were grown under a 12 h photoperiod with day/night temperature regime of $25^oC/19^oC$ and light intensity (PPFD) at plant height of ~500 µmol m^{-2} s^{-1}. All plants were grown in 20 cm diameter pots in commercial compost (Fafard 3B, Sun Gro Horticulture, Agawam, MA, USA) and were watered every 2-3 days.

4.2.2 Titratable Acidities

Measurements of leaf titratable acidity were made using leaf tissue from samples taken at dawn and dusk for 14 agave cultivars, 3 biological replicas for each. See section 2.2.3.

4.2.3 Soluble sugar analysis

Carbohydrate analysis was determined using a colorimetric method (Dubois *et al.*, 1956). Simple sugars give an orange yellow precipitate when treated with phenol and concentrated sulfuric acid. Analyses were performed on methanol extracts obtained as described previously (See section 3.2.3). The volume of methanol extract analyzed must fall within the linear range of glucose calibration. Exactly 0.1 ml of plant extract was added to 0.4 ml H_2O and 0.5 ml 5% phenol and then 2.5 ml concentrated sulphuric acid. Samples were mixed with a glass rod and left to cool for 15 minutes. Readings were taken at 483 nm using a spectrophotometer and compared with glucose standards of known concentration (See section 2.2.5).

4.2.4 HPLC analysis of sugars and fructans

High performance liguid chromatography (HPLC) was used to measure the concentrations of glucose, fructose, and sucrose present at dawn and dusk in samples taken from the 14 *Agave* species (3 biological replicas each) in mol/L using the method described by (Adams *et al.*, 1992). HPLC analysis was conducted by injecting 20 µl of each de-salted sample via a Rheodyne valve onto a Carbopac PA-100 column (Dionex, Sunnyvale, California, USA). Approximately, 100 µl of sample was placed into an analysis vial so as to ensure optimal immersion of the auto-sampler syringe. Sample components were eluted from the column isocratically using 100mM NaOH (de-gassed by helium) flowing at 1 ml/min for 8 min at room temperature. The chromatographic profile was recorded using pulsed amperometric detection with an ED40 electrochemical detector (Dionex, Sunnyvale, California, USA). Elution profiles were analysed using the Chromeleon software package (Thermo Fisher Scientific Inc., MA, USA). Daily reference curves were obtained for glucose, fructose and sucrose by injecting calibration standards with concentrations of 10

p.p.m. for each sugar. Total fructans were calculated using the subtraction method following acid hydrolysis (Liu *et al.*, 2011) see section 2.2.4).

4.2.5 Leaf Osmotic Pressure

Leaf sap extracts were analyzed for osmotic pressure using a Gonotec Osmometer 300. Leaf sap from obtained by crushing thawed leaf tissue in a garlic press. Exactly 50 µl of sample was placed in an eppendorf tube, inserted into the osmometer and the reading taken in mosmoles, for all 14 *Agave* species (3 biological replicas each). The osmometer was calibrated usingsugar standards supplied by the manufacturer of the osmometer (Gonotec GMbH, Berlin Germany).

4.2.6 Leaf succulence and specific leaf area

Succulence (kg m^{-2}) was measured by punching 3 discs of known area from one mature, fully expanded leaf of each *Agave* cultivar with 4 biological replicates taken for each species. Disc fresh weight was recorded. The same discs were dried at 70^{o}C, to constant weight then dry weight was recorded. Specific leaf area (SLA) was calculated as;

SLA = Leaf Area (cm^2)/ dry weight (g) [4.1]

4.2.7 Statistical Analysis

A correlation matrix was constructed for the 14 species of *Agave*. The variables were grouped on a fresh weight basis and area basis (see appendix C&D). Analysis was conducted using SPSS 19 statistical package, using Pearson's correlation which indicates strength and direction (+,-) of the correlation, p-value <0.05 and p-value <0.01 (2-tailed).

A correlation matrix is a good tool to investigate relationships between variables tested. It can display coefficients for more than one pair of variables at a time, and can compute partial correlation coefficients without the unneeded regression output. The grey shading of cells in the correlation matrix table

(Appendix C) indicates correlations of interest, and the dark blue cells, indicates the significance of the correlation.

4.3 Results

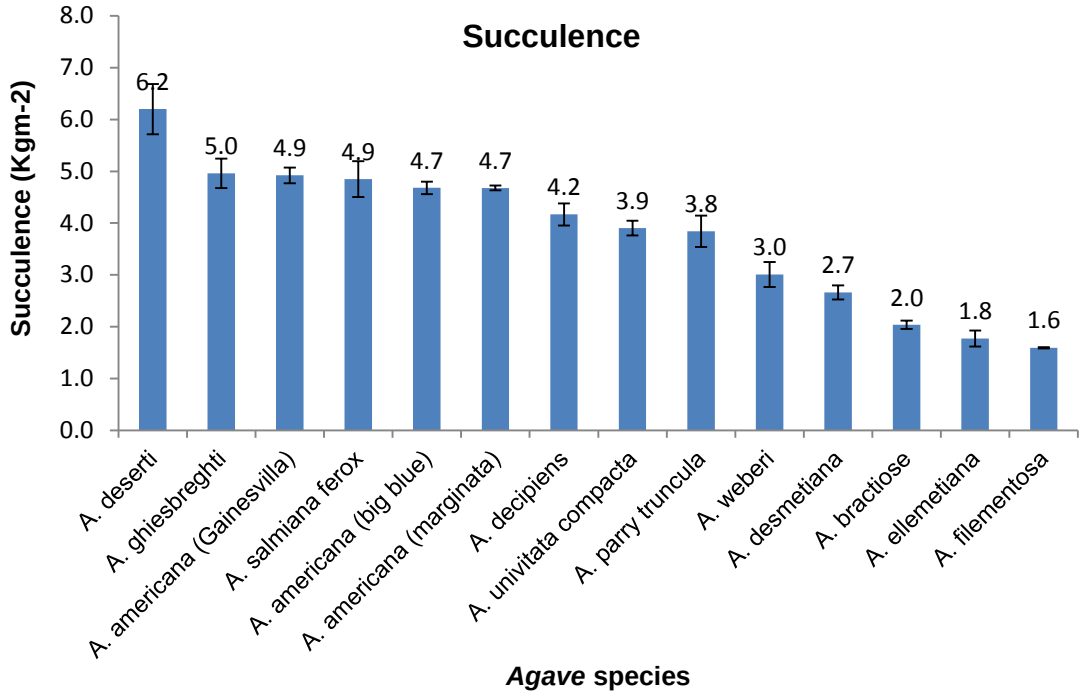


Figure 4.1 Mean values of leaf succulence across 14 different species of *Agave*. Each value is the mean of 4 biological replicates ± standard errors of mean.

The 14 different species of *Agave* that were studied showed a 4-fold range in leaf succulence (Fig. 4.1). Measurements of dawn and dusk titratable acidity were made to assess the magnitude of CAM in the different species and this was subsequently compared against the degree of leaf succulence.

4.3.1 Titratable Acidities

Titratable acidity analysis identified nocturnal acid accumulation as a marker for CAM expression in all 14 species of *Agave* assessed by differences in acidity measured at dawn and dusk, both on a leaf area basis (mmol m^{-2}) Figure 4.2 A, and on a leaf fresh weight basis (μmol H^+ g^{-1} fwt) Figure 4.2B

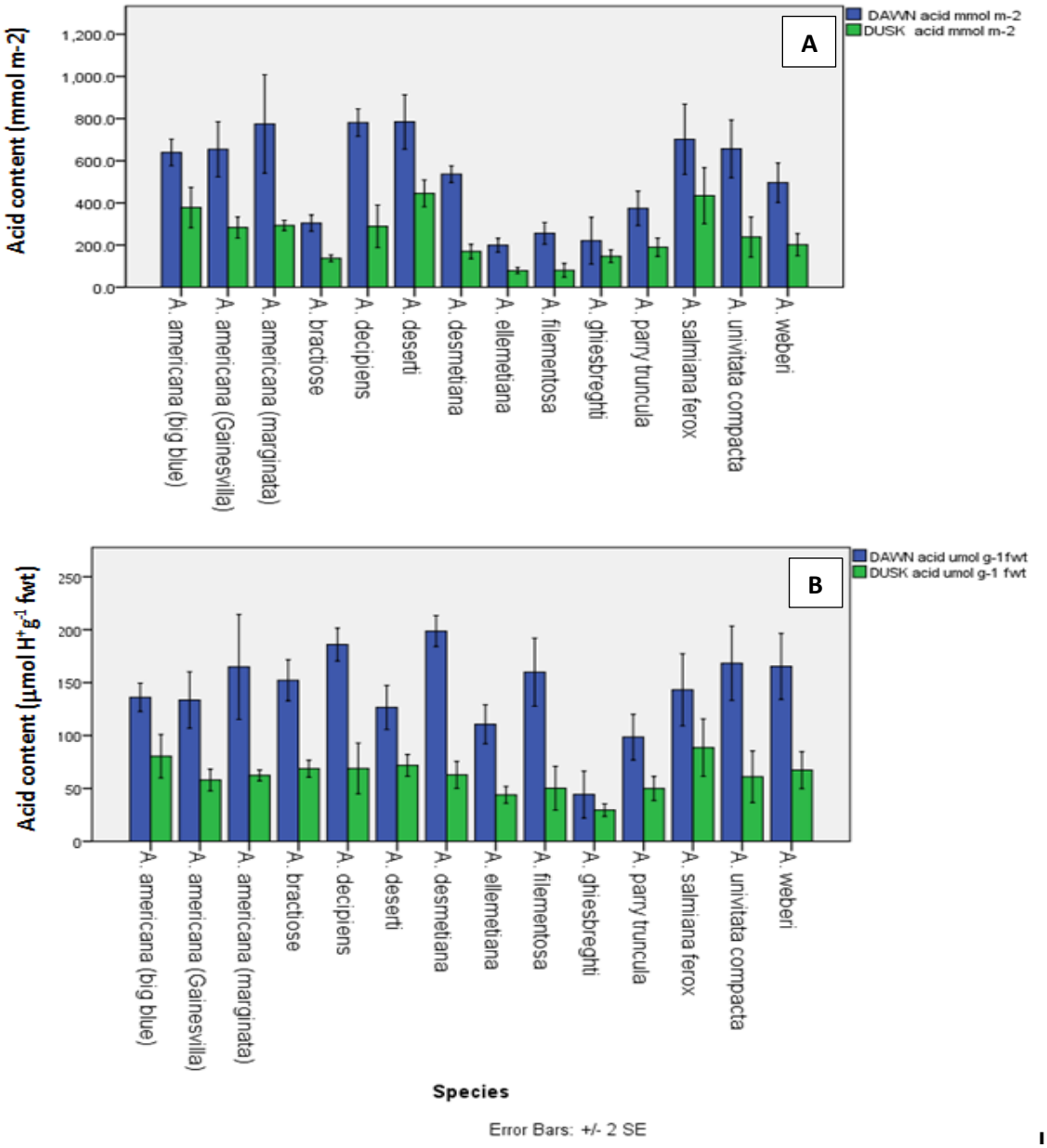


Figure 4.2 Day/night changes in acid content in 14 *Agave* species varying in succulence. (A) Data is expressed on leaf area basis (mmol m^{-2}). (B) Data expressed on leaf fresh weight basis (μmol H^+ g^{-1} fwt), for dawn and dusk samples (n = 3 ± standard errors).

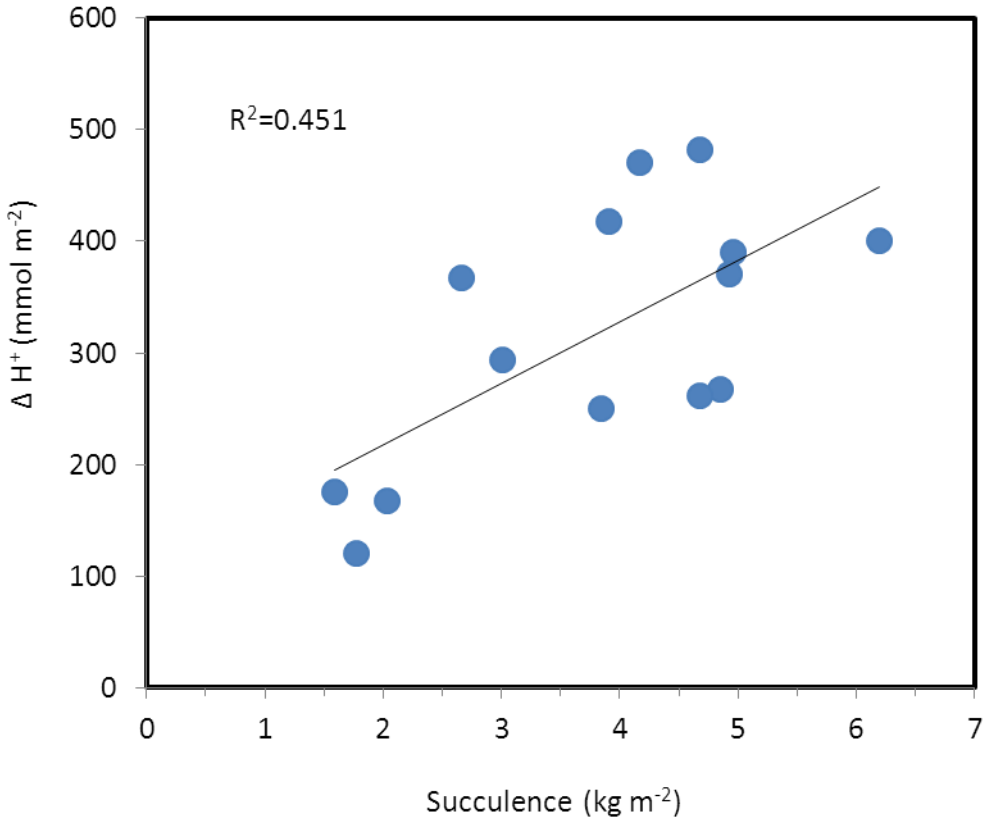


Figure 4.3 Correlation of CAM activity (measured as overnight accumulation of acidity (ΔH^+) with leaf succulence across 14 different species of *Agave*.

CAM activity expressed as the overnight accumulation of acidity was positively correlated with leaf succulence across the 14 species of *Agave* (Fig. 4.3). Moreover, acid content measured at dawn and at dusk was positively correlated with leaf succulence (Pearson's =0.364, p= 0.018) p value< 0.05.

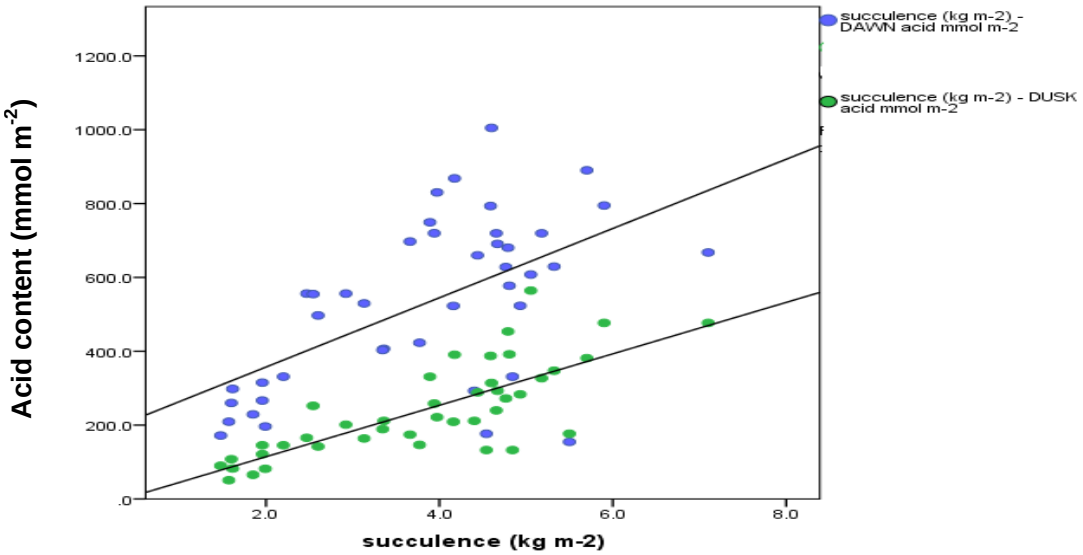


Figure 4.4 Multiple scatterplots of mean dawn and dusk acid contents measured as mmol m^{-2}, directly correlated with leaf succulence, dawn (Pearson's correlation=0.579, significance =0.000, R^2 =0.464) and dusk (Pearson's correlation=0.777, significance 0.000, R^2=0.690) p-value<0.05. N=42

Leaf acid contents measured at dawn and at dusk were also compared with leaf osmotic pressures at measured at comparable time points.

Dusk acid content levels had a direct correlation with dusk osmotic pressure, with a significance of 0.002 (p-value<0.05). Increasing levels of acid may facilitate osmotic water uptake and hence may act as a possible additional benefit to CAM in nocturnal storage of water. However, dawn acid content was not significantly correlated with dawn osmotic pressure (p-value=0.06; Figure 4.5).

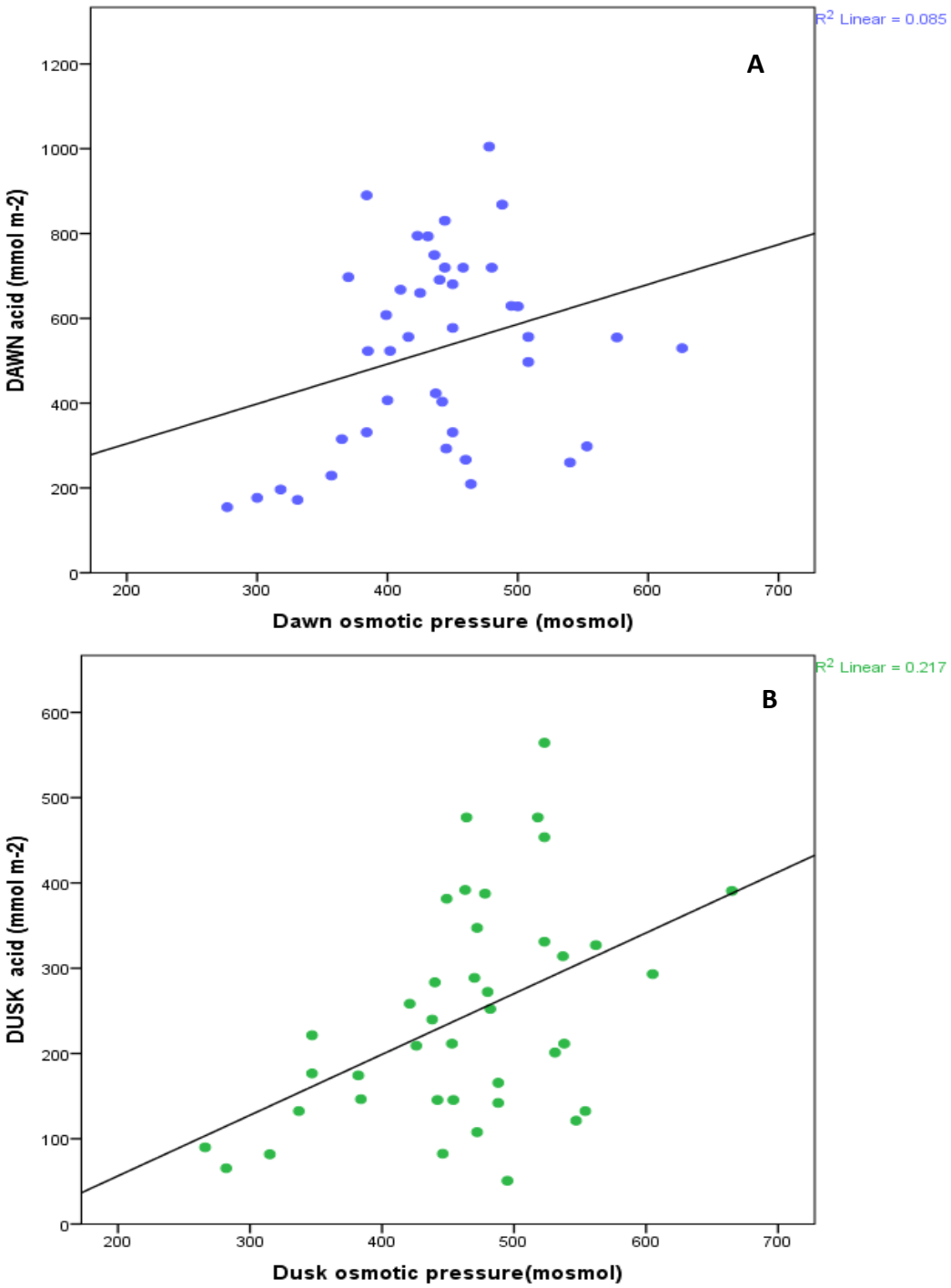


Figure 4. 5 Scatterplot of dawn acidity verses dawn osmotic pressure (A), and dusk acidity verses dusk osmotic pressure, (B) across 14 different species of Agave. Dawn values: (Pearson's correlation=0. 292 significance =0.060 R^2= 0.0.082) and dusk values: (Pearson's correlation=0.466, significance 0.002, R^2= 0.279) p-value<0.05. N=42

4.3.2 Soluble Sugars Analysis

Agave samples were analysed for their total soluble sugar content using phenol/sulphuric acid method (Dubois *et al.*, 1956), and isocratic HPLC analysis was used to identify the different sugars (glucose, fructose, sucrose, fructan).

All *Agave* species demonstrated the same trend of accumulating soluble sugars over the day. The most succulent species, *A. deserti* accumulated the highest amount of soluble sugars expressed on an area basis (Figure 4.6A) whilst *A. americana marginata* accumulated the most soluble sugars on a fresh weight basis (Figure 4.6B). Mean total sugars for dawn and dusk measured on an area basis (mmol glc m^{-2}) (Fig 4.6.A) were directly correlated with succulence (Kg $m^{-2;}$ see correlation analysis in Figure 4.7). However, when compared on fresh weight basis, succulence was not significantly correlated with the amount of soluble sugars with (p-value=0.359, Pearson's=0.145 for soluble sugars at dawn, and p-value=0.159, Pearson's=0.145 for soluble sugars at dusk, data not shown).

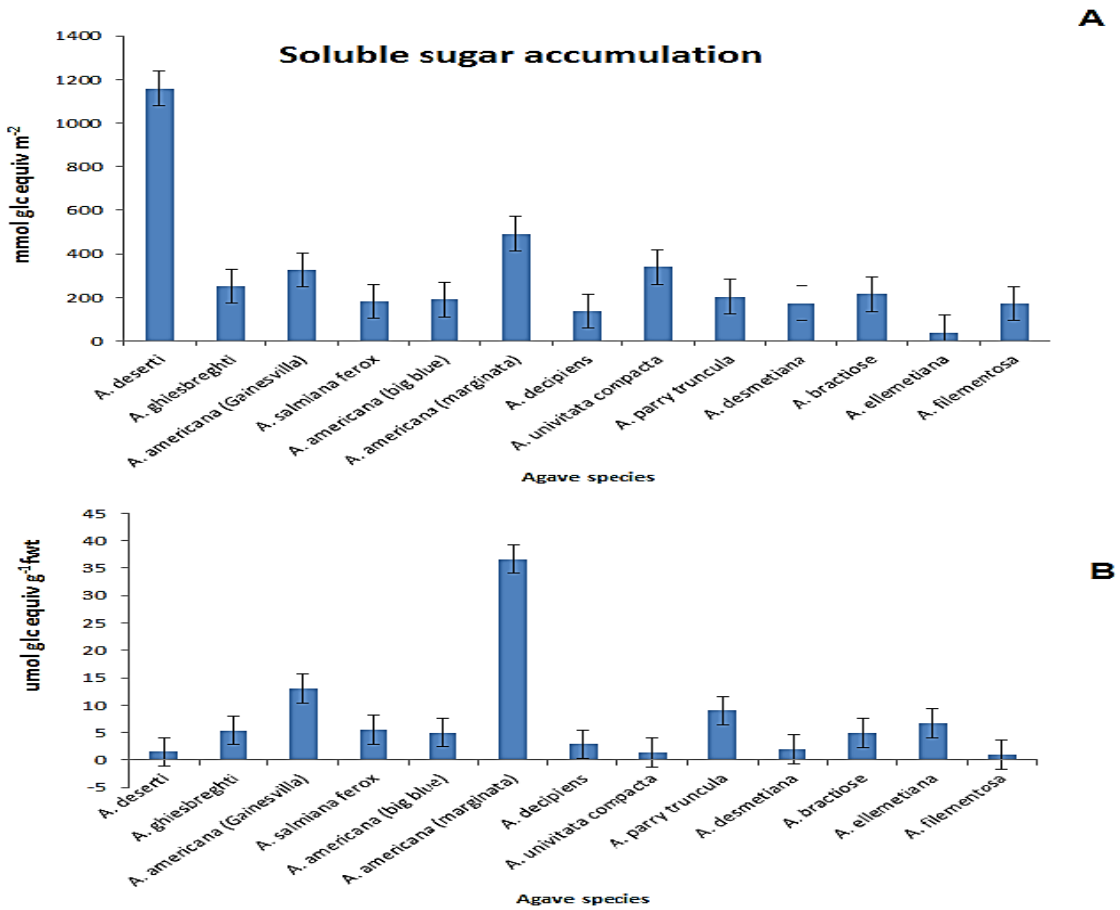


Figure 4.6 Day-time **s**oluble Sugar accumulation for 14 *Agave* species varying in succulence, (A) on area basis (g m^{-2}) and (B) on fresh weight basis umol glc equiv g-1 fwt. (n = 3 ± standard errors for error bars indicated).

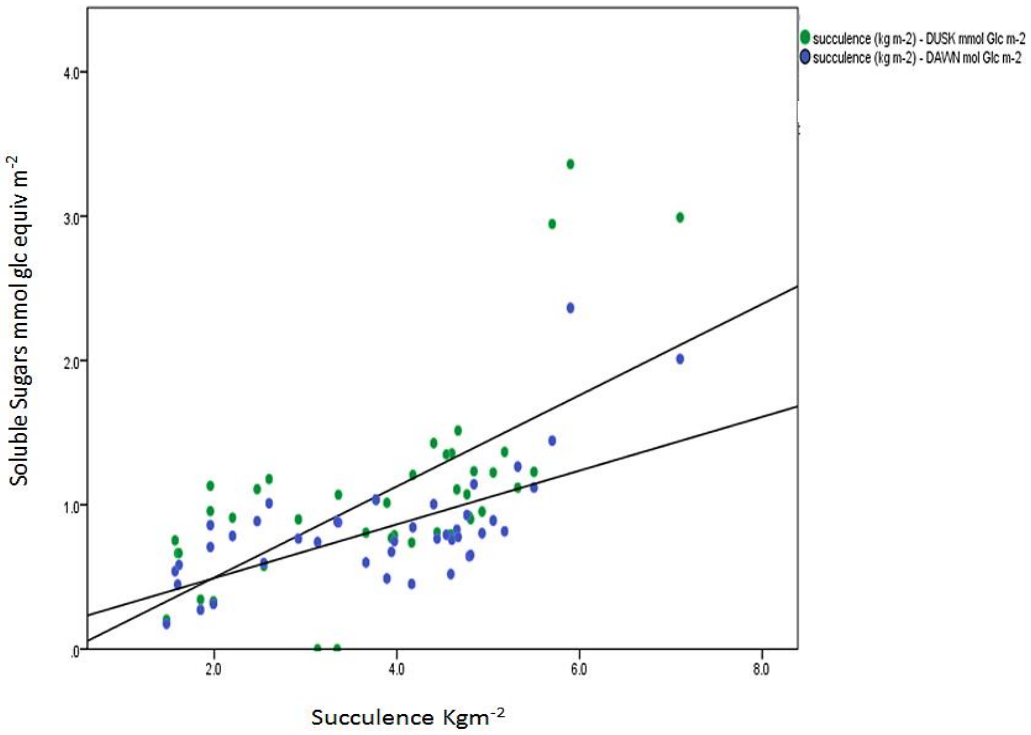


Figure 4.7 Scatterplot of mean total sugars for dawn and dusk measured on an area basis (mmol glc m^{-2}), directly correlated with succulence (Kg m^{-2}), dawn (Pearson's correlation=0.651, significance =0.000, R^2=0.472) and dusk (Pearson's correlation=0.660, significance 0.000, R^2=0.488) p-value<0.05 N=42

The total soluble sugar contents of leaf sap were measured at dawn and dusk, and were compared with osmotic pressures of leaf sap made at comparable time points Dusk total soluble sugar content showed a direct correlation with dusk osmotic pressure with significance of 0.029 (p-value<0.05). Osmotic pressure increased with dusk sugar content and could be important in driving changes in leaf osmotic pressure during the night (Figure 4.8).

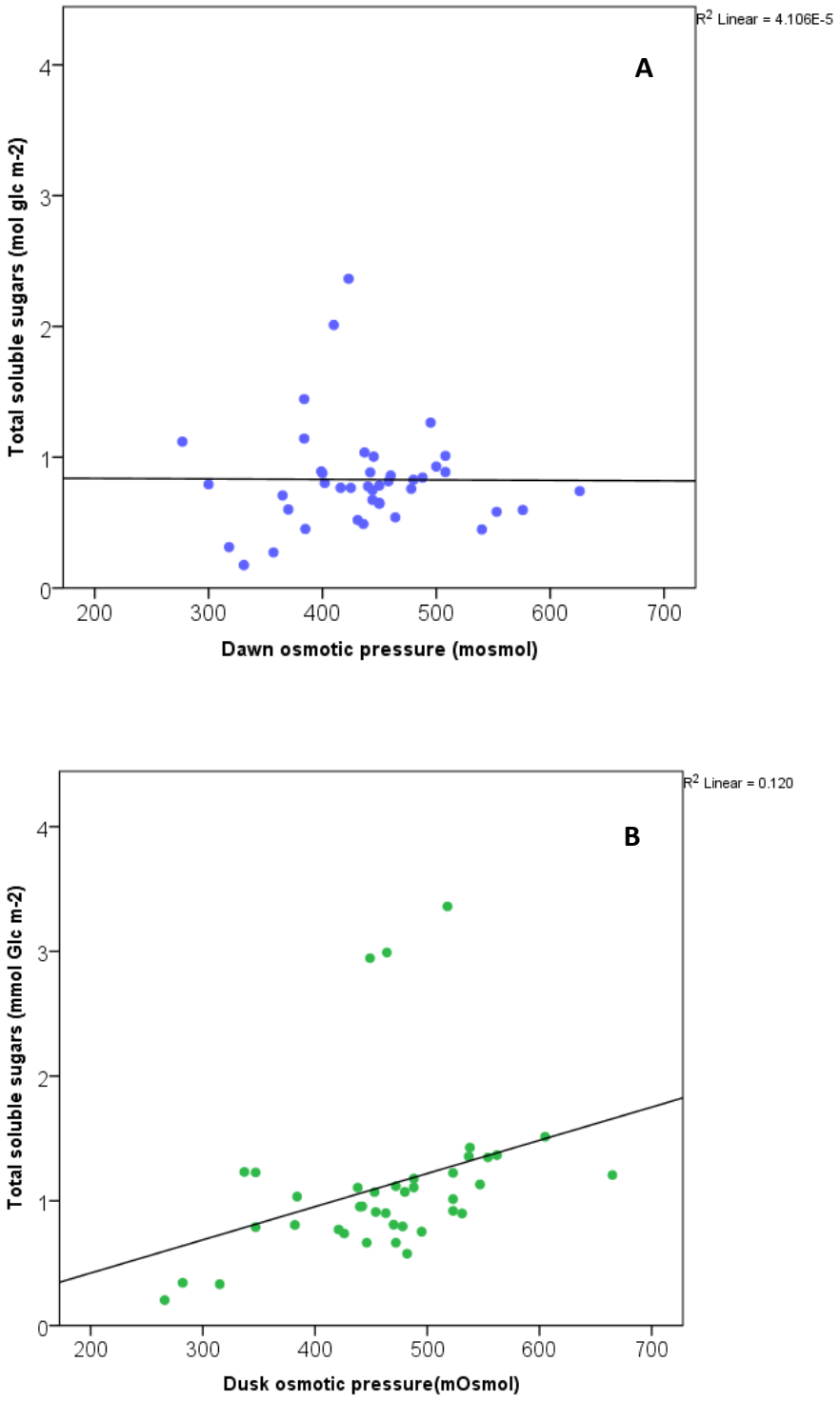
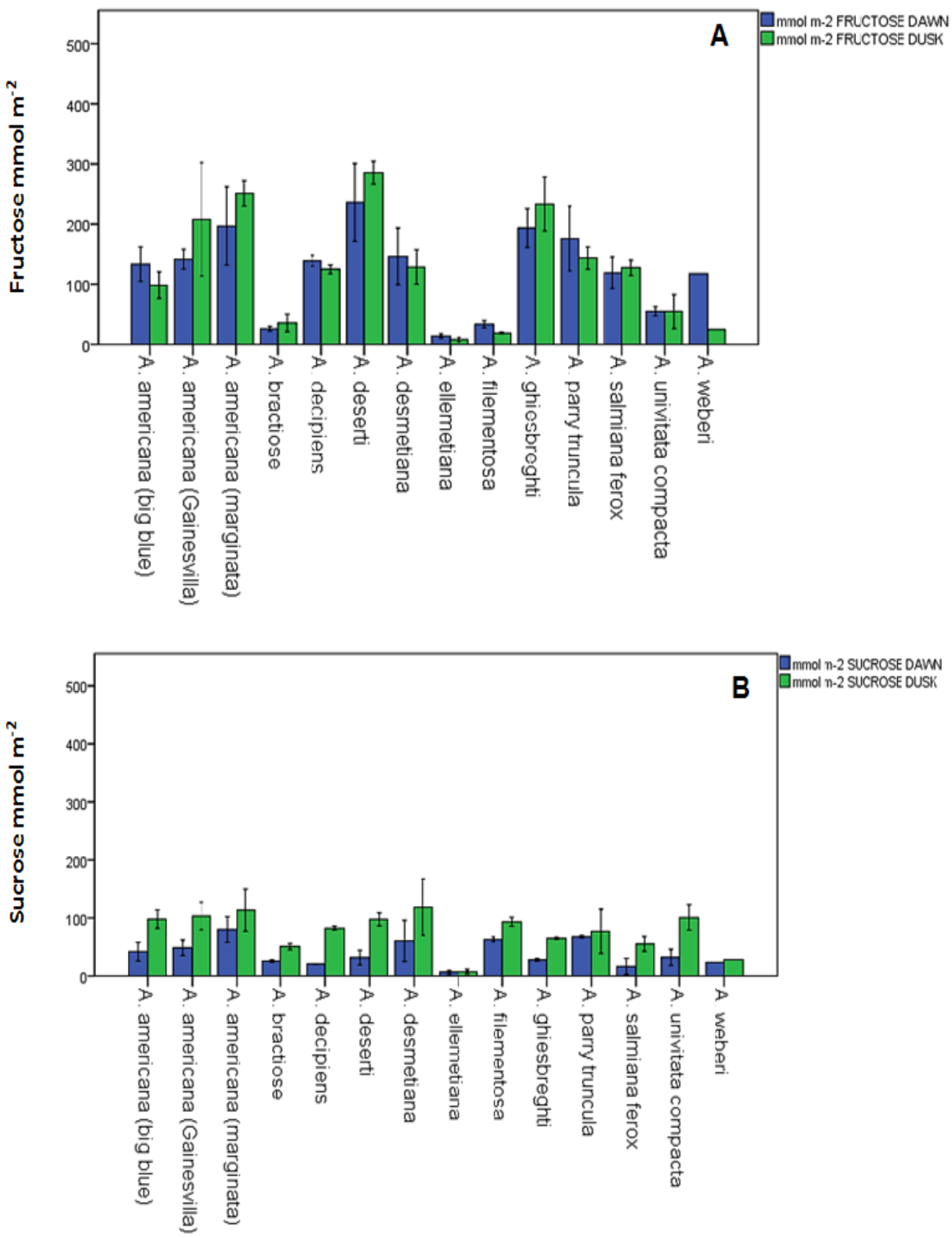


Figure 4.8 Scatterplot of dawn total soluble sugars verses dawn osmotic pressure, Figure (4.8A), and dusk total soluble sugars verses dusk osmotic pressure, Figure (4.8B). Total soluble sugars were measured as mmol glc equiv m^{-2}.Dawn values: (Pearson's correlation=-0. 006 significance =0.968 R^2= 4.106E-5) and dusk values: (Pearson's correlation=0.346, significance 0.029, R^2= 0.120) p-value<0.05 N=42

The composition of the soluble sugar pool was analysed using HPLC. (glucose was the most abundant soluble sugar in most of the *Agave* species, followed by fructose then sucrose (Figure 4.9). However, *A. desmetiana* had high levels of fructans exceeding glucose content (Figure 4.9D). Over-night depletion of sucrose had an inverse relationship with nocturnal acid accumulation (Pearson = -0.367, sig=0.017, R^2=0.135) and succulence (Pearson= -0.436, sig=0.004, R^2=0.186), correlation significant at p-value<0.05 (Figure 4.10). This data suggests that sucrose was a source of substrate for dark CO_2 uptake and thus the major substrate for production of PEP for PEPC activity. However, overnight depletion of fructan also occurred and displayed a positive relationship with nocturnal acid accumulation (Figure 4.11).


A
mmol m-2 FRUCTOSE DAWN
mmol m-2 FRUCTOSE DUSK
Fructose mmol m-2
500
400
300
200
100
0
A. americana (big blue)
A. americana (Gainesvilla)
A. americana (marginata)
A. bractiose
A. decipiens
A. deserti
A. desmetiana
A. ellemetiana
A. filementosa
A. ghiesbroghti
A. parry truncula
A. salmiana ferox
A. univitata compacta
A. weberi
B
mmol m-2 SUCROSE DAWN
mmol m-2 SUCROSE DUSK
Sucrose mmol m-2
500
400
300
200
100
0
A. americana (big blue)
A. americana (Gainesvilla)
A. americana (marginata)
A. bractiose
A. decipiens
A. deserti
A. desmetiana
A. ellemetiana
A. filementosa
A. ghiesbreghti
A. parry truncula
A. salmiana ferox
A. univitata compacta
A. weberi

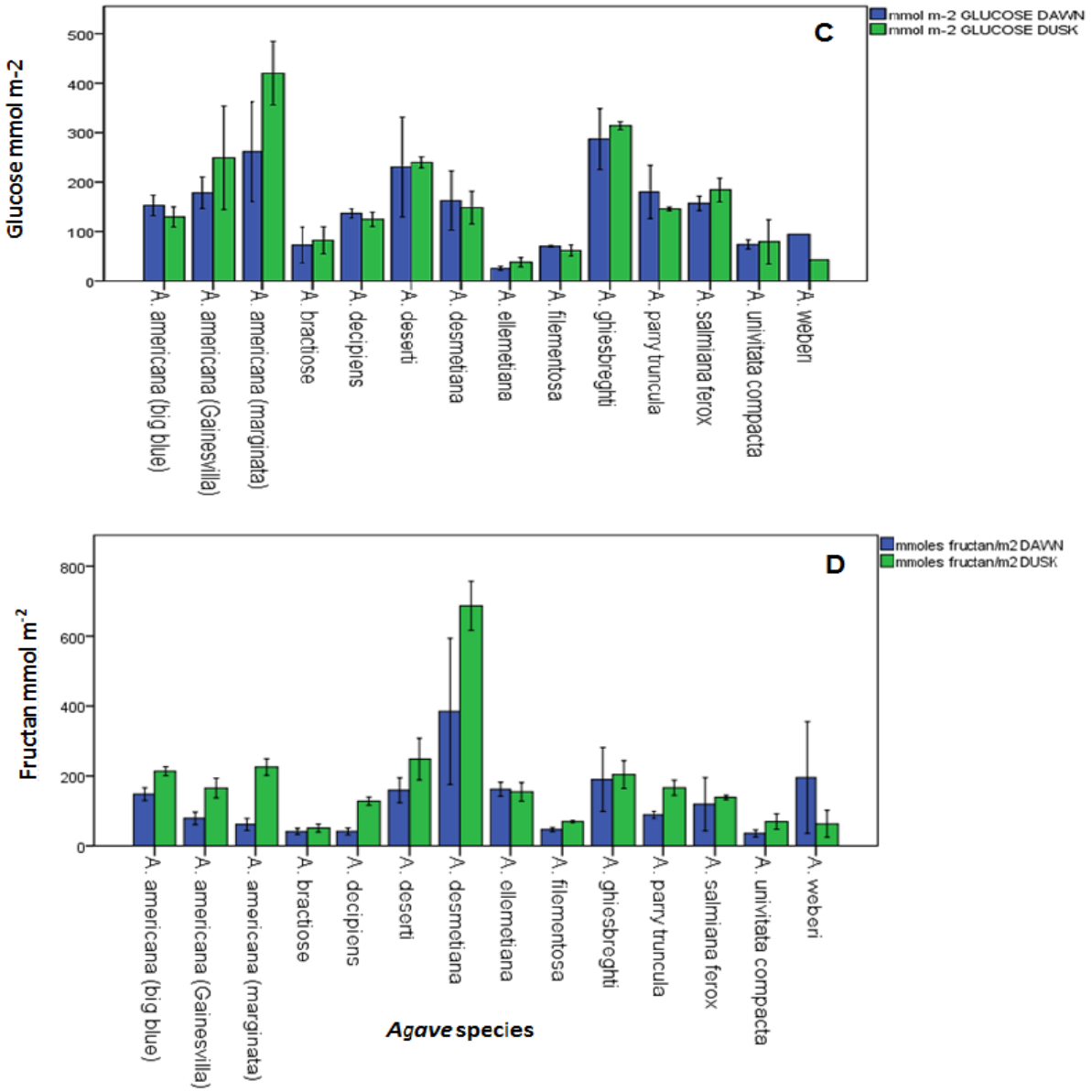


Figure 4.9 Day/night changes in (A) fructose, (B) sucrose and (C) glucose (D) fructans, in 14 species of *Agave* varying in succulence, on an area basis (mmol m^{-2}). (n = 3 ± standard errors).

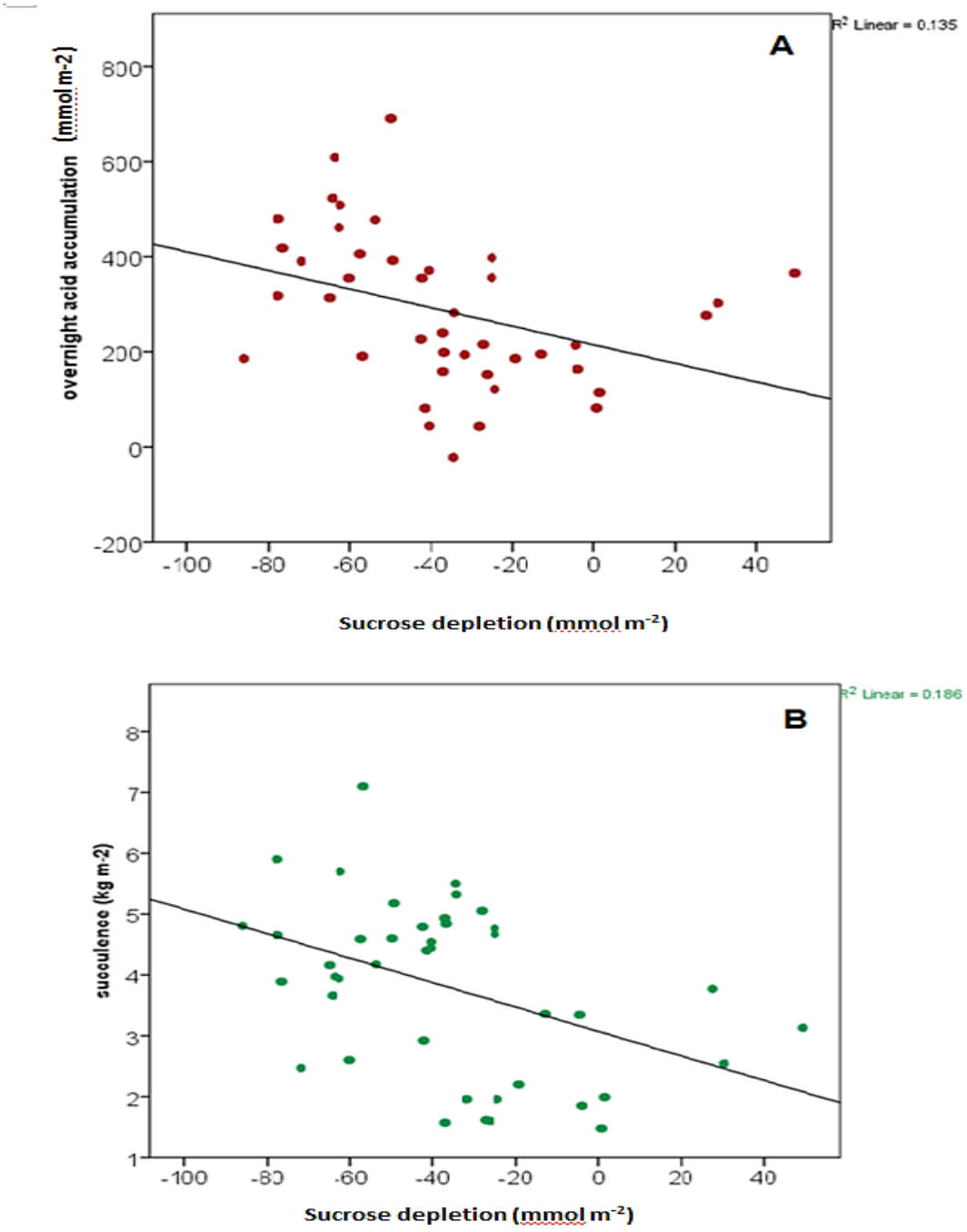


Figure 4.10 Scatterplot of nocturnal sucrose depletion (mmol m^{-2}) correlated with (A) nocturnal acid accumulation (mmol m^{-2}) (Pearson =-0.367, sig=0.017,) (B) succulence (Kg m^{-2}) (Pearson= -0.436, sig=0.004). N=42

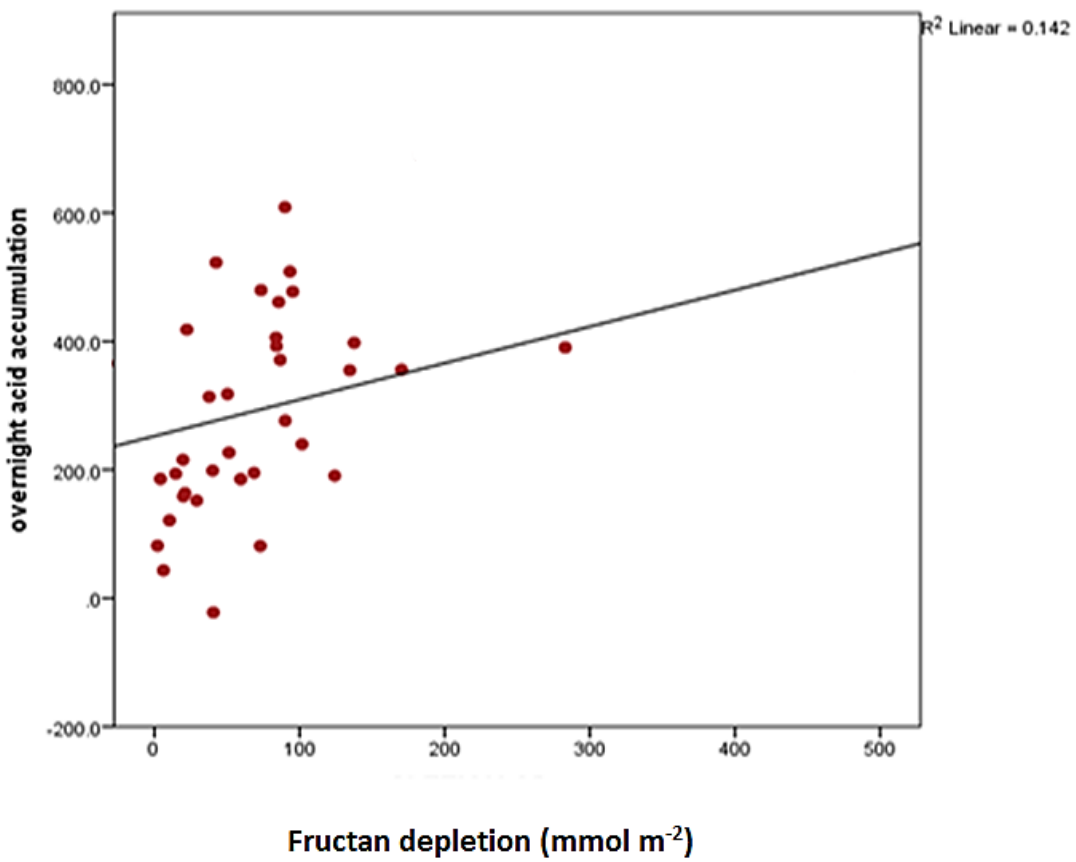


Figure 4.11 Scatterplot of nocturnal fructan depletion (mmol m^{-2}) correlated with nocturnal acid accumulation (Pearson's=0.377, sig=0.014 with p-value<0.05. N=42

4.3.4 Specific leaf area and CAM in *Agave*

Specific leaf area (SLA) measurements were taken for the 14 species of *Agave*, which showed a significant inverse relationship with succulence, and the magnitude of CAM (R^2= 0.113, Pearson's correlation= -0.436, sig= 0.004) (Figure 4.12).

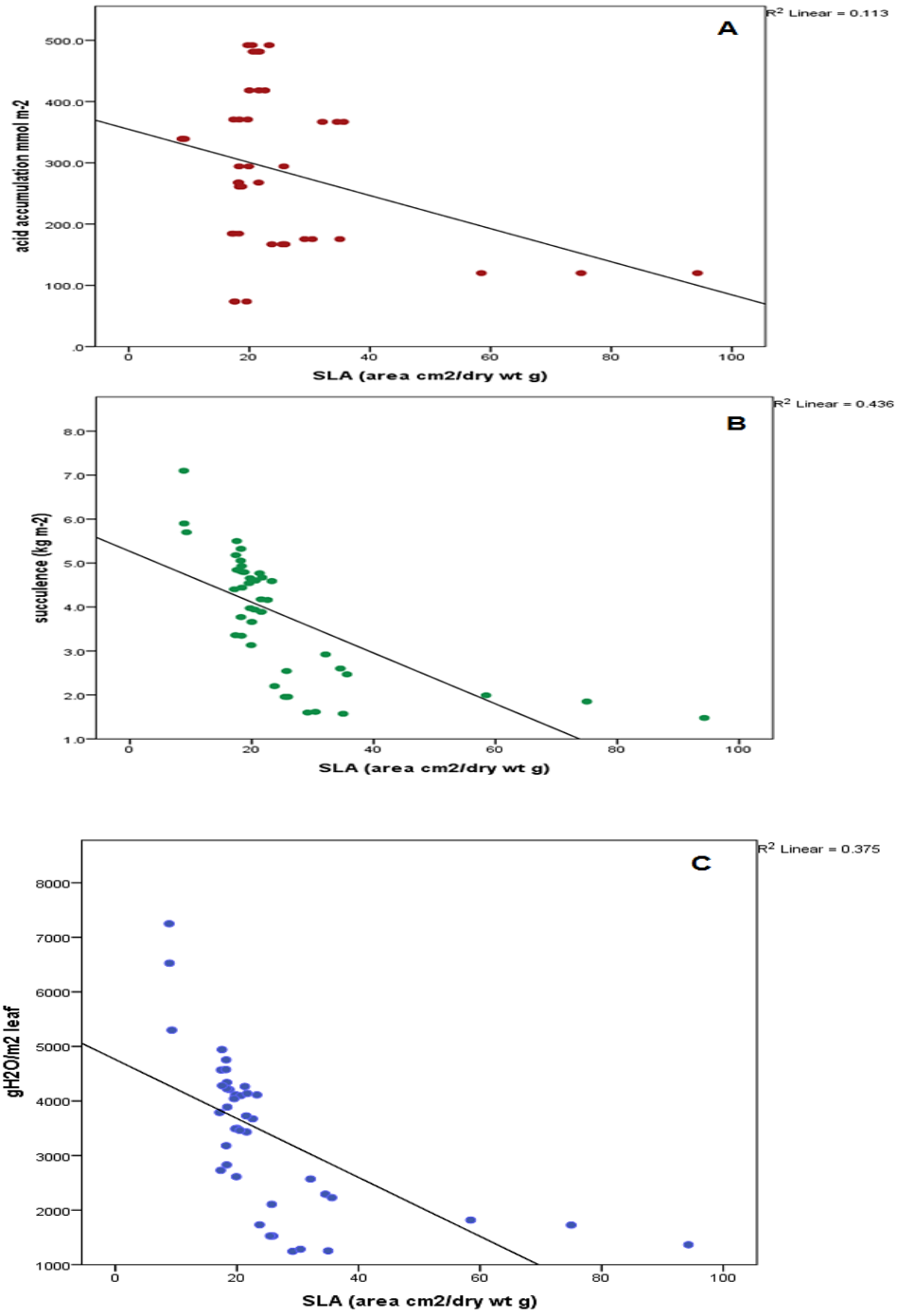


Figure 4.12 Scatterplot of inverse correlation of SLA with (A) acid accumulation expressed on area bases m^2, R^2= 0.113, Pearson's correlation= -0.436, sig= 0.004(B) succulence (Kg m-2), R^2=0.436, Pearson correlation= -0.661, sig= 0.006(C) SLA and leaf water content (g H_2O/m2 leaf). Pearson's correlation= -0.611, sig= 0.00).

Succulence gave a strong positive correlation with leaf water content (g H_2O/m_2 leaf). R^2=0.980. See Figure 4.13.

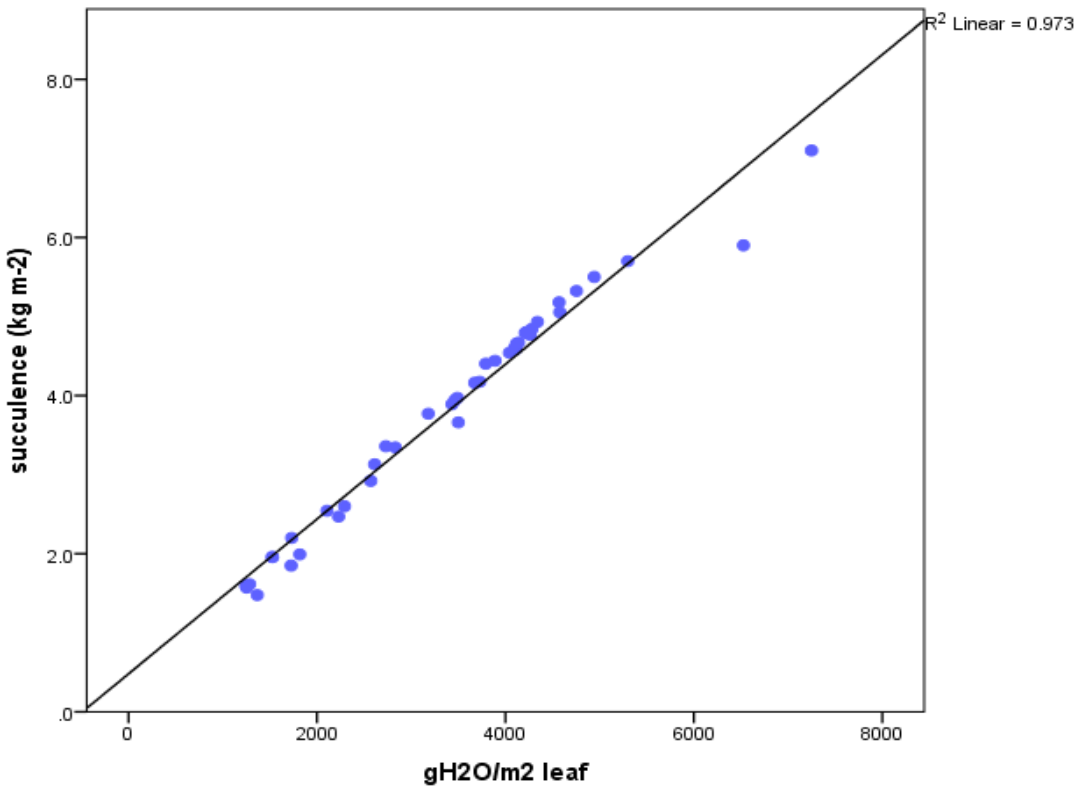


Figure 4.13 Positive correlation between succulence (Kg m^{-2}) and leaf water content (gH_2O/m_2), R^2= 0.973, Pearson's = 0.980, significance=0.00

Total soluble sugar content measured at dusk was also calculated on a dry weight basis and the portion of the soluble sugars required for the measured overnight accumulation of acids was calculated on the understanding that 1 mole of glc equivalents will give 1 mole of malic acid (or 2 H^+, Figure 3.14). *A.deserti* and *A. desmetiana* showed the highest total sugar contents (on a dry weight basis) and *A. deserti* (the most succulent species) invested more sugars into CAM (on a dry weight basis) compared to the other *Agave* species. Fructan content was also measured on dry weight basis. *A.desmetiana* showed the highest fructan content on a dry weight basis; Figure 4.15).

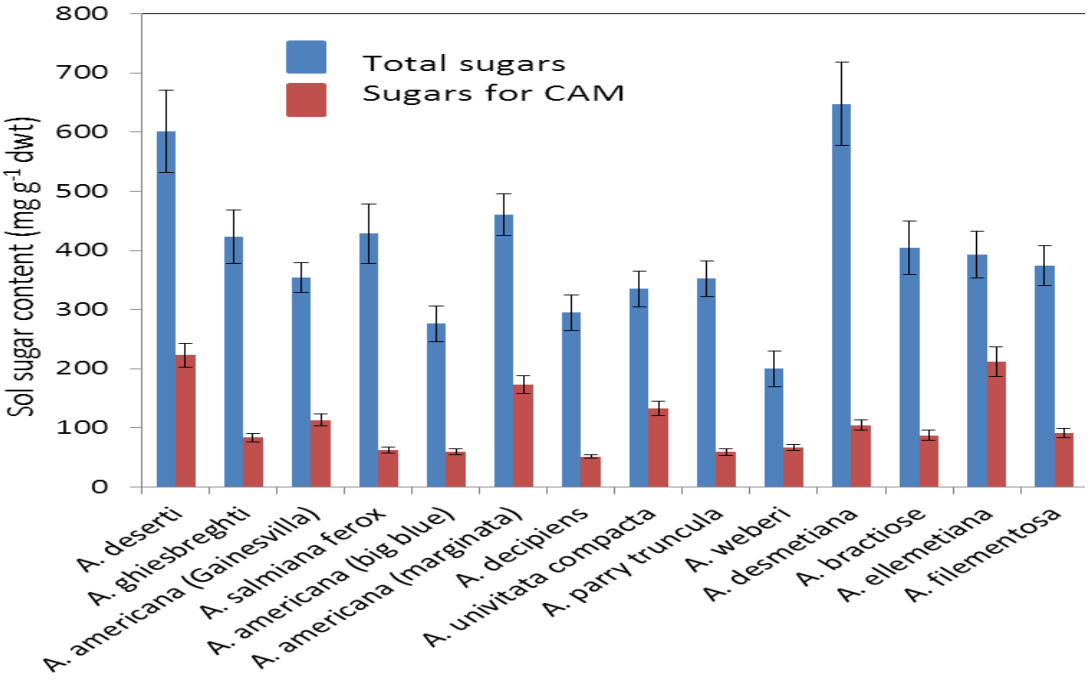


Figure 4. 14 The total soluble sugars and the contribution from these sugars to on dry weight basis across 14 *Agave* species. (n = 3 ± standard errors)

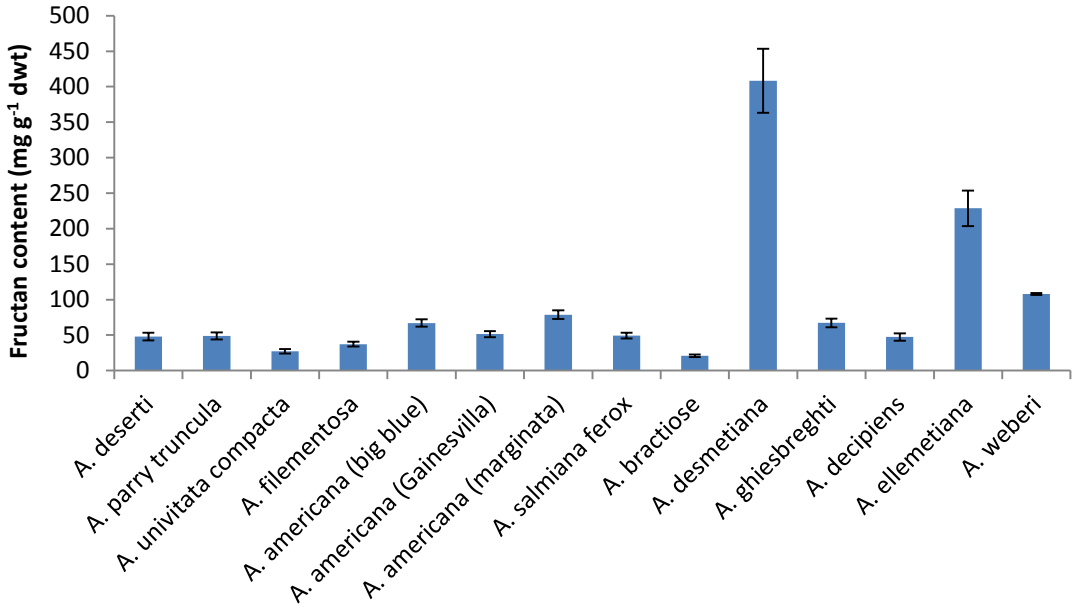


Figure 4.15 Fructan content on dry weight basis across 14 *Agave* species (n = 3 ± standard errors).

4.4 Discussion

This study compared the potential of 14 different *Agave* species, varying in succulence, under well watered conditions, as sources of bioethanol from marginal lands, by assessing the capacity for CAM and sugar content.

4.4.1 Leaf morphology alters commitment to CAM

Leaf succulence is an important anatomical trait in CAM plants and is a key morphological correlate of the capacity for CAM (Winter *et al.*, 1983; Borland *et al.*, 1998). As predicted, thicker, more succulent leaves of *Agave* showed an increased commitment to CAM, manifested as overnight accumulation of acidity as well as the acid contents measured at dawn and at dusk. Large vacuoles provide capacitance for nocturnal storage of malic acid and act as water reservoirs, enhancing photosynthetic carbon gain and reducing vulnerability to water stress (Smith *et al.*, 1996; Osmond *et al.*, 1999; Borland *et al.*, 2000). The data presented in this chapter also showed a clear correlation between succulence and leaf water content. High succulence may lead to tight cell packing and low intracellular air space (Maxwell *et al.*, 1997; Nelson *et al.*, 2005), enhancing photosynthetic efficiency by restricting CO_2 efflux during the decarboxylating (phase III) of CAM (Maxwell *et al.*, 1997; Borland *et al.*, 2000). This will enhance CAM function during times of severe drought that might limit uptake of atmospheric CO_2 (Borland *et al.*, 2000).

Specific Leaf Area (SLA) is a key leaf functional trait that has been widely used to provide information on plant growth rate and resource-use strategy in C3 plants (Garnier, 1992; Lambers and Poorter, 1992; Reich, 1993; Vendramini *et al.*, 2002). SLA is considered the best candidate for inclusion in large screening program for comparative databases (Vendramini *et al.*, 2002) .Variation in SLA depends on leaf water content (LWC), which has a close correlation with tissue density (Witkowski and Lamont, 1991; Garnier and Laurent, 1994) and leaf thickness (LT) (Witkowski and Lamont, 1991; Shipley, 1995; Cunningham *et al.*, 1999; Pyankov *et al.*, 1999; Wilson *et al.*, 1999). As a good indicator for leaf thickness and tissue density (Vile *et al.*, 2005) SLA generally shows an inverse relationship with succulence. The data presented above confirm the inverse relationship of SLA with both succulence and the magnitude of CAM for the 14

species of *Agave* studied (i.e. SLA was lower which means thicker, denser leaves) for those *Agave* species which showed increased acid accumulation). In previous studies (Vendramini *et al.*, 2002), results suggest that SLA is a better predictor of species resource-use strategy than leaf water content (LWC) in succulents. In addition, SLA serves to elucidate converging strategies in carbon assimilation and nutrient conservation (Vendramini *et al.*, 2002). The carbon and nutrients invested in a certain area of light intercepting foliage varies, and plants with lower SLA might have a higher level cost for light interception (Poorter *et al.*, 2009). Plants with this strategy tend to inhabit drought and limited nutrient environments as exemplified by *Agave.* Low SLA is a key trait which acts as an additional benefit to *Agave* living in marginal lands. In a comparative study on *Peperommia* and *Clusia*, cross sections of water storage parenchyma (WSP) were inversely correlated with the capacity of CAM (Gibeaut and Thomson, 1989; Borland *et al.*, 1998). This might suggest that the large cells of WSP in *Peperomia* could not resist losing water to the environment under extreme conditions (Kaul, 1977). Therefore, thicker cuticle and lower surface areas are more effective in reducing water loss under extreme exposure as in *Clusia* (Borland *et al.*, 1998) and probably also in *Agave* .

In this screening of 14 *Agave* species, increased levels of acid accumulated overnight were accompanied by an increase in leaf osmotic pressure which could expedite osmotic water uptake by cells. In a previous study on the cactus *Cereus validus*, malate concentration and stem osmotic pressure increased during night time CO_2 fixation, indicating that changes in malate affected the water relations of the succulent stems (Lüttge and Nobel, 1984), which could act as an additional benefit of CAM for nocturnal water storage. *Agave* species accumulate soluble sugars and fructans rather than insoluble and osmotically inactive starch, which can also influence the osmotic pressure and osmotic adjustment of leaf cells (Olivares and Medina, 1990). The data presented above showed a positive correlation of nocturnal accumulation of soluble sugars with an increase of overnight osmotic pressure. Studies on *Fourcroya humboldtiana* demonstrated relatively high osmotic pressures due to the accumulation of osmotically active soluble carbohydrates (Olivares and Medina, 1990). This

nocturnal increase in osmotic pressure could be crucial for maintaining turgor during dark CO_2 uptake in the water-limited habitats that *Agave* frequents.

4.4.2 Plasticity of carbohydrate source pools driving the nocturnal CO_2 uptake in *Agave*

Reserve carbohydrates in CAM represent a substantial investment of resources which are essential for nocturnal CO_2 uptake whilst at the same time, carbohydrates have to support other metabolic activities such as acclimation to abiotic stress, dark respiration and growth (Ceusters *et al.*, 2009). CAM plants are biochemically diverse in the carbohydrate species which are degraded at night. They range from species that use cytosolic mono, di or oligosaccharides to species that use chloroplastic starch (Christopher and Holtum, 1996). Nocturnal breakdown of carbohydrates generates the 3C substrate PEP for PEPC. Carbohydrate turnover is an essential component determining the magnitude of CAM (Borland and Dodd, 2002). The variations in carbohydrate source used to provide PEP for nocturnal CO_2 uptake between different CAM species is probably the result of constraints imposed by CAM and diversity in biochemistry resulting from different evolutionary histories (Christopher and Holtum, 1996).

Agave species accumulate fructans that are synthesised from sucrose and are accumulated in vacuoles of the leaf parenchymatous cells. The data presented in this chapter indicated that nocturnal breakdown of fructan content had a positive relationship with the magnitude of CAM across 14 species of *Agave.* Evidence from previous studies on *A. americana* suggested that fructans are not broken down during the dark period to provide PEP as a substrate for nocturnal CO_2 fixation (Raveh *et al.*, 1998). The same study indicated that diel fluctuations in sucrose could account for more than 83% of carbon needed for nocturnal PEP regeneration. Findings in Chapter 2, showed that sucrose was the major sugar used for nocturnal acid production in *Agave* species under investigation. In chapter 2, stoichiometric analyses of sugar breakdown and PEP requirements for CAM indicated that of the 3 *Agave* species studied in that chapter, only *A. americana* showed a shortfall in sugar depletion, implying that some nocturnal fructan depletion may be required in this species to provide

PEP. In this chapter, *A. desmetiana* had the highest fructan content on a dry weight basis, which is important in terms for bioenergy harvesting perspective.

On the other hand, it has been reported for other species of *Agave*, such as *A. guadalajarana* that diel fluctuations in leaf glucose, fructose and sucrose could not account for the carbon needed for night time PEP production, thus this species required an alternative carbohydrate such as fructan to provide nocturnal PEP (Christopher and Holtum, 1996). This was similar to results conducted on *Agave humboldiana* which showed an inverse relationship between fructan and malic acid (Olivares and Medina, 1990).

Flexibility in the major carbohydrate source used for the sustainability of dark CO_2 uptake could be a key attribute for bioenergy feedstocks like *Agave* which are capable of maintaining carbon acquisition under environments with limited precipitation.

4.5 Conclusions

The data presented in this chapter has confirmed that under well watered conditions inter-specific variations in the magnitude of expression of CAM in *Agave* are dependent on leaf succulence. The day/night changes in malic acid and soluble sugar contents also affect the cell sap osmotic pressure and water relations of *Agave*. Increasing levels of malic acid uptake facilitate osmotic uptake of water by cells, which is an additional benefit of CAM to nocturnal water storage (Lüttge and Nobel, 1984). Accumulation of osmotically active soluble carbohydrates can contribute to high osmotic pressures (Olivares and Medina, 1990), Soluble sugars serve as the precursors for nocturnal organic acid synthesis (Borland and Griffiths, 1989) and may also contribute to water stress tolerance in *Agave*.

Agave displays flexibility in the use of carbohydrate source pools to sustain dark CO_2 uptake. Some species appear to use fructans and others sucrose as substrate for dark CO_2 uptake. This is of importance in terms of vacuolar sugar transporters which are hypothesized to play a key regulatory role in determining sucrose turnover for CAM and fructan accumulation. Thus, vacuolar sugar transporters could represent future targets for genetic engineering of increased

sugar content for plants grown for bioenergy (Antony *et al.*, 2008; Antony and Borland, 2009; Borland *et al.*, 2009). This topic will be addressed in Chapter 5.

Chapter 5

Vacuolar sugar transporter identification in *Agave americana marginata*

5.1 Introduction

In leaves, mesophyll cells harbour large central vacuoles in which sugars, hydrolytic and biosynthetic enzymes, inorganic ions, organic acids, amino acids and secondary compounds (Maeshima, 2001; Martinoia *et al.*, 2002) are stored. In CAM plants, these central vacuoles, which are surrounded by a single permeable membrane (i.e. the tonoplast) are large in size and can occupy 80-95 % of total cell volume (Winter *et al.*, 1993; Neuhaus, 2007). From the diversity of compounds and enzymes located in the vacuole, this organelle can be described as a core structure for energy management, accumulation of nutrients and reserves, regulation of cellular osmotic pressure, detoxification and ecological interactions (Neuhaus, 2007). Several of the compounds found in the vacuole accumulate by secondary active transporters against an existing concentration gradient; this process is driven by electrochemical gradients generated by two types of proton pumps; a vacuolar type (V-type) H^+-ATPase and H^+-PP_iase (Rea and Sanders, 1987; Kluge *et al.*, 2003) which are present on the tonoplast membrane (Hedrich *et al.*, 1989; Maeshima, 2000; Maeshima, 2001). Typical organic compounds which accumulate in the vacuole are carbohydrates, fructans and carboxylic acids. In CAM plants, malate enters the vacuole either by an anion channel specific for $malate^{2-}$ (Hafke *et al.*, 2003) or by a solute carrier (Emmerlich *et al.*, 2003). The central vacuole also enables cells to reach a large size, allows chloroplasts to be distributed around the cell periphery for optimal light capture and efficiency and it allows the cell to keep cytosolic concentrations of ions and metabolites optimal for metabolism (Boller and Wiemken, 1986; Martinoia, 1992; Martinoia *et al.*, 2000; Maeshima, 2001).

In CAM plants, the vacuole serves as a storage reservoir for malic acid which accumulates as a consequence of dark CO_2 uptake. In CAM species, an equivalent of 17% of total cell dry mass may cross the tonoplast everyday (Holtum *et al.*, 2005). The three major components of the tonoplast are V-ATPases and V-PPases that catalyse the transport of H^+ into the vacuole (Marquardt and Lüttge, 1987) and aquaporins (water channels).

Agave species use soluble sugars to provide the substrate (PEP) for dark CO_2 uptake (Black *et al.*, 1996) as observed in Chapters 2 and 4. Thus, vacuolar

sugar transporters likely play a key role in the diel operation of the CAM cycle in *Agave* (Kenyon *et al.*, 1985; Christopher and Holtum, 1998). The capacity of the vacuole as a sink for carbohydrate may be an important determinant of CAM expression and has important implications for plant growth and productivity. Up to 20% of leaf dry weight contributes as carbohydrate reserves for CAM (Black *et al.*, 1996), but the potential of high productivity is not compromised, with some *Agave* species productivity rivalling sugar cane (Bartholomew and Kadzimin, 1977; Nobel, 1996).

Sugar transporters have been recognised as key targets for regulatory roles in long distance and subcellular distribution and partitioning of assimilates (Williams *et al.*, 2000; Lalonde *et al.*, 2004). Thus in CAM plants, sugar transporters represent an important checkpoint in regulating partitioning of photo-synthetically fixed carbon between supply of substrate on one hand and for nocturnal carboxylation and export for growth on the other hand (Antony and Borland, 2009). As outlined in the general introduction (Fig 1.21) it has been proposed that in CAM plants which store vacuolar soluble sugars, transport of sucrose into the vacuole would occur during the day whilst export of hexoses would occur at night to fuel the production of PEP (Antony and Borland, 2009). Examination of the proteome of vacuolar membranes of *Arabidopsis* cells provided first evidence on the molecular nature of a vacuolar sucrose carrier (Endler *et al.*, 2006). The first transport proteins involved in the movement of monosaccharides across the tonoplast have been identified which belong to the Tonoplast Monosaccharide Transporter (TMT) group (Wormit *et al.*, 2006) and belong to the monosaccharide transporter (-like) (MST) gene family (Lalonde *et al.*, 2004). These are integral membrane proteins which are localized to the tonoplast membranes (Wingenter *et al.*, 2010). *At*TMT were directly identified from *Arabidopsis* with 12 predicted transmembrane α helices and comprised of two units of six helices connected by central loop varying in length (Lemoine, 2000). The *At*TMT transporters are believed to operate using a proton-coupled anti-port mechanism, allowing active transport and accumulation of hexoses (glucose and fructose) in the vacuole especially when induced by cold, drought or salinity. These stimuli promote sugar accumulation in *Arabidopsis* (Wormit et al., 2006). To date, the transporters responsible for sucrose and hexose transfer across the tonoplast membrane have not been identified in *Agave*.

The central aim of this chapter was to develop a method to identify candidate vacuolar sugar transporters in *Agave*. The first step was to isolate a tonoplast-enriched protein fraction, exploiting as a guide the activity of two known vacuolar markers, ATPase and PP_iase of leaf vesicles of *Agave americana marginata*, and their sensitivity to inhibition by known inhibitors. Secondly, a proteomics GeLCMSMS approach was used to analyse the tonoplast-enriched fraction with the aim of identifying vacuolar sugar transporter proteins. The focus on identifying vacuolar sugar transporters was due to the hypothesis that these play key regulatory roles in determining sugar turnover for CAM and fructan accumulation.

5.2 Materials & Methods

5.2.1 Plant material

The *Agave* species under investigation was *A*. a*mericana marginata*, seen in Figure 5.1. This species was chosen since an extensive transciptome and proteome database has been created for it by the Plant Systems Biology group at the Oak Ridge National Laboratory. All plants were maintained under controlled conditions of a 12 hour photoperiod and day/night temperatures of 28/22^{O}C with a photon flux density of 300 µmol m^{-2} s^{-1}. Soil was made up in 127 mm pots containing a mixture of 1 part sharp sand (J. Arthur Bower's, UK), 4 parts John Innes No. 3 (JI no. 3), 1 part gravel. Plants were watered twice a week.

Figure 5.1 Plants of *Agave americana marginata* used for tonoplast isolation.

5.2.2 Tonoplast extraction and purification

The method for tonoplast extraction was based on previous work on tonoplast extraction from the CAM species *Kalanchöe daigremontiana* and *Ananas comosus* (Bettey and Smith, 1993; McRae *et al.*, 2002). Leaf numbers 3 and 4 (numbered from the centre of the rosette) of *A. americana marginata* were harvested 3 to 4 hours after commencement of the light period, which is Phase III of CAM cycle where maximum rate of decarboxylation occurs (Christopher and Holtum, 1996). Leaf tips, spines and leaf bases were removed. Approximately 100 g fresh weight of leaves, were sliced transversely at 3mm intervals. The sliced fresh tissue was suspended in ice cold extraction buffer made up as outlined in Table 5.1.

Table 5.1 Chemicals used for tonoplast extraction (McRae *et al.*, 2002)

Chemical (Sigma Aldrich,USA)	**MWT**	**STOCK**	**FINAL CONC. in 250 ml buffer**
Mannitol	182.17	1M (45.54g/250ml)	450 mM (112.5 ml)
$MgSo_4$	246.48	0.3M (18.486g/250ml)	3 mM (2.5 ml)
EDTA	368.4	0.2M (7.368g/100ml)	2mM (2.5 ml)
PVP	40000		0.5% (1.25g/250 ml) added to extraction buffer
Tris-base	121.14	1M Tris pH 8 (30.28g/250ml)	100mM (25 ml)
DTT	154.25	0.5M (0.7712g/10ml)	10mM (5 ml)
PMSF	174.19	1M (0.34g/2ml DMSO)	1mM (250μl)
Bovine albumin serum			0.5% (1.25g/250 ml)

In a cold room maintained at 5 ^{o}C, tissue was homogenised with 6-8 repetitions of 3 second bursts in a blender (Coline, model: 18-4518-3, Clas Ohlson). The homogenate was strained through one layer of Miracloth (Calbiochem, San Diego, CA, USA). The homogenate was centrifuged in polycarbonate bottles with aluminium caps (70 ml, Bechman coulter Inc, USA) at 15,000g for 15 minutes at 4^{o}C (Optima LX-100 Ultracentrifuge, Bechman Coulter Inc, USA). Supernatant was collected and centrifuged at 80,000g for 50 min at 4^{o}C. The resulting pellet was suspended in a buffer made up as outlined in Table 5.2.

Table 5. 2 Chemicals used for pellet suspension (glycerol storage medium)

Chemical	MWT	STOCK	FINAL CONC. in 100 ml buffer
Glycerol	92.10	2M (36.84ml/200ml)	1.1M (55 ml)
EDTA	368.4	0.2 M (7.368g/100ml)	1mM (500µl)
Tricine	179.2	0.1 M pH8 by 200mM Tris bis propane(4.48g/250ml)	10mM (10 ml)
Tris-bis-propane	282.33	200mM (5.64g/100ml)	Added to adjust pH 8 of Tricine
DTT	154.25	0.5M (0.7712g/10ml)	2mM (400 µl)

Another method of extraction was tested but first centrifuging at 21,000g for 20 minutes at 4°C and then at 100,00g for 50 minutes at 4°C. Resulting pellets were re-suspended in glycerol storage medium (Table 5.2). Samples were kept in -80°C. Figure 5.2 summarises the main steps used in the two extraction methods used.

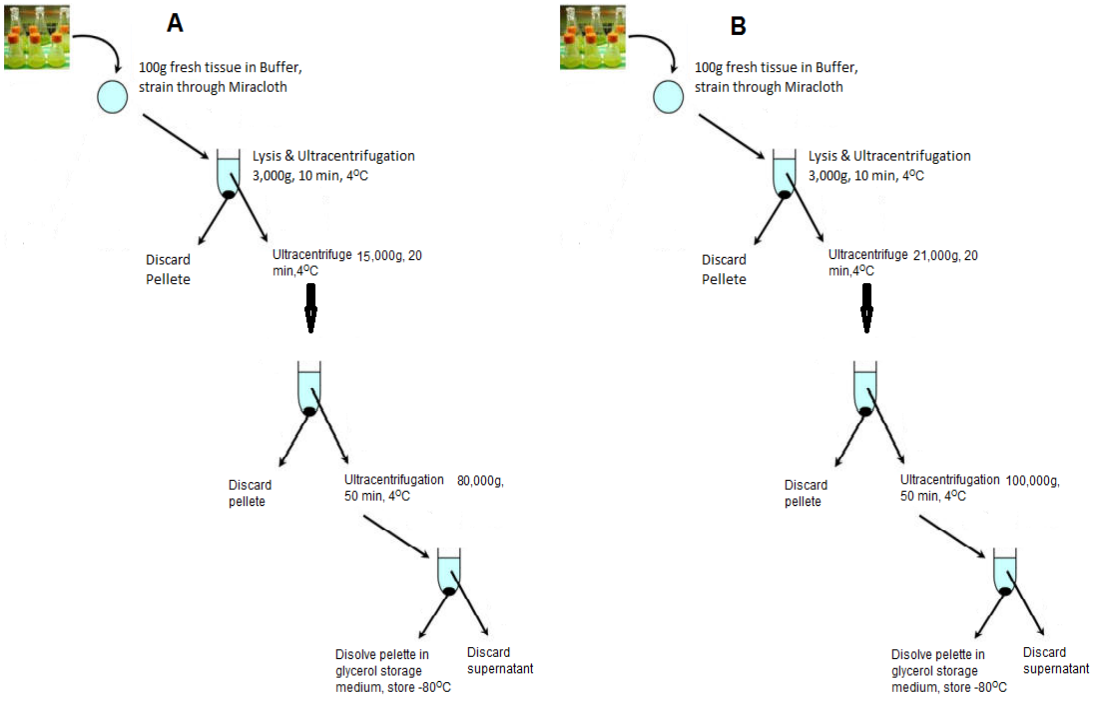


Figure 5.2 Diagram of tonoplast extraction and purification steps. Different ultracentrifuge speeds were tested, as described above. In (A) the extract was spun at 15,000 then finally at 80,000g, and (b) the extract was spun at 21,000g and finally at 100,000g.

5.2.3 Kinetics of ATPase and Pyrophosphatase hydrolytic activity assays

To assess the purity of tonoplast in the membrane vesicle preparation, methods based on those of (Smith *et al.*, 1984; McRae *et al.*, 2002) were used to measure ATPase by colorimetric determination of P_i liberation from ATP. Different protein concentrations of membrane preparation were measured to check that the assays were optimized in terms of substrates. Protein ranged from 2.5μg, 1.5 μg, 1 μg and 0.5 μ. Phosphate standards from 0-40 nmol NAH_2PO_4 were prepared and added to a reaction medium containing chemicals in a total volume of 130 μl, as outlined in table 5.3.The addition of 200 mM Tris-base to Tricine was used to bring pH of the assay mixture to 8.0. Brij-58 detergent made vesicles permeable to substrates.

Table 5. 3 Chemicals used in ATPase Assay

Chemical	Molecular Weight	Stock	Final concentration/130ml
Tricine	179.2	100m M (1.792g/100ml)	50mM
KCL	74.55	500m M (3.72g/100ml)	50mM
$MgSO_4$	246.48	100m M (2.464g/100ml)	3mM
EDTA	368.4	10m M (0.368g/100ml)	0.1mM
Na_2MoO_4	241.95	10m M (0.02419g/10ml)	0.1mM
Brij-58	1122	0.15mg/ml	0.0195mg/ml

The reaction was initiated by the addition of 20 μl of 3 mM ATP and incubated for 10, 20 and 30 minutes at 38^{o}C. The reactions were stopped by the addition of 150 μl of 12% sodium dodecyl sulphate (W/V) and 300 μl of phosphate determining reagent. Phosphate determining reagent was made up in two parts. The first part contained 340 mM D-ascorbic acid and 1 M HCL and the second part contained 30 mM $(NH_4)_6Mo_7O_{24}$. These two were combined in equal volumes immediately before use, each (150μl), with an incubation period of 3 minutes. After incubation, 450 μl of citrate reagent containing 680 mM trisodium citrate, 1.5 mM sodium *meta* arsenite and 2% (v/v) glacial acetic acid was

added, with an incubation for 10 minutes. The absorbance was measured at 850 nm using a spectrophotometer (GENESYS 10 VIS, UK) and cuvettes with pathway length of 1 cm. The same was done for hydrolytic pyrophosphatase (PP_iase) activity. The difference was that 1 mM NA_2MoO_4 was added to the ATPase reaction and reactions were initiated with 20 μl of 500 μM $NaPP_i$ instead of ATP. Control samples with no inhibitors were run along with samples with two different inhibitors; KNO_3 (50mm), an inhibitor of vacuolar ATPases, and in the other samples containing NaN_3+ $Na3VO_4$ both at 100μM, which are inhibitors of mitochondrial and plasma membrane ATPases (McRae *et al.*, 2002)

5.2.4 Protein estimation

Protein contents of the membrane preparations were determined by a colorimetric assay as described by (Bradford, 1976). This is a protein determination method which involves protein binding to Coomassie Brilliant Blue G-250 (Bradford Reagent), causing a shift in absorption maximum of the dye from 465 to 595 nm. Samples were analysed with a spectrophotometer to determine their absorbance at 595 nm (see section 3.2.4.2).

5.2.5 Discontinuous SDS-PAGE gel preparation for protein separation

Proteins were separated by apparent molecular mass, using polyacrylamide gel electrophoresis SDS-PAGE; (Laemmli, 1970) with a vertical Mini-Protean II ™ gel system (Bio-Rad Laboratories Ltd, Hemel Hempstead, Hertfordshire). The multiphasic system employs a separating gel in which samples are fractionated, and a lower percentage stacking gel added above it. A fuller description of reagents used and running conditions are given in sections 3.2.4.3 and 3.2.4.4.

5.2.5 Digestion of proteins from Coomassie-stained gels with trypsin including reduction and alkylation

The protein gel was washed with 70% ethanol for 1 min. A photo copy of the gel was made, marking and labelling bands of interest before cutting them from the gel. The gel was kept hydrated while excising the bands and cut out with a scalpel into smaller pieces, which allows more trypsin to penetrate and increases the yield of peptides which result in a better signal on the MS. Trypsin cleaves on the C-terminal side of arginine and lysine and peptides fragment in a more predictable manner throughout the length of the peptide by putting the basic residues at the C-terminus (Johnson, 2006). Gel pieces were washed 2x

with 200 mM NH_4HCO_3, once with 60% acetonitrile in 200 mM NH_4HCO_3 (30 min incubation with shaking), 50 mM NH_4HCO_3 (30 min incubation with shaking), followed by dehydration with acetonitrile. After this procedure, the gel pieces have shrunk and are white in colour.

Proteins in the gel pieces were reduced with 50 µl of 10mM DTT (AppliChem A1660, 0025) in 100 mM NH_4HCO_3, which resulted in swelling of the bands and clearing in colour. The samples were incubated at 56ºC for 1 hour and then spun in an eppendorf centrifuge at 10,000 rpm for 10 seconds and supernatant was removed by pipetting.

Alkylation of samples was done by adding 50µl of freshly prepared 50 mM iodoacetamide (AppliChem**,** A1101, 0025) in 100 mM NH_4HCO_3, and incubated in the dark for 30 minutes, allowing the lodoacetamide to alkylate all cysteine residues. Gel pieces were pelleted in an eppendorf centrifuge at 10,000 rpm for 10 seconds and supernatant was removed by pipetting, then washed with 200 µl of 100 mM NH_4HCO_3 for 15 mins on a shaker at 1000 rpm, 37ºC, followed by a spin in eppendorf centrifuge for 10,000 rpm for 10 seconds. Supernatant was discarded. Samples were washed with 200 µl of 50 mM NH_4HCO_3/MeCN (50/50 v/v), for 15 mins at 37ºC, 1000 rpm, resulting in shrinkage of samples and turning white. A final spin in eppendorf centrifuge was done at 10,000 rpm for 10 seconds and supernatant was discarded by pipetting.

Dehydration of samples was done by the addition of 70µl of MeCN to dehydrate again for 5minutes at 37ºC, 1000rpm, and placed in eppendorf centrifuge at 10,000rpm for 10 seconds, followed by removing the supernatant by pipetting. Samples were dried under vacuum in a Speedvac (Eppendorf).,

The next step was Trypsin digestion. To a 10 µl aliquot of Trypsin (Promega, Madison, WI,USA) (Shevchenko *et al.*, 1996), the addition of 250µl of 50mM NH_4HCO_3 / 1mM $CaCl_2$. The amount of 30 µl was added to each white, shrunk sample with an incubation of 5 minutes. An addition of 30µl 50mM NH_4HCO_3 to the sample was made and samples were placed in the Thermomixer (Eppendorf) at 37ºC and 1000 rpm with aluminium foil to prevent condensation of the buffer in the top of the tube. A further 30µl of the Trypsin/ 50mM NH_4HCO_3 / 1mM $CaCl_2$ was added to each white, shrunk sample, and left overnight.

After overnight incubation, 10 µl 5% Trifluoroacetic acid (Sigma, USA) was added to samples (to stop the tryptic digest) and samples were left standing at room temperature for 2 minutes. Samples were placed in Eppendorf centrifuge at 10,000 rpm for 10 seconds. The supernatant containing the digested peptides that had been eluted from the gel was transferred to individually labelled 500 µl Eppendorf tubes. Gel pieces were covered with 20 µl of 2% Trifluoroacetic acid/ 60% Acetonitrile, vortexed and left standing for 10 minutes, then spun in an eppendorf centrifuge at 10,000 rpm for10 seconds eluting very hydrophobic peptides. Peptides were transferred to labelled tubes and gel pieces were placed in sonication bath (VWR, PA, USA) until gel pieces were shrunk again.

Acetonitrile (20 µl of 100%) was added to the gel pieces, vortexed and left standing for 5 minutes and placed in an eppendorf centrifuge at 10,000 rpm for 10 seconds, eluting possible remaining peptides. All the peptide containing fractions for one sample were pooled and transferred to newly labelled tubes, then and placed in a Speedvac, drying samples down to remove the acetonitrile.Trifluoroacetic acid (10µl of 1%) was added to the dry residue and tubes were vortexed thoroughly in preparation for mass spectrometer analysis. Samples were transferred into labelled MS vials taking care to avoid transferring gel pieces which can cause damage to the HPLC.

This work was carried out at the Newcastle University Protein and Proteome Analysis facility (NUPPA), Devonshire Building, Devonshire Terrace, Newcastle upon Tyne NE1 7RU, under the supervision of Dr Achim Treumann (Director of NUPPA) and Samantha Baker. The subsequent identification of peptides by LC-MS/MS as detailed below was carried out by NUPPA.

5.2.6 Identification of proteins by Liquid Chromatography-Mass Spectrometry (LC-MS/MS)

5.2.6.1 High Performance Liquid Chromatography

HPLC was performed on a Dionex Ultimate 3000 nano HPLC system (Thermo, Hemel Hampstead, UK), coupled to a Thermo LTQ XL Orbitrap mass spectrometer. The following HPLC conditions were used:

Column: PepMap (Thermo, Hemel Hampstead, UK) column (3 um RP C18 particles, 75 um ID x 250 mm length). Solvents: A, 0.05% formic acid, B 0.05% formic acid in 80% acetonitrile. Samples were loaded onto a PepMap trap column (300 um ID x 10 mm) at a flow rate of 25 ul/min for 3 min. Flow rate: 300 nl/min

Gradient:

Time [min]	%B
0	4
3	4
90	35
102	65
103	95
109	95
109.1	4
120	4

5.2.6.2 Mass Spectrometry

Mass spectrometry is a very effective proteomics tool for identification and quantitation of proteins. The coupling of LC to MS employs ion pair reversed phase chromatography and it also employs nano-HPLC systems with small column diameters which operate at low flow rates giving the advantage of working with small quantities(Mallick and Kuster, 2010). Since both HPLC and

electrospray ionisations (ESI) operate in the liquid phase, no sample collection step is required, avoiding losses.

Precursor spectra were acquired in Orbitrap at a resolution of 60,000. At every time point the 10 most intense precursor ions (excluding singly charged ions) were fragmented in the LTQ linear trap. Normalised collision energy was 35.0, isolation width was 2.0 Da, activation Q was 0.25 and activation time 30 us. Mass accuracy was corrected using the silica ion at m/z 445.120023 as a lockmass (Olsen *et al.*, 2005).

5.2.6.3 Data Processing, Data Analysis and Search Parameters

Raw data were converted into peak-lists in mgf (mascot generic format) using msconvert from the Proteowizard suite (Kessner *et al.*, 2008). The search engine used was X! Tandem Sledgehammer (2013.09.01.2), with a local installation of the global proteome machine (ftp://ftp.thegpm.org/projects/gpm/gpm-xe-installer/). The database searched was the *Agave deserti* proteome (agave_deserti_proteins.fa), downloaded from (http://datadryad.org/resource/doi:10.5061/dryad.h5t68) on July 23rd, 2014 (Westbrook *et al.*, 2011). Annotations (agave_deserti) pfam_interpro_annotations.txt) were downloaded from the same website and associated with identified proteins using a Microsoft Access database. Further annotations were obtained using manual protein blast provided by NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against plant proteins in the uniprot knowledgebase.

Multiple sequence alignment was followed out on V-ATPase V-PPiase and the identified sugar transporters using Clustal Omega (Sievers *et al.*, 2011) from the website (www.ebi.uk/Tools/msa/Clustalo). Clustal Omega is a multiple sequence alignment bioinformatics program, producing biologically meaningful multiple sequence alignment of divergent sequences which are coupled with Cladograms to establish evolutionary relationships. (See Appendix F for alignments).

Fixed modifications were set to carbamidomethyl on C, precursor ion tolerance was set to +/- 10 ppm, product ion tolerance was 0.6 Da, isotope error was set to ‘yes’, refinement was set to ‘yes’, with the following parameters: first round of refinement (deamidation on N,Q, phosphorylation on S,T,Y, oxidation

on M,W, methylation on C,D,E,H,R,K), second round of refinement (methylation on N, Q, dioxidation on M,W, dehydration on S,T, carbamidomethylation on H,D,E,K, lack of carbamidomethylation on C). In an attempt to account for using a not very well annotated database with proteins for a related species, rather than an acknowledged reference proteome, we utilised the option of allowing for single amino acid polymorphisms at the refinement stage of the X!Tandem search. Figure 5.3 summarises the mass spectrometry/proteomic experiment.

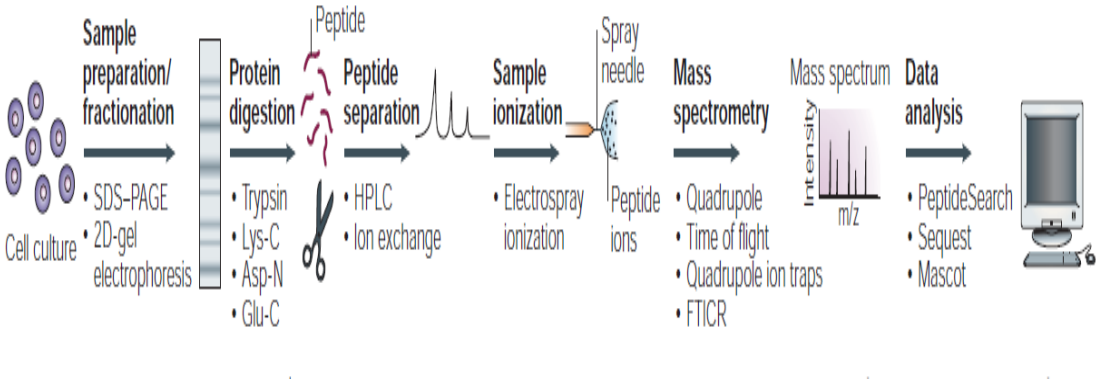


Figure 5.3 Pipeline of Mass-spectrometry/proteomic experiment. Protein extracted from *Agave*, purified by SDS-PAGE. Desired gel lanes are excised and cut in several slices, and digested. Finally, the peptide sequencing data were obtained from the mass spectra and searched against protein databases using a number of database searching programs. Scheme adopted from (Steen and Mann, 2004)

5.3 Results

In this study, a method was developed to generate a protein fraction from *A. americana marginata* that was enriched in tonoplast proteins. This fraction was characterised using biochemical and proteomic approaches.

The relative proportions of vacuolar, mitochondrial and plasma membranes in the isolated membrane preparations were estimated by measuring the inhibition kinetics of ATPase in vesicle preparations. Vacuolar ATPases (V-ATPases) are sensitive to inhibition by potassium nitrate (KNO_3) as low as millimolar

concentrations but are insensitive to inhibition by either sodium orthovanadate (Na_3VO_4) or sodium azide (NaN_3) (Wang and Sze, 1985). In contrast, plasma membrane and mitochondrial ATPases are insensitive to KNO_3 but show a sensitivity at micro-molar concentrations to Na_3VO_4 and NaN_3 (Gallagher and Leonard, 1982; Wang and Sze, 1985).

5.3.1 Inhibition kinetics of ATPase

Four different protein concentrations of membrane prepared from leaves of *A. americana marginata* tonoplast-enriched preparations (ranging between 0.5 – 2.5 μg/300μl) were tested to find the optimal assay conditions for demonstrating the kinetics of ATPase activity. This was determined before adding known ATPase inhibitors. A protein loading of 1.5μg was found to give the most consistent results when assayed (i.e. there was no substrate limitation and reaction was linear for up to 30 mins as illustrated below). ATPase activity was measured in nano katal (nkat mg^{-1} protein) (Fig. 5.4)

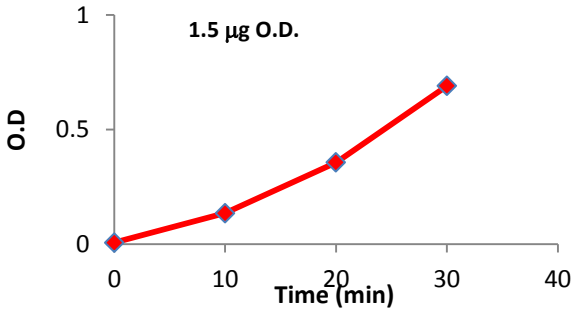


Figure 5.4 Linear activity of ATPase measured as a change in optical density (OD) at 850 nm wavelength for 1.5μg membrane protein extracted from leaves of*A. americana marginata*. At this protein input, ATPase activity was linear for up to 30 minutes, with calculated ATPase activity of 20.1 nkat mg^{-1}protein.

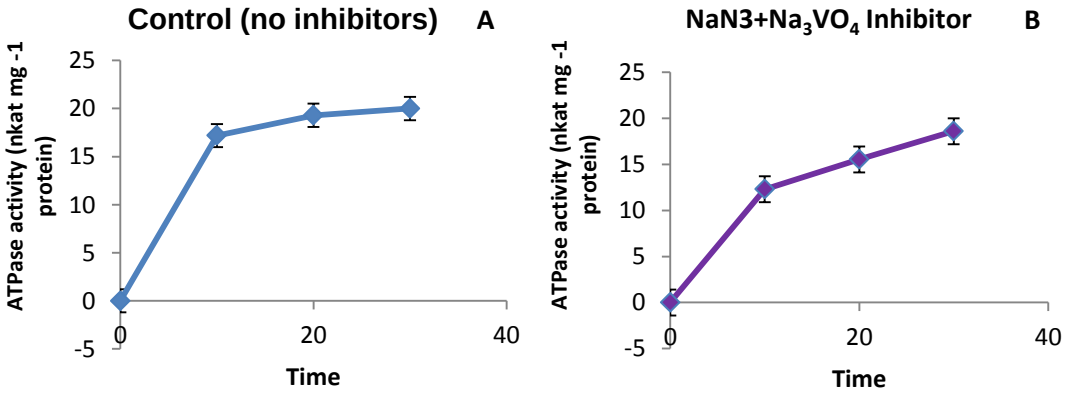
The proportion of ATPase activity in the membrane protein extract that could be attributed to vacuolar, i.e. V-ATPase activity was ~ 91.5%, if estimated as KNO_3 sensitive activity, and ~ 93%, if estimated as NaN_3 and $NaVO_4$ insensitive activity (Table 5.4).

Table 5. 4 ATPase and PP_iase of *Agave Americana marginata* leaf vesicles and the sensitivity of ATPase activity to inhibition by known ATPase inhibitors

Treatment	ATPase activity (nkat mg^{-1}protein)	Inhibition (%)	*PP_iase activity (nkat mg^{-1}protein)
Control(no inhibitors)	20.1 ± 1.2	-	5.9
KNO_3 (50mM)	1.7 ± 0.49	91.5	-
NaN_{3+} $NaVO_4$	18.8 ± 1.42	7	-

Rates are sums of activities of inside and outside facing ATPases, the assays included the detergent Brig-58 which makes vesicles permeable to substrates. ATPase values represent the mean ± S.E (n=3).
*Only one sample for PPiase activity was measured.

The vesicle membrane preparations exhibited features expected for a fraction highly enriched in tonoplast membrane with ATPase activity of 20.1 ± 1.2 nkat protein which was inhibited 91.5% by 50 mM KNO_3, an inhibitor of vacuolar ATPase, but was only 7% inhibited by 100 μM NaN_3 and 100μM Na_3VO_4, inhibitors of mitochondrial and plasma membrane ATPases respectively (Gallagher and Leonard, 1982; Wang and Sze, 1985). Vesicles exhibited a kinetic gradient that was maintained for up to 30 minutes. Figure 5.5 shows that inhibition increased with incubation time.



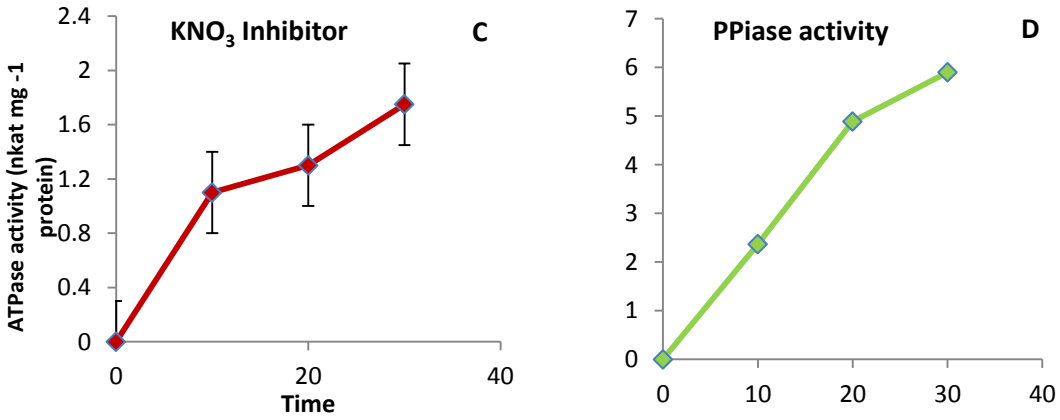


Figure 5.5 Time course kinetics of ATPase activity in membrane vesicles prepared from leaves of *A. americana* marginata with and without addition of known ATPase inhibitors; (A) control contains no inhibitors. (B). ATPase activity inhibition by NaN_3+Na_3VO_4. (C). ATPase activity inhibition by KNO_3. and (D) PPiase activity. ATPase activity measurements represent the mean ± S.E (n=3). PP_iase (n=1).

5.3.2 Protein fractionation by discontinuous SDS-PAGE analysis

Proteins which made up the isolated tonoplast-enriched membrane fraction from *A. americana marginata* were separated using SDS-PAGE gels. Exactly 15 μg of membrane protein was loaded and separated on a 12% acrylamide gel stained with Coomassie Blue ® G-250 (Biorad, USA; Fig. 5.6). Different membrane fraction preparations obtained from different centrifugation speeds (15,000, 21,000, 80,000 and 100,00g) was compared. At 15,000 and 21,000g, the extract was predicted to contain mitochondria, chloroplasts and nuclei. At 80,000 g and 100,000 g, the samples should be comprised predominantly of tonoplast membrane. Samples were run on the gel and bands of interest were cut out to check for the presence of tonoplast proteins by LC-MS/MS analysis. Major bands from SDS-PAGE migration (lane 4) were cut out, between 55 and 40 kDa. This led to the identification of lane 4 as a membrane fraction that was enriched for tonoplast proteins. Following this preliminary experiment, the remainder of the lane was sliced into 5 additional bands and each of these bands was subjected to in gel trypsin digestion followed by LCMSMS analysis.

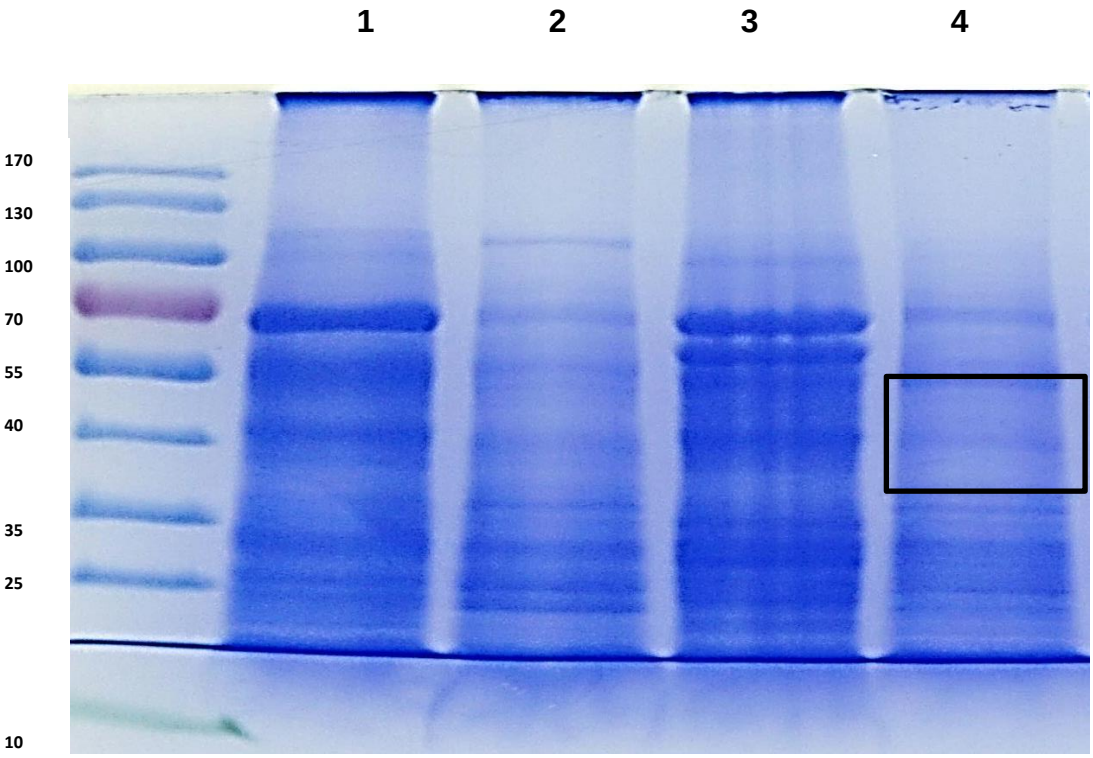
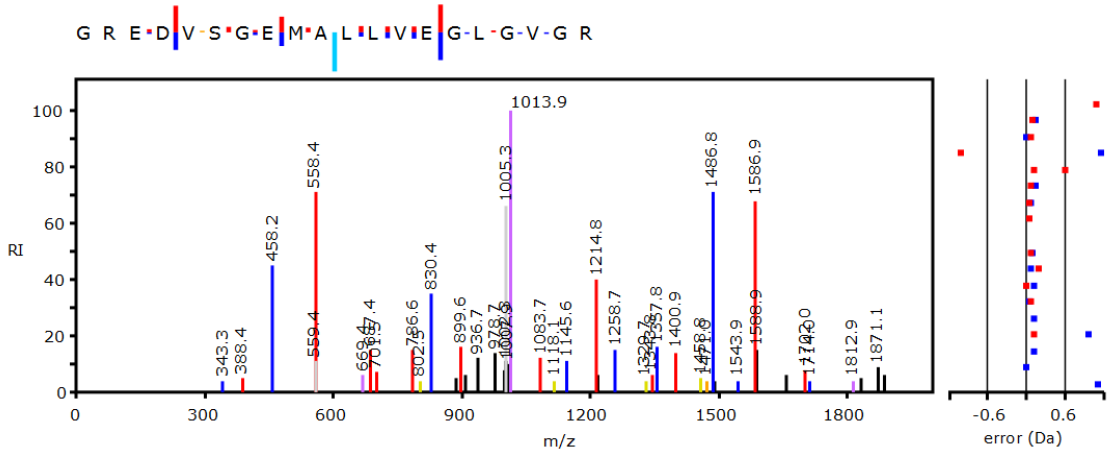


Figure 5.6 SDS-PAGE gel showing separation of proteins obtained as a result of different centrifugation speeds. Numbers on the left represent the size of the molecular mass markers in kDa. Lane 1: Proteins from the spin at 15,000 g. Lane 2: Proteins from the spin at 21,000 g. Lane 3: Proteins from the spin at 80,000 g. Lane 4: Proteins from the spin at 100,000 g. The black frame in lane 4 shows the bands which were excised (from 55 to 40 kDa) for subsequent LC-MS/MS identification. The remainder of Lane 4 was sliced into 5 additional bands and each of these bands was subjected to in gel trypsin digestion followed by LCMSMS analysis.

5.3.3 LC-MS/MS ANALYSIS for peptide identification

The analysis of one lane (lane 4; Fig. 5.6) of an SDS-PAGE gel containing a tonoplast-enriched protein fraction yielded a total of 1296 protein identification events (8657 peptides at a peptide level false positive rate of less than 1%) (Gupta *et al.*, 2011) from 6 SDS-PAGE gel bands. Due to the identification of many products from several gene loci, this corresponds to 934 gene products that were identified in this sample. It was encouraging to observe that subunits of vacuolar ATPases were amongst the most confidently identified proteins in the sample, detected in relatively high abundance (as judged by spectral counts), confirming that we are dealing with a tonoplast-enriched fraction. The presence of heat shock proteins and PEP carboxylase shows that

this was not a completely pure tonoplast preparation, but this was not unexpected. See Figure 5.8 for predicted molecular weight of identified gene products in *Agave americana,* which was a result from 6 bands cult from the SDS-PAGE gel).



matched/total:		# ions: 68%		intensity: 85%		μ: 0.20, σ: 0.42 Da
bond	^{+1}y	$^{+1}y^{-17}$	$^{+1}y^{-18}$	^{+1}b	$^{+1}b^{-17}$	$^{+1}b^{-18}$
$^{G}1$	1987.028	1970.001	1969.017	58.029	41.002	40.018
$^{R}2$	1830.927	1813.900	1812.916	214.130	197.103	196.119
$^{E}3$	1701.884	1684.857	1683.873	343.172	326.146	325.162
$^{D}4$	1586.857	1569.831	1568.846	458.199	441.173	440.189
$^{V}5$	1487.789	1470.762	1469.778	557.268	540.241	539.257
$^{S}6$	$^{+1,2}$1400.757	1383.730	1382.746	644.300	627.273	626.289
$^{G}7$	1343.735	1326.709	1325.725	701.321	684.295	683.311
$^{E}8$	1214.693	1197.666	1196.682	830.364	813.337	812.353
$^{M}9$	1083.652	1066.626	1065.642	961.404	944.378	943.394
$^{A}10$	1012.615	995.588	994.604	1032.441	1015.415	1014.431
$^{L}11$	899.531	882.504	881.520	1145.526	1128.499	1127.515
$^{L}12$	786.447	769.420	768.436	1258.610	1241.583	1240.599
$^{V}13$	687.378	670.352	669.368	1357.678	1340.651	1339.667
$^{E}14$	558.336	541.309	540.325	1486.721	1469.694	1468.710
$^{G}15$	501.314	484.288	483.304	1543.742	1526.716	1525.732
$^{L}16$	388.230	371.204	370.220	1656.826	1639.800	1638.816
$^{G}17$	331.209	314.182	313.198	1713.848	1696.821	1695.837
$^{V}18$	232.140	215.114	214.130	1812.916	1795.889	1794.905
$^{G}19$	175.119	158.092	157.108	1869.937	1852.911	1851.927

Figure 5.7 MS/MS spectrum for peptides of interest (sugar transporter protein, Locus6095v1rpkm49.88_8). Peptide sequence is shown at the top of each spectrum, as well on the left under (bond), with the annotation of the identified matched amino terminus-containing ions (b ions) and the carboxyl terminus-containing ions (y ions) (Roepstorff and Fohlman, 1984). For clarity, only major identified peaks are labelled. m/z on x-axis, mass to charge ratio, and RI on y-axis, Relative Intensity.

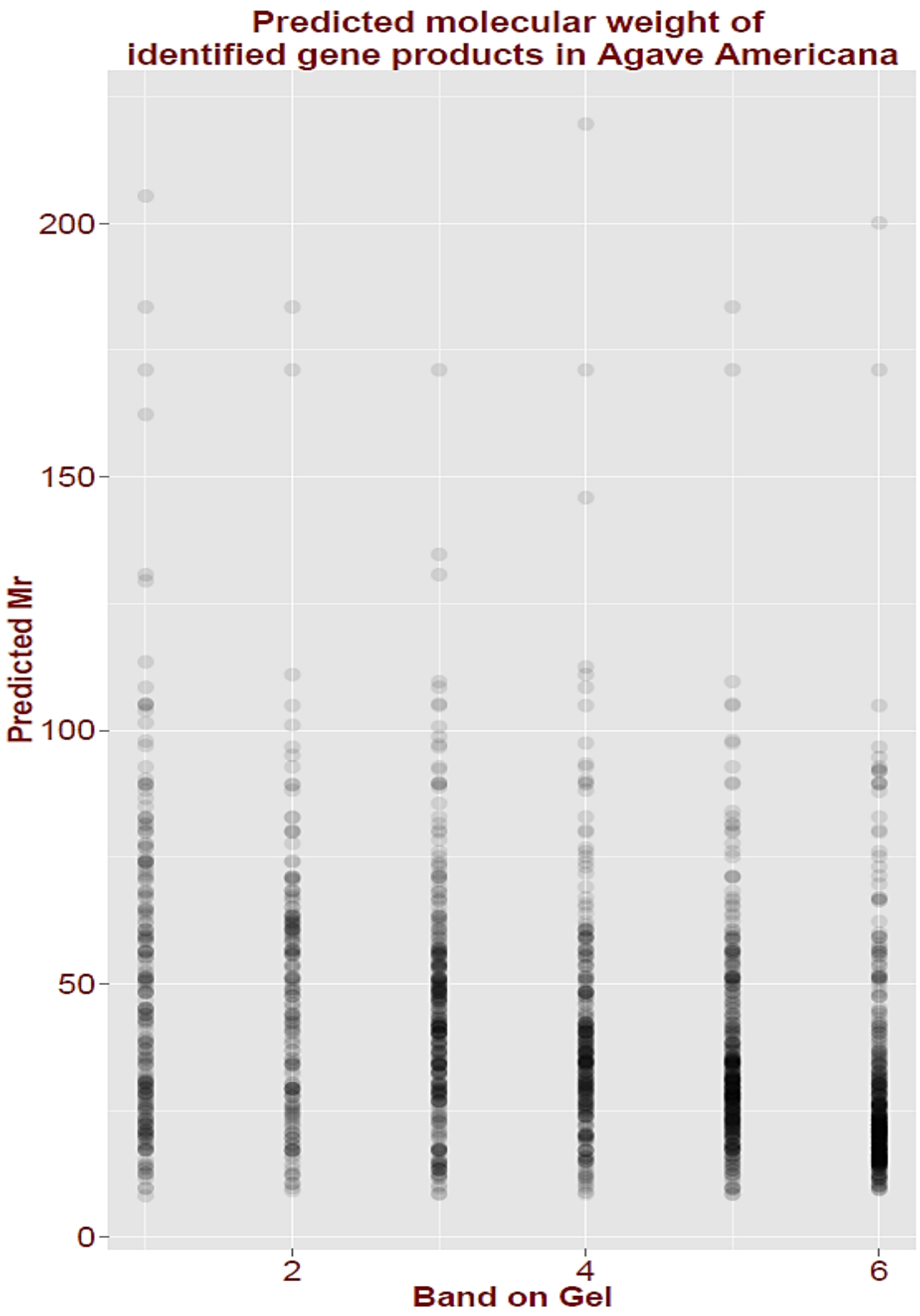


Figure 5.8 Scatter plot of predicted molecular weight for each identified gene product in *Agave americana*. With higher molecular weights in band 1 of the gel, and the lower molecular weights in band 6 of the gel

5.3.4 Sugar transporter annotations in proteomics results

Manual searching of the pfam annotation column and the interpro description column in the list of protein identification events was conducted to identify proteins that could be linked to carbohydrate biosynthesis or metabolism. This approach identified 36 proteins that were annotated as being linked to saccharide biosynthesis, carbohydrate metabolism or transport. Six of these protein identification events stood out with the pfam annotation “Sugar_tr” – and upon closer investigation it turned out that these corresponded to 5 different gene loci and to 4 different proteins

Overall, the proteomics analysis identified 934 protein events (see Appendix E). Out of those, 4 proteins were identified as containing a sugar transporter domain. The sugar transporter proteins identified are as follows:

1. Identifier: Locus20314v1rpkm10.75_5
   Log(e): -13.8
   E Value: 3.80E-40
   Protein length: 328
   Pfam description: Sugar_tr
   Interpro description: General substrate transporter
   GO terms: Cellular Component: integral to membrane (GO:0016021), Molecular Function: transmembrane transporter activity (GO:0022857), Biological Process: transmembrane transport (GO:0055085)
   Sequence:

MGIGGGWQLAWKWSERDGADGTKEGGFKRIYLHPEGVAGSQRGSIVSLPGAGVQG
SEVFQAVALVSQPAVYSKELMEQHPIGPAMLHPLETASKGPRWGDIFDAGVKHALFV
GIGIQILQQFAGINGVLYYTPQILEQAGVGVLLSNIGISSDSTSILISVLTTLLMLPSIGVA
MRLMDISGRRSLLLATIPVLIVTLVILVIANLVNLGSVLHAVLSTISVIVYFCFFVMGFGPI
PNILCAEIFPTHVRGICIAICALTGWIGDIIVTYTLPLMLSSIGLAGVFGIYAIVCIVSLLFVF
LKVPETKGMPLEVITEFFAIGAKQAAGN

   A. thaliana annotation: At4g35300

This protein was similar to the Tonoplast Monosaccharide Transporter 2 (TMT2) (http://thbiogrid.org/14966).

2. Identifier: Locus6095v1rpkm49.88_8
   Log(e): -53.3
   E Value: 1.10E-50
   Protein length: 523
   Pfam description: Sugar_tr

   Interpro description: General substrate transporter

   GO terms: Cellular Component: integral to membrane (GO:0016021), Molecular Function: transmembrane transporter activity (GO:0022857), Biological Process: transmembrane transport (GO:0055085)

   Sequence:

MGAVLIAIAAAIGNLLQGWDNATIAGSVLYIKKEFNLESEPAIEGLIVAMSLIGAT
VITTFSGAISDAFGRRPMLIVSSLLYFLSGIVMFCSPNIYVLLLARLIDGLGIGLSV
TLVPMYISETAPSDIRGLLNTLPQFTGSCGMFLSYCMVFGMSLRVKPDWRLML
GVLSIPSLLYFALTIFYLPESPRWLVSKGRMIEAKHVLQRLRGREDVSGEMALL
VEGLGVGRETSIEEYIIGPADELPDEEDPTAESEKIMLYGPEAGQSWVAQPVK
GHSVLGSALGVVSRQGSTANRNIPLMDPLVTLFGSVHEKAPEIGGSMRSILFP
NFGSMFSAAGQQSRSEQQWDEEIIQREGEDYVSDAERSDSDDNLQSPLLSR
QTTSMEGKDMVPPPSNGGTLGMRRVSLMLGTSGEAVSSMGIGGGWQLAWK
WSERDGADGTKGGFKRIYLHPEGVPGLQRGSTVSLPGADVQGSEVIRAAALV
SRPAFYSKELMEQHPVGPAMVHPLETASKGPRWGDLFDAGVQHA

   A. thaliana annotation: At3G51490

This protein was also similar to Tonoplast Monosaccharide Transporter 2 (TMT2) in *A.thaliana*.

3. Identifier: Locus7701v1rpkm38.97_6
   Log(e): -3.3
   E Value: 2.50E-66
   Protein length: 400
   Pfam description: Sugar_tr

   Interpro description: General substrate transporter

   GO terms: Cellular Component: integral to membrane (GO:0016021), Molecular Function: transmembrane transporter activity (GO:0022857), Biological Process: transmembrane transport (GO:0055085)

   Sequence:

MSFRGDESGGEDGGLRKPFLHTGSWYRMGMGSRQSSLMDKSSSGSVIRDS
SVSVVLCTLIVALGPIQFGFTGGYSSPTQDAIIKDLGLSISEFSIFGSLSNVGAMV
GAIASGQIAEYIGRKGSLMIASIPNIIGWLAISFAKDSSFLYMGRLLEGFGVGVIS

YTVPVYIAEIAPQNMRGGLGSVNQLSVTIGIMLAYIFGMFLPWRLLAVMGVLPC
TVLIPGLFFIPESPRWLAKMGMMEDFEASLQVLRGFDTDISVEVNEIKRSVASG
TRRTTIRFSDLKQRRYKLPLMIGIGLLVLQQLSGINGILFYANNIFKAAGVSSSA
GATCGLGAIQVIATGFTTWLLDRAGRRLFLIISSAGMTASLLLVAIVFYLKGVITE
DSKFYFILGVLSLVGLVAY

A. thaliana annotation: At1G75220

This protein was similar to Sugar transporter, Early Response to Dehydration (ERD6-like 6) in *A. thaliana.*

4. Identifier: Locus834v1rpkm277.18_5
   Log(e):-19.6
   E Value: 1.70E-48
   Protein length:506
   Pfam description: Sugar_tr
   Interpro description: General substrate transporter
   GO terms: Cellular Component: integral to membrane (GO:0016021),
   Molecular Function: transmembrane transporter activity (GO:0022857),
   Biological Process: transmembrane transport (GO:0055085)
   Sequence:

MGFFTDAYDLFCISLVTKLLGRIYYHVDGSETPGVLPPNVSAAVNGVAFCGTLL
GQLFFGWLGDKMGRKRVYGMTLMLMVICSVASGLSFGHKAKGVMATLCFFR
FWLGFGIGGDYPLSATIMSEYANKKTRGAFIAAVFAMQGFGILTGGAVALIVSA
AFKNEFKAPTYEQNAVASTVPEADYVWRIILMFGALPAAMTYYWRMKMPETA
RYTALVAKNAKQAAADMSKVLQVEIEAEQEKVEKIATSEANTFGLFTKEFAKR
HGLHLLGTTTTWFLLDIAFYSQNLFQKDIFSAIGWIPKAKTMNAIEEVFRIARAQ
TLIALCGTVPGYWFTVGLIDVIGRFTIQMMGFFFMTVFMLGLAIPYHHWTLKGN
HIGFVVMYAFTFFFANFGPNSTTFIVPAEIFPARLRSTCHGISAAAGKAGAIIGSF
GFLYAAQNQDKAKADHGYPAGIGVRNSLFVLAGCNLLGLFFTLLVPESNGKSL
EEMSRENEDEEQAGGNPNSRTVPV

A. thaliana annotation: At3G54700

This protein was similar to an inorganic phosphate transporter 1-7 in *A.thaliana.* This is another 12 TMT protein, also a part of the Major Facilitator Superfamily (MFS), which includes sugar transporters.

5. Identifier: Locus3753v1rpkm79.87_8

This protein is highly homologuous to the Monosaccharide-sensing protein 2 (or 3) in *Arabidopsis thaliana* (http://www.uniprot.org/uniprot/Q8LPQ8), a 12

transmembrane domain sugar transporter that has been localised in a proteomic study to the vacuolar membrane(Jaquinod *et al.*, 2007) .

Log(e): -43.7
E Value: 2.60E-42
Protein length: 598
Pfam description: Sugar_tr
Interpro description: General substrate transporter
GO terms: Cellular Component: integral to membrane (GO:0016021), Molecular Function: transmembrane transporter activity (GO:0022857), Biological Process: transmembrane transport (GO:0055085)
Sequence:

MFLSYCMVFSMSLLPQPNWRLMLGVLSIPSLLYFALTIFYLPESPRWLVSKGR
MTEAKKVLQRLRGREDVAGEMALLVEGLGVGGETSIEEYIIGPANDLNDEHAP
AADKEQITLYGPEEGQSWIARPAKGQSMLGSALGIISRHGSMENQGSIPLMDP
LVTLFGSVHENLPQSGSMRNSMFPNFGSMFSFAADQHPKTEQWDEEHGQR
EGDGYASDSTGGDSDDNLHSPLLSRQTTSIEGKDIAPHGTHGSTLNMGRNSS
LLQGTSGDAMGIGGGWQLAWKWSERDGADGKKEGGFKRIYLHEGVPSSHR
GSLVSLPGGDVPEETEYVQAAALVSQPALYSKELMNQHPVGPAMVHPSEEAA
KGPRWTDLLEPGVRHALVVGIGIQILQQFSGINGVLYYTPQILEQAGVGILLSNL
GISSTSASILISGLVTLLMLPSIGIAMKFMDVAGRRSLLLSTIPVLILTLVILVLSNV
MDFGQVAHAVLSTISVIVYFCCFVMGFGPIPNILCSEIFPTRVRGVCIAICALTF
WIGDIIVTYTLPVMLDSIGLAGVFGIYAVVCIISLVFVFLKVPETKGMPLEVITEFF
AVGARQPGRT

5.3.5 Multiple sequence alignments for V-ATPase, V-PP$_i$ase and sugar transporters in *A.americana*

Spectra from LC/MS/MS for V-ATPase and V-PP$_i$ase were compared. Multiple sequence alignment uncovered redundancy which is genome loci that are listed more than once. For the V-ATPase, 8 different loci were found and can be seen in Table 5.5. Also 7 different loci were found for V-PP$_i$ase (Table 5.6). Total peptides indicate the abundance of V-ATPases are much higher than those of V-PPiases in *A. americana*. The clustal alignment for both proteins shows that they are most likely to correspond only to one gene each (see cladograms in Figure 5.9).

Table 5.5 V-ATPase loci found in *A. americana* tonoplast

No	Band	Identifier	rI*	Mwt (kDa)	Total peps	Pfam description	Interpro description
1	1	Locus706v1rpkm311.27_14	40	92.8	153	V_ATPase_I	ATPase, V0/A0 complex, 116kDa subunit
2	3	Locus15040v1rpkm17.17_2	8	14.4	18	V-ATPase_C	ATPase, V1 complex, subunit C
3	5	Locus3216v1rpkm92.68_6	32	41.4	102	V-ATPase_C	ATPase, V1 complex, subunit C
4	5	Locus8278v1rpkm36.22_7	22	42.3	69	V-ATPase_C	ATPase, V1 complex, subunit C
5	6	Locus18798v1rpkm12.27_12	16	51.2	35	V-ATPase_H_C	ATPase, V1 complex, subunit H, C-terminal
6	5	Locus10055v1rpkm28.77_6	13	32.6	42	V-ATPase_H_C	ATPase, V1 complex, subunit H, C-terminal
7	3	Locus4457v1rpkm67.94_2	10	11.8	26	V-ATPase_H_C	ATPase, V1 complex, subunit H, C-terminal
8	6	Locus2992v1rpkm99.31_8	34	38.7	88	V-ATPase_H_N	ATPase, V1 complex, subunit H

Table 5.6 V-PP$_i$ase loci found in *A. americana* tonoplast

No	Band	Identifier	rI	Mwt (kDa)	Total peps	Pfam description	Interpro description
1	1	Locus18589v1rpkm12.49_1	4	12.7	4	H_PPase	Pyrophosphate-energised proton pump
2	1	Locus195v1rpkm706.72_5	27	44	51	H_PPase	Pyrophosphate-energised proton pump
3	6	Locus106v1rpkm921.90_2	30	15.2	42	H_PPase	Pyrophosphate-energised proton pump
4	6	Locus2238v1rpkm129.65_6	29	40.9	56	H_PPase	Pyrophosphate-energised proton pump
5	1	Locus2512v1rpkm117.37_7	20	56.3	30	H_PPase	Pyrophosphate-energised proton pump
6	1	Locus3621v1rpkm82.92_6	10	43.2	10	H_PPase	Pyrophosphate-energised proton pump
7	6	Locus847v1rpkm274.83_10	29	79.9	88	H_PPase	Pyrophosphate-energised proton pump

*: Number of peptides found

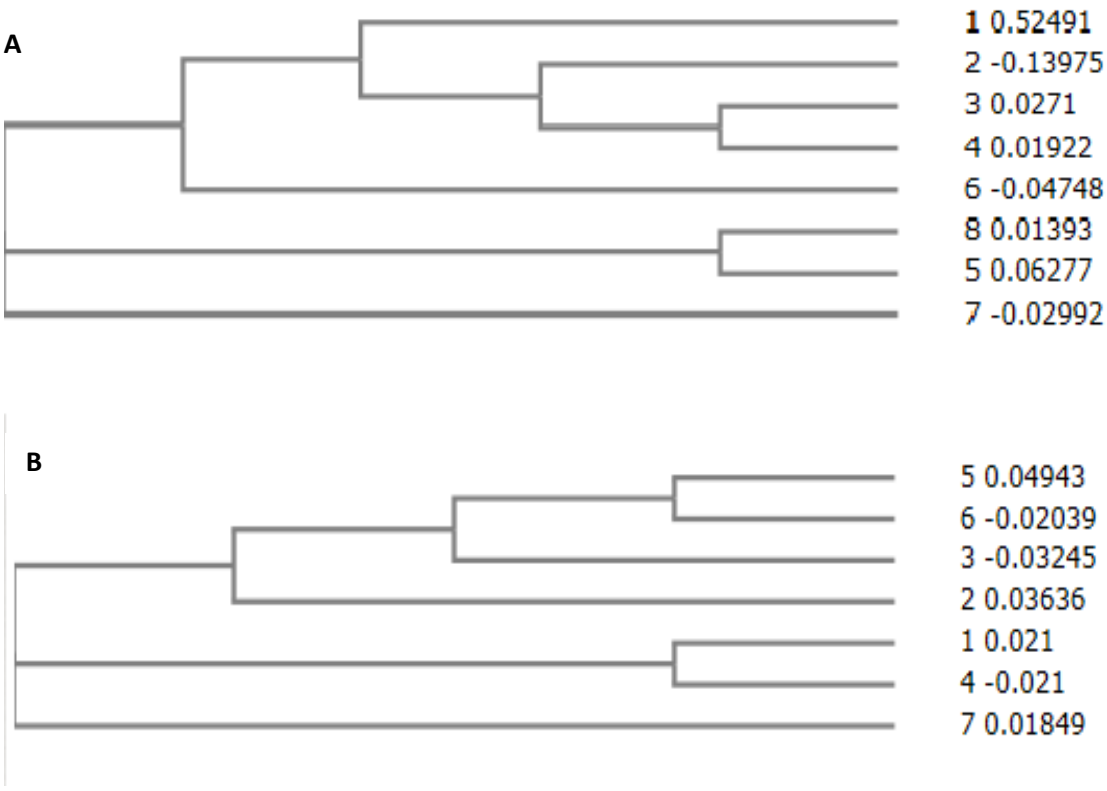


Figure 5.9 Cladograms for the multiple sequence alignment for (A) V-ATPase and (B) V-PPiase loci from tonoplast of *A. americana.* The numbers on the right correspond to the different loci in Table 5.5 and Table 5.6

For V-ATPase, similarities were found between Locus3216v1rpkm92.68_6 and Locus8278v1rpkm36.22_7, and between Locus2992v1rpkm99.31_8 and Locus18798v1rpkm12.27_12, with Locus4457v1rpkm67.94_2 being far related from the rest of the identifiers for V-ATPase.

For V-PP$_i$ase, similarities are shown in loci Locus2512v1rpkm117.37_7 and Locus3621v1rpkm82.92_6 and between Locus18589v1rpkm12.49_1 and Locus2238v1rpkm129.65_6, with loci Locus847v1rpkm274.83_10 distantly related to the rest.

Identified sugar transporters were the least abundant proteins when compared with the two vacuolar pumps, V-ATPase and V-PP$_i$ase. Five different loci for putative sugar transporters were identified from *A. americana* tonoplast as mentioned previously (see Appendix F for multiple sequence alignments).

Cladogram for identified sugar transporters was also constructed. Close similarities are shown between Locus20314v1rpkm10.75_5 and Locus7701v1rpkm38.97_6, and Locus3753v1rpkm79.87_8 been the farthest related to the other sugar transporters. See Figure 5.10

Figure 5.10 Cladogram for loci of 5 different sugar transporters.

2.3.6 Interrogation of transcript and proteome data related to identified sugar transporters in *A. americana*

The putative sugar transporters identified from the tonoplast membrane prepared from leaves of *A. americana marginata* were used to search Transcriptome and Protein databases for *A. americana* (Biosciences Research group at the Oakridge National Laboratory in Tennessee). These data bases contain information relating to transcript and protein abundances from mature leaves of *A. americana marginata* sampled at 4 hour intervals over a 24 light/dark cycle. The transcriptome data base also contains information pertaining to global transcript abundances in young, C3 leaves and other plant tissues (e.g. meristem, stem, root, rhizome). Three out of the 4 sugar transporters were identified in the *A. americana* transcriptome database. Transcript sequences producing significant alignments were selected (see Table 5.7 for score (bits) and E values of chosen sequences). For the first TMT2 (*A. thaliana* annotation: *At4g35300*) transcript and proteome abundances are shown in Figure 5.8 for mature leaves. The transcript abundances for young leaves at different time intervals and different tissues are shown in Figure 5.12.

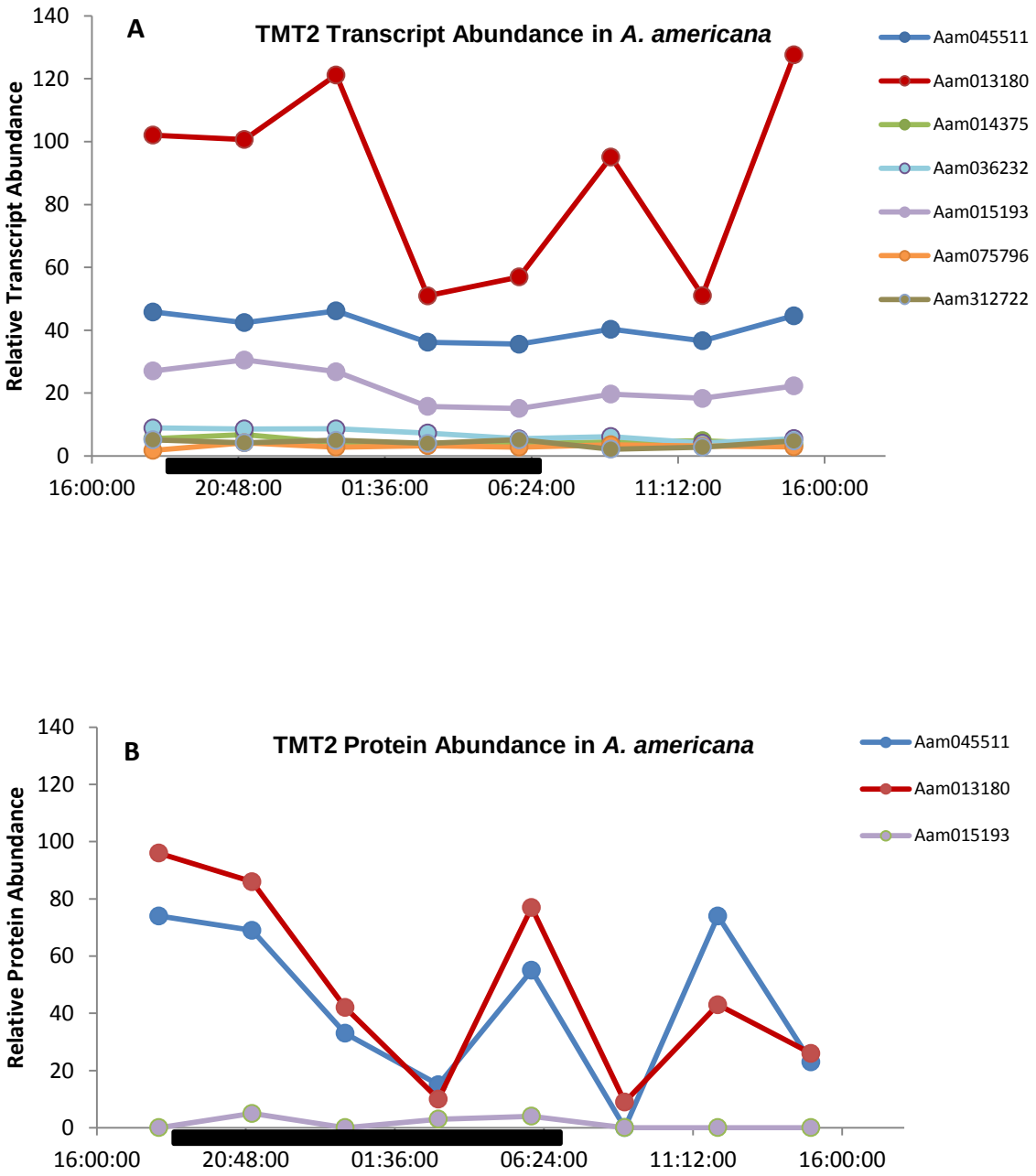


Figure 5.11 Transcript (A) and protein (B) abundance profiles for TMT2 over 24 h time course in mature leaves of *A. americana*, showing the highest abundance for *Aam 013180* sequence for both transcriptome and protein profiles. Dark period was between 6:30pm to 06:30 am indicated by black bar on x-axis.

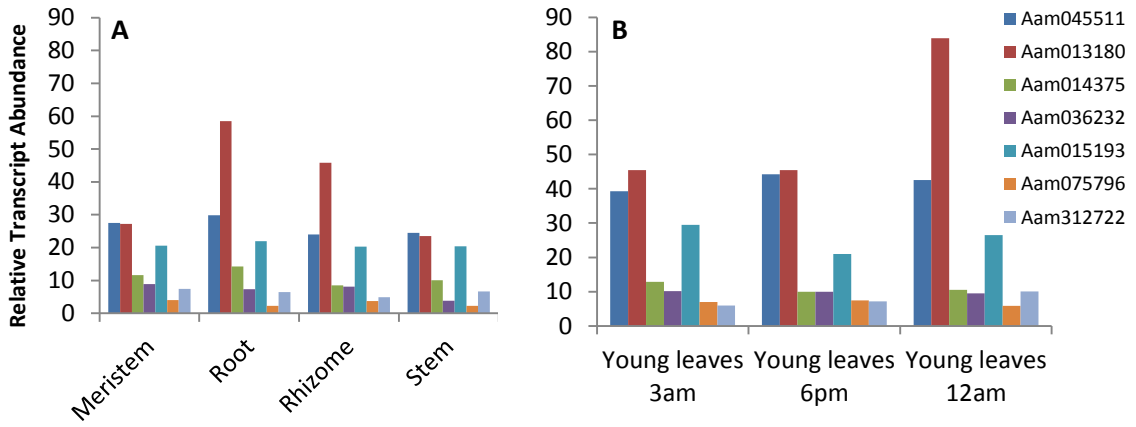


Figure 5.12 Distribution of TMT2 (*A. thaliana* annotation: *At4g35300*) transcript abundance in different tissue of *A. americana* (A) and in young leaves (B). Both display high abundance of Aam *013180* sequence in root and young leaves (12am).

Some 7 transcript sequences were found to correspond to TMT2. Transcript sequence *Aam013180* showed the highest abundance in mature leaves and peaked at times 12am and 3pm in the diel cycle. This transcript also peaked at 12 am in the young (C3 leaves). Sequence *Aam 013180* also had the highest protein abundance in mature leaves although diel patterns of transcript and protein abundance did not exactly mirror each other (Fig. 5.11). In general sequence Aam013180 had the highest transcript abundance compared to the other TMT2-like transcripts in roots and rhizomes. Whilst transcript abundance of *Aam 01318* was generally higher in leaves compared on the other tissues, its existence in roots and rhizomes suggests that it does not appear to have a CAM-specific function.

For the ERD6-like protein identified from the *Agave* tonoplast preparation, (*A. thaliana* annotation: *At1G75220*) 5 transcripts with sequence similarity were identified from the *A. americana marginata* transcriptome database (Fig. 5.13). *Aam081118* showed the highest transcript abundance and showed higher expression in mature and young leaves compared to other tissues (e.g. roots, meristems, rhizomes, stems). *Aam 12894* was more abundant and the highest of the ERD6-like proteins in root tissue, Figure 5.10. No sequence match for ERD6- like was found in the proteome database of *A. americana* implying that it is a low-abundance protein.

Transcript and protein abundance for the TMT inorganic phosphate transporter 1-7 (*A. thaliana* annotation: At3G54700) is shown in Figure 5.14. Some 11 transcripts were found to show similarity to the TMT inorganic P transporter. *Aam 013446* was the transcript in highest abundance in mature and young leaves and this transcript encoded the protein with highest abundance for the TMT inorganic P translocator in mature leaves (Fig. 5.14). Transcript abundance of *Aam013446* in roots, meristems, stems and rhizoids was much lower than that in young leaves and particularly mature leaves. Thus, this protein may well have a CAM-specific function.

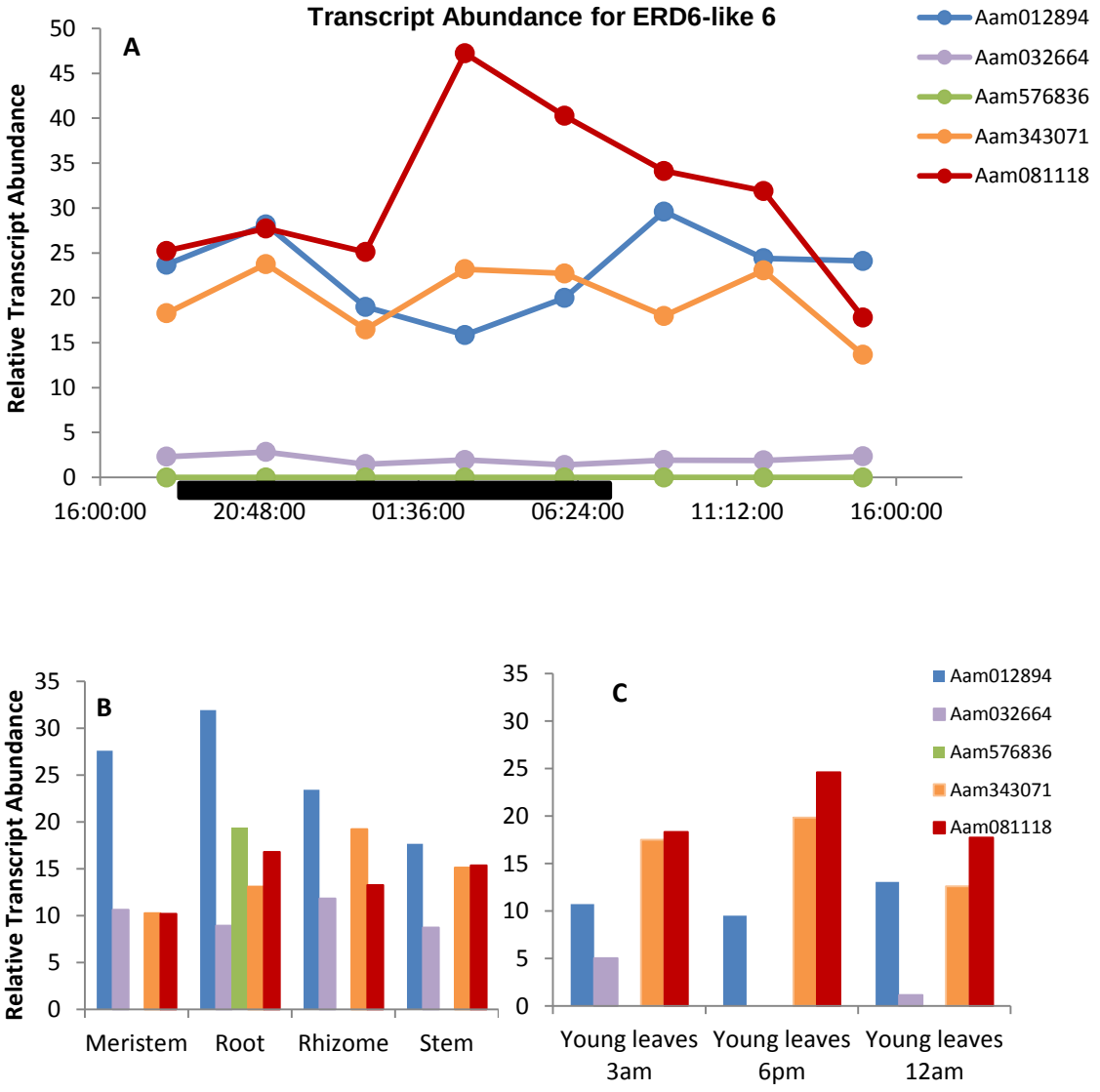
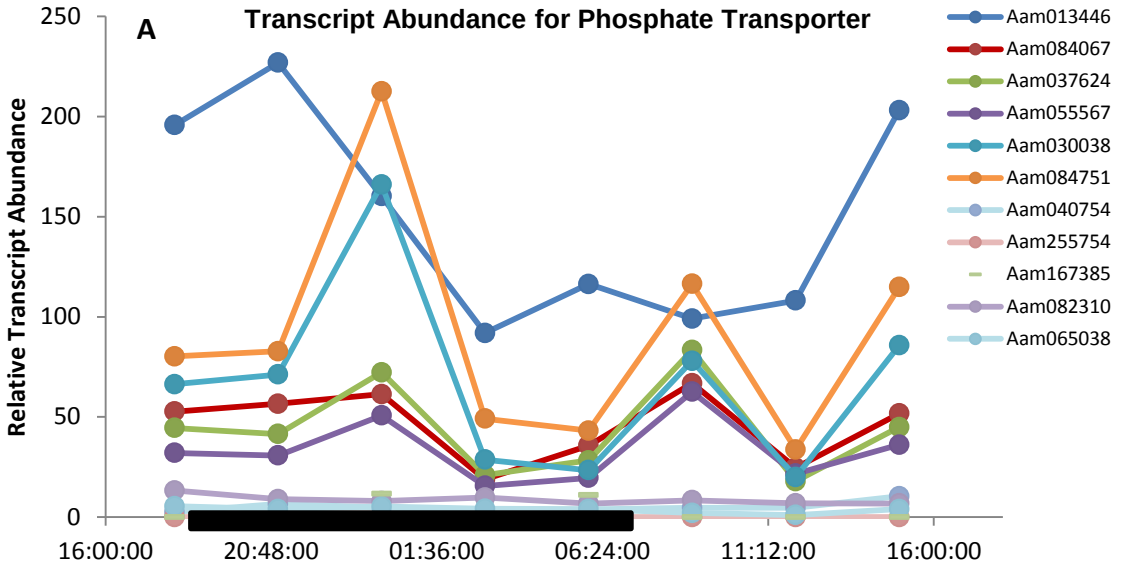
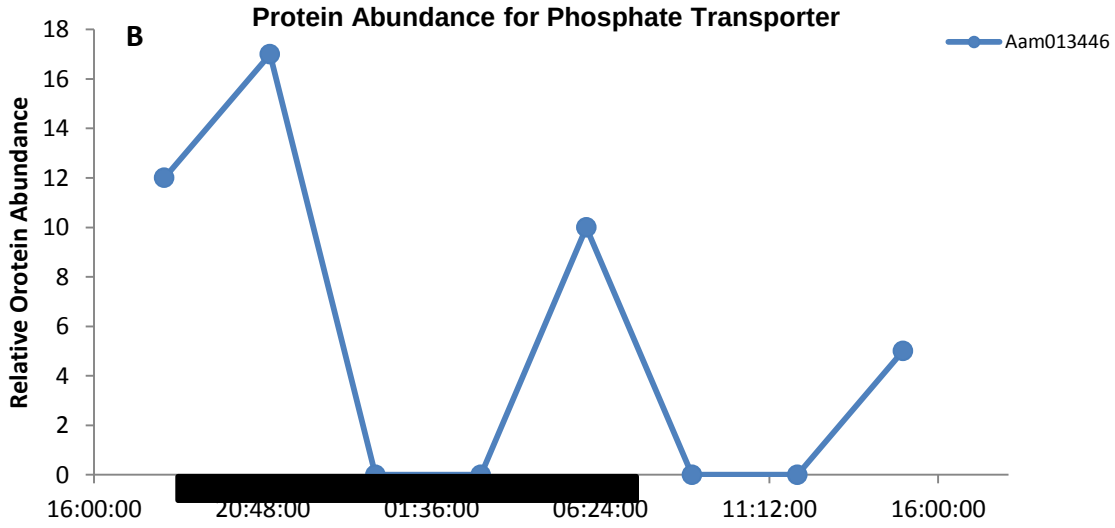


Figure 5.13 Transcript abundance profiles for ERD6-like over 24 h time course in A. mature leaves of *A. americana*, showing the highest abundance of *Aam 081118* sequence. Also shown are transcript abundances for different tissues and C. young leaves at 3 time points.


A
Transcript Abundance for Phosphate Transporter
Relative Transcript Abundance
250
200
150
100
50
0
16:00:00
20:48:00
01:36:00
06:24:00
11:12:00
16:00:00
Aam013446
Aam084067
Aam037624
Aam055567
Aam030038
Aam084751
Aam040754
Aam255754
Aam167385
Aam082310
Aam065038


B
Protein Abundance for Phosphate Transporter
Relative Orotein Abundance
18
16
14
12
10
8
6
4
2
0
16:00:00
20:48:00
01:36:00
06:24:00
11:12:00
16:00:00
Aam013446

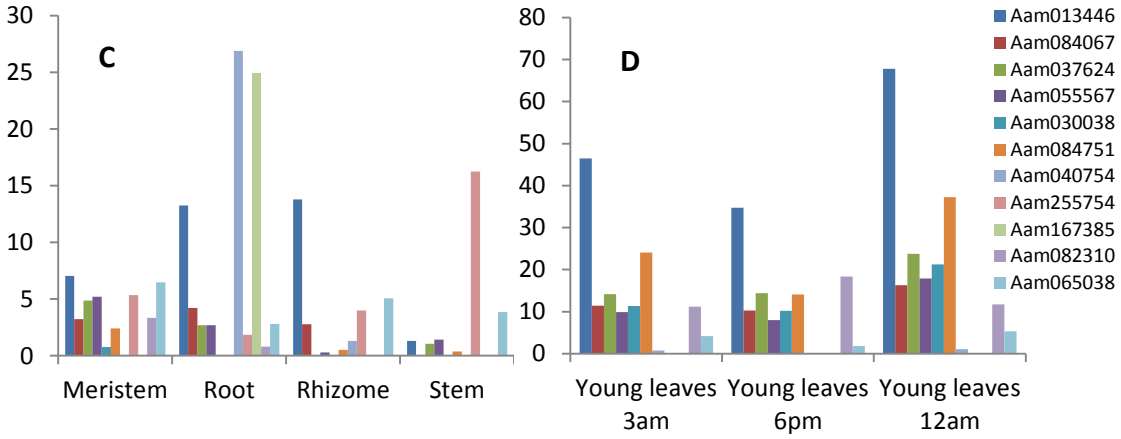


Figure 5.14 Transcript and protein abundance profiles for TMT inorganic phosphate transporter over 24 h time course in A. mature leaves of *A. americana*, showing the highest abundance of *Aam013446* sequence for both transcriptome (A) and protein (B), as well for young leaves at 12am (D).(C) distribution of transcript abundance in different plant tissues of *A. americana*.

Table 5.7 Sugar transporter sequences producing significant alignments

Sequence	Score (bits)	E Value
TMT2		
Aam045511	80	5e-013
Aam013180	62	1e-007
Aam312722	54	3e-005
Aam075796	54	3e-005
Aam015193	54	3e-005
Aam036232	48	0.002
Aam014375	48	0.002
ERD6		
Aam012894	96	7e-018
Aam032664	58	1e-006
Aam576836	42	0.084
Aam343071	42	0.084
Aam081118	42	0.084
Phos.Trans		
Aam013446	151	1e-034
Aam084067	143	3e-032
Aam037624	109	4e-022
Aam055567	92	9e-017
Aam030038	92	9e-017
Aam084751	84	2e-014
Aam040754	68	1e-009
Aam255754	66	5e-009
Aam167385	66	5e-009
Aam082310	50	3e-004
Aam065038	50	3e-004

5.4 Discussion

In this chapter, a method was developed to identify vacuolar sugar transporters in *Agave.* The approach combined biochemical assays to check on the purity of tonoplast membrane isolated via differential centrifugation and this was followed up with a proteomics approach to identify putative sugar transporter proteins.

5.4.1Tonoplast purity

Isolating sufficient amounts of a tonoplast-enriched membrane fraction that exhibits adequate purity is an essential pre-requisite for conducting informative proteome analysis to identify candidate vacuolar transporters. The enrichment and purification of tonoplast vesicles from *Agave* was based on methods of tonoplast extraction reported for the CAM species *Kalanchoe daigremontiana* and *Ananas comosus* (Bettey and Smith, 1993; McRae *et al.*, 2002). The vesicle preparations exhibited features expected for a fraction highly enriched in vacuolar membranes. ATPase activity was inhibited by more than 91% by 50 mM KNO_3, an inhibitor of vacuolar ATPase, but was only 7% inhibited by 100 μM NaN_3 and 100μM Na_3VO_4 which are inhibitors of mitochondrial and plasma membrane ATPases respectively (Gallagher and Leonard, 1982; Wang and Sze, 1985). Such data are consistent with studies on tonoplast isolated from leaves of pineapple (*Ananas comosus*) (McRae *et al.*, 2002). In the present study, specific PP_iase activity in *Agave* was relatively low compared with V-ATPase activity (Table 5.4) and a similar trend was observed in pineapple (McRae *et al.*, 2002). The specific activity for V-ATPase in *Agave* (20.1 nkat mg^{-1}protein) was higher than reported literature values for other species, which range between 1 and 5 nkat mg^{-1}protein for many C3 species, and 15.7 nkat mg^{-1}protein for the CAM plant *Ananas comosus (McRae et al., 2002)*. *Agave* displayed higher ATPase activity than the CAM species *Kalenchoe daigremontiana* (1.9 nkat mg^{-1}protein)(White and Smith, 1989) . Generally, V-PPiase activity is high in young tissues but in some cases such as in grape berries, the V-PP_iase is the predominant vacuolar proton pump in

mature plant cells (Martinoia *et al.*, 2007) even though they poses very acidic vacuoles (pH<3) (Terrier *et al.*, 1997). *A. americana marginata* PP_iase activity at 5.9 nkat mg^{-1}protein was within the range (although at the low end) of recorded values for other species which can range between 4.0 to 25.8 nkat mg^{-1}protein (Sarafian and Poole, 1989). It proved very difficult to obtain reliable measurements of the activity of V-PP_iase in *A. americana marginata*. This was later confirmed with the spectra counts from LC/MS/MS for the number of total peptides of V-ATPase and V-PP_iase, with the latter in relatively low abundance compared to the V-ATPase. It has been reported by Chen and Nose (2004) that CAM species such as *Kalanchoe* which accumulate starch as the major carbohydrate used for malic acid synthesis, have higher V-PP_iase activity than V-ATPase. The reverse is the case in pineapple which utilizes hexose (Chen and Nose, 2004). Potentially, other species of *Agave* such as *A. deserti* and *A. tequilana* with known proteome databases could be analysed for the abundance of vacuolar pumps and sugar transporters, to see if they follow the same trend. In some CAM species such as *Kalanchoë daigremontiana,* the activity of the tonoplast V-PP_iase is higher than that of the V-ATPase (Marquardt and Lüttge, 1987). The activity of each enzyme seems to depend on the type of carbohydrate used for nocturnal malate synthesis. *K. daigremoniana* uses starch to provide PEP for malate synthesis. In *Agave*, sucrose, glucose and fructose are stored in the vacuole during the day and at night these sugars are converted into precursors required for malate synthesis (Holtum *et al.*, 2005) From the results described in this chapter, it is evident that the membrane extraction method was reliable in obtaining a membrane fraction from the leaves of *Agave* that was enriched in tonoplast proteins.

5.4.2 Qualitative tonoplast proteome analysis for *Agave*

Tonoplast proteome analysis represents an analytical strategy that combines traditional biochemical methods of fractionation with more modern tools for protein identification. In this work, 934 proteins were identified by mass spectrometry, giving a broad view of the tonoplast membrane proteome and providing an important platform for the subsequent functional analyses of tonoplast proteins and transporters. The proteomics confirmed the quality of the

tonoplast preparations from *Agave* leaves. .Key and abundant tonoplastic proteins were identified such as several sub-units of the V-ATPase and PP_iase, both known to be abundant and important proteins that facilitate transport across the vacuolar membrane. However the presence of heat shock proteins and PEP carboxylase (an abundant cytosolic protein) indicates that the membrane tonoplast preparation was not 100% pure. Other proteins identified were categorized as having diverse functions that included: transporters, stress response, signal transduction, metabolism, cellular transport, protein synthesis, cytoskeleton, glycosyl hydrolase, unclassified and contaminants. Contaminants that were found in the membrane preparation seemed to be predominantly cytosolic proteins, in particular ribosomal proteins. Also, a few mitochondrial and chloroplast proteins were detected. Other transporters identified belong to the ABC transporter family. One of the subfamilies found within this transporter family was 'pleiotropic drug resistance' (PDR). Other transporters present on the tonoplast enriched membrane fraction from Agave included integral membrane proteins, nodulin like proteins, glucose-6-phosphate translocator (GTP), proton dependent oligopeptide transporter (OPT), Nramp transporter, sodium/calcium exchanger and potassium transporter (see Appendix E for all proteins identified from the *Agave* tonoplast enriched preparation).

Five candidate vacuolar sugar transporters were identified from the *Agave* membrane preparation which belong to the monosascharide transporter (-like) (MST) gene family. There are 53 members of this gene family within the *Arabidopsis* genome (Lalonde *et al.*, 2004). MST transporters possess 12 transmembrane domains.

In *Arabidopsis*, *AtTMT* transporters are believed to operate by a proton coupled anti-port mechanism which facilitates the active transport and accumulation of hexoses (glucose and fructose) in the vacuole, often in response to stresses such as drought, salinity and cold, which promote sugar accumulation (Wormit *et al.*, 2006). A homologue of *Arabidopsis* AtTMT2 was found in *Agave americana marginata* (Locus20314v1rpkm10.75_5) and was also identified in *A. comosus* fruit and root (AcMST2; (Antony and Borland, 2009) Transcript abundance of TMT2 was higher in mature (i.e. CAM-performing) leaves of *A. americana marginata* compared to young C3 leaves. However, this gene was also expressed in roots and rhizomes so it would not appear to have a CAM-

specific function in *Agave*. For this gene in pineapple (AcMST2), transcript abundance was similar in both leaves and fruits whilst AcMST1 was more highly expressed in fruit and root, implying differences in physiological functions between the MST-vacuolar transporters. However, the function of these and the *Agave* TMT2 candidate vacuolar hexose importer in the operation of CAM is not clear (Antony and Borland, 2009). There is evidence of stimulation of transcript abundance of *Arabidopsis* AtTMT by glucose (Wormit *et al.*, 2006) but there is no evidence of day/night regulation of the transcript abundance of vacuolar sugar transporters in *Arabidopsis* or pineapple (*A .comosus*) (Antony *et al.*, 2008). There did appear to be a diel change in transcript abundance of the TMT2 gene in mature and young leaves of *A. americana marginata*. There was a peak in transcript abundance in the middle of the dark period (12 am) in both leaf ages and again at 3 pm (towards the end of the light period) in mature leaves. However, the diel changes in transcript abundance were not mirrored by changes in protein abundance of the MST2 in *Agave*. Thus, it was difficult to reconcile the diel changes in transcript abundance of MST2 with a CAM-like function since a vacuolar hexose importer would be predicted to be most active during the day whilst a hexose exporter would be predicted to be most active at night. So far, there is little evidence to indicate that the MST vacuolar transporters could operate both as importers and exporters of hexoses.

Another MST sequence from *Agave* (Identifier: Locus7701v1rpkm38.97_6) showed homology to a distinct subfamily of MST genes in *Arabidopsis* designated AtERD6-LIKE 6, which is an aquaporin. Aquaporins are channel proteins present in the plasma and vacuolar membranes of plant cells, where they facilitate the transport of water and/or small neutral solutes (urea, boric acid, silicic acid) or gases (ammonia, carbon dioxide). (Johnson and Ryan, 1990; Maeshima, 1992).

It has been reported that *Kalanchoë daigremontiana*, a typical CAM plant, contains only very low amounts of vacuolar aquaporins. This might be expected for a CAM plant with minimum fluctuation of water content (Maeshima *et al.*, 1994). It has also been suggested that AtERD6 homologues(Identifier: Locus7701v1rpkm38.97_6) could play a role in the transport of sugars out of the vacuole (Büttner, 2007). The proposed model for vacuolar sugar transport in the leaves of *A. comosus* (McRae *et al.*, 2002) suggests the existence of a

tonoplast localised hexose transporter that permits efflux of glucose and fructose providing substrate for nocturnal CO_2 uptake (Antony and Borland, 2009). Transcript abundance for ERD6-LIKE 6 (Identifier: Locus7701v1rpkm38.97_6) was found to be higher in mature and young leaves of *A. americana* compared to non-CAM tissue (meristem, root, rhizome). Moreover, there was a distinct diel change in transcript abundance of the ERD6-LIKE 6 homolog in *A. americana marginata* which peaked in the middle of the dark period. This pattern of gene expression would be consistent with a proposed function of export of hexoses at night to provide substrate for dark CO_2 uptake in *Agave*. Further work is required to characterise the transport activity of ERD6-LIKE 6 in *Agave* and to compare physiological characteristics and energetic requirements of hexose transport across the tonoplast in leaves and stems of *Agave.*

The phosphate transporter (Locus834v1rpkm277.18_5) identified from the *Agave* tonoplast-enriched preparation was highly homologous with an inorganic phosphate transporter 1-7 in *A. thaliana.* This is another 12 TMT protein, which belongs to the Major Facilitator Superfamily (MFS), which includes sugar transporters. Transcript abundance of this phosphate transporter was higher in mature (i.e. CAM-performing) leaves of *A. americana marginata* compared to young C3 leaves. Aam 013446 was the transcript in highest abundance in mature and young leaves and this transcript encoded the protein with highest abundance for the TMT inorganic P trans-locator in mature leaves. Transcript abundance of Aam013446 in roots, meristems, stems and rhizoids was much lower than that in young leaves and particularly mature leaves. Thus, this protein may well have a CAM-specific function.

5.5 Conclusions

Results presented in this chapter, demonstrated that the combination of tonoplast proteomics alongside the interrogation of diel transcriptome data is a potentially powerful approach to identify candidate vacuolar sugar transporters in *Agave*. This proof of concept now needs to be developed and a more exhaustive proteomics analyses of the tonoplast membrane should be encouraged in order to identify more candidate sugar transporters which could play key regulatory roles in determining sucrose/hexose turnover for CAM as well as fructan accumulation. Such sugar transporters could represent future targets for genetic engineering of increased sugar content for plants grown for bioenergy.

Chapter 6: General Discussion

Kuwait is diversifying its energy supply by exploring the viability of different sources of renewable energy that are capable of withstanding the challenges of Kuwait's harsh climate. Succulent species of *Agave* (Agavaceae), which show high water-use efficiency, drought durability and impressive rates of biomass production (Simpson *et al.*, 2011b), represent potential bioenergy feed stocks for semi-arid, abandoned, or degraded agricultural lands which are required in order to ensure a sustainable biomass production system. The aim of this thesis was to use a combination of physiological, biochemical and proteomic approaches to start identifying traits for the improvement of *Agave* species for biomass production on arid lands.

6.1 High leaf succulence is associated with increased magnitude of CAM in *Agave*

Several authors are in agreement of the close relationship between the magnitude of CAM photosynthesis and leaf succulence with large vacuoles, providing capacitance for nocturnal acids and acting as water reservoirs. In general, a positive relationship was found between the magnitude of CAM photosynthesis and high leaf succulence across the various *Agave* species examined in this thesis (chapters 2-4). *Agave* incorporates several anatomical and physiological adaptations with CAM expression being the most important character, ensuring survival and growth under water limiting conditions. The data collected for *Agave*, indicated that the magnitude of CAM increased with leaf succulence, manifested in a higher ΔH^+, and higher rates of nocturnal net CO_2 uptake and the magnitude of CAM increased with leaf age from young to mature. Older and more succulent leaves are more committed to CAM compared to younger, thinner leaves. This was in agreement with Griffiths *et al* (2008) on CAM dicot *Kalanchoe*. Similar observations were found in the monocot *Fourcroya humboldiana.* CAM activity was also measured on a fresh weight basis. Results showed that the least succulent *A. attenuata* expressed higher CAM than succulent *A. americana*. These findings highlight the importance of units used for CAM expression. Succulence can also buffer against water limiting conditions and maintain growth. This was supported by data showing that under drought conditions, the more succulent *A. americana* produced more leaves compared to *A. attenuata*. Hence, the more succulent

Agave species potentially can outperform less succulent species under field conditions.

Leaf succulence in *Agave* appears to be a key trait for optimizing carbon gain, by accommodating large vacuolar capacities for malic acid which maximizes the amount of CO_2 taken up by PEPC during phases I and II (Osmond *et al.*, 1999). An increase in succulence is accompanied by tight cell packing; this seems to enhance CAM efficiency by reducing CO_2 leakage in phase III. Reduced internal CO_2 conductance (g_i) may promote overall carbon gain by limiting efflux of CO_2 released from decarboxylation of malate during the day (Nelson *et al.*, 2005). Reduced g_i does not appear to limit atmospheric CO_2 uptake in phase I because vacuolar capacity and PEP availability are probably the main controls over night time CO_2 acquisition (Maxwell *et al.*, 1997; Osmond *et al.*, 1999; Borland *et al.*, 2000).

Leaf succulence also seems to determine how plastic CAM expression can be as first observed in *Kalanchoë* species varying in succulence by (Dodd *et al.*, 2002). In the data presented in Chapter 2 of this thesis, the least succulent *Agave* species (*A. attenuata*) displayed similar behavior to the thin leaved *K. pinnata* showing high plasticity in photosynthetic expression under 20% and 70% F.C. *Agaves* face many challenges living in arid lands with different environmental factors such as high rates of evaporation, and drought. For the work described in this thesis, the proportion of net dark CO_2 uptake to day-time uptake increased under drought conditions in all 3 *Agave* species. Under well watered conditions, Phase II was reduced for the 2 succulent species, and the least succulent *A. attenuata* showed that net CO_2 uptake was dominated by day-time, C_3 fixation. Hartsock and Nobel (1976) observed the plasticity of *A. deserti* when under well watered conditions, which are able to to change from CAM to C_3 as manifested in daytime CO_2 uptake and no day/night acid fluctuations. Photosynthetic plasticity has been observed in *A. tequilana* young and adult plants adjusting daytime carbon gain and during the night (Pimienta-Barrios *et al.*, 2001).

The link between CAM and leaf succulence has prompted much debate on how these biochemical and morphological traits evolved, i.e. did they evolve concurrently or separately? (de Santo *et al.*, 1983) suggested that CAM is not

inextricably linked with succulence, but that CAM and succulence are often associated only because both are adaptive traits in arid environments. For example, several species within the genus *Peperomia*, had high succulence but showed low CAM activity, while some species within *Cyphostemma*, showed low succulence but exhibited high CAM activity, with overnight acid accumulation as high as 332 μeq/g fresh weight measured in the thin leaves of *Cissus* species. This suggests that in some CAM plants (i.e. *Cissus* and *Cyphostemma)*, succulence is a new acquisition allowing plant species to spread from wet tropics to arid environments. Further research work is required to establish if succulence was a trait found in the progenitors of the *Agave* genus which then led towards a predisposition to develop CAM (Sage, 2002).

The degree of leaf succulence also appears to have implications for stomatal patterning. In general, previous studies have indicated that more succulent species show lower stomatal density than less succulent species (Sayed, 1998). Some preliminary results obtained for *Agave* however have added a further layer of complexity to this observation. Stomata in *Agave* occur on both surfaces of the leaves (amphistomatous; see Figure 6.1 for stomatal impressions). Stomatal density was found to be significantly higher in the adaxial (upper) leaf surface (compared to the lower leaf surface) of the least succulent *A. attenuata* (Pearson's= -.804, sig=0.000). However, in the two more succulent species, stomatal densities were almost the same on both surfaces of the leaf (see Figure 6.2). If total stomatal density of upper and lower surfaces are combined, then total stomatal density was significantly higher in the least succulent *A. attenuata* (Pearson's= -.755, sig= 0.000) and this is in agreement with Sayed (1998).

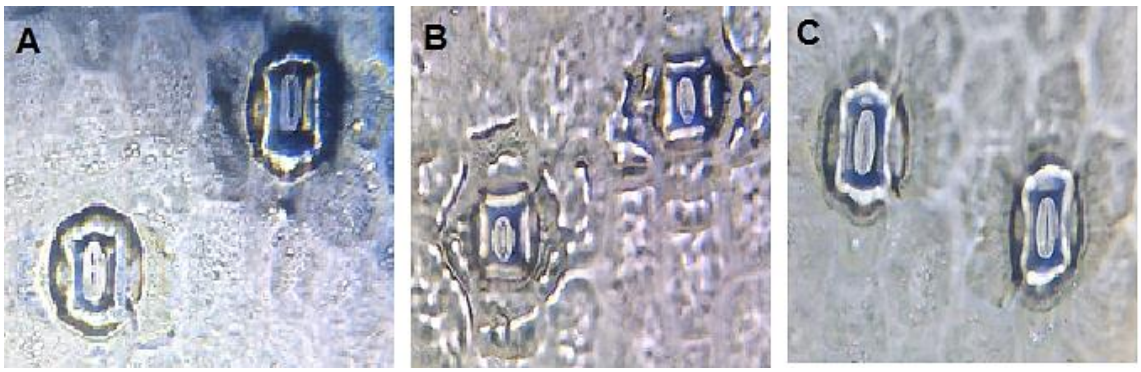


Figure 6. 1 Stomatal impressions taken from the abaxial (lower) surfaces of leaves for 3 species of Agave under the light microscope at 40X magnification. (A) *A. americana*, (B) *A. angustifolia* and (C) *A. attenuata*

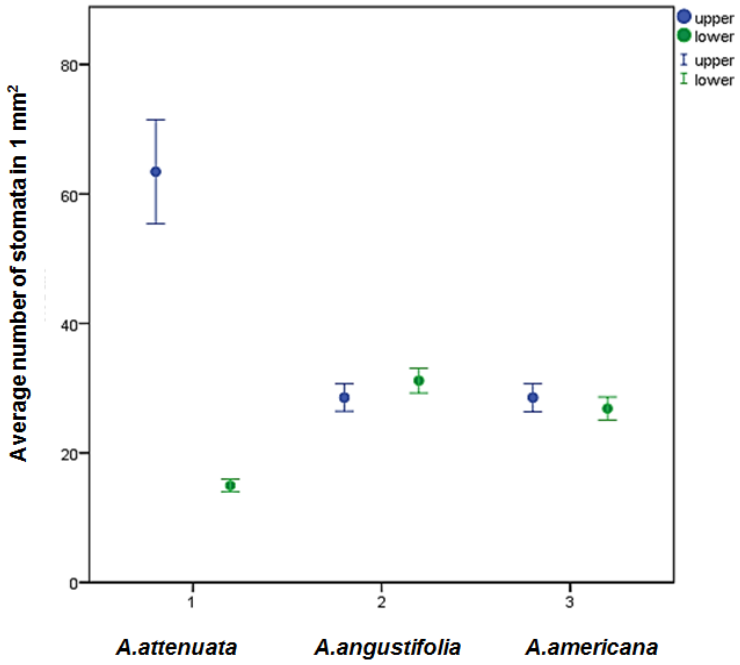


Figure 6.2 Stomatal density and distribution on both leaf surfaces in 3 *Agave* species varying in succulence, N=24, (sig=0.000, p-value< 0.05, Pearson's= -.804).

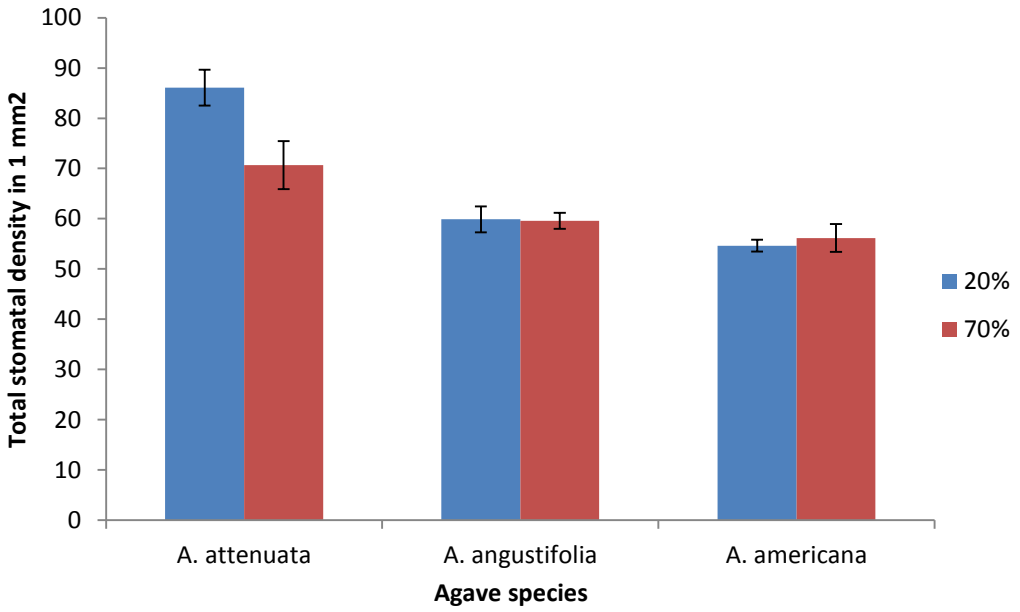


Figure 6.3 total stomatal density of upper and lower surfaces are combined in 3 *Agave* species varying in succulence, N=24.

Owen and Griffiths (2013) showed higher cumulative and instantaneous phase I (night-time) CO_2 uptake in *Agave tequilana* compared with *Kalanchoë daigremontiana,* This data stressed the importance of CO_2 conductance across the stomata and mesophyll which must be taken into consideration for CAM species. Although succulence is considered to impose constraints on CO_2 diffusion as discussed above (Maxwell *et al.*, 1997), Owen and Griffiths (2013) showed that the highly succulent *A. tequilana* had a higher stomatal density and higher chlorenchyma airspace compared with the less succulent *K. daigremontiana.* The much higher stomatal density provides a strong basis for increasing conductance of CO_2 through the stomata. Thus, high stomatal density and low chlorenchyma dry mass may be important traits for facilitating high instantaneous phase I CO_2 uptake in highly succulent species such as *Agave* and contributing towards the potentially high productivity of these species (see Figure 6.4).

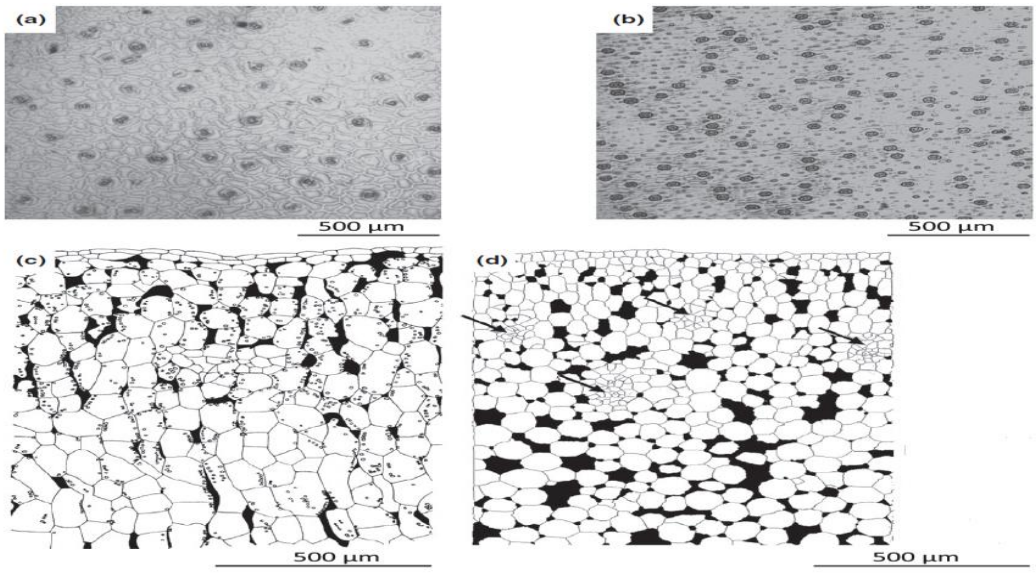


Figure 6.4 Comparison of stomatal density and chlorenchyma airspace in *A. tequilana* and *K. daigremontiana*. (a) Stomatal impression of *Kalenchoe daigremontiana* with average adaxial and abaxial stomatal density of 17 stomata mm^{-2}; (b) Stomatal impression for *A. tequilana* with average abaxial and adaxial stomatal density of 41 stomata mm^{-2}; (c) Leaf cross section of *K. daigremontiana* with average airspace 8.8% (black) and mesophyll conductance= 0.05 mol m^{-2} s^{-1} bar^{-1} (Maxwell *et al.*, 1997) (d) leaf cross section of *A. tequilana*, average chlorenchyma airspace 14.3% (black) and vascular bundles identified by arrows. Taken from (Owen and Griffiths, 2013).

The hypothesis of succulence and its relationship with the magnitude of CAM was further tested over a wide range of *Agave* species. In Chapter 4, the screening of 14 *Agave* species showed thicker, more succulent leaves were more commitment to CAM, and showed a clear correlation between succulence and leaf water content. Measurements of specific Leaf Area (SLA) confirmed the inverse relationship with both succulence and the magnitude of CAM for the 14 species of *Agave* studied. This trait (SLA) is a better predictor of species resource-use strategy than leaf water content (LWC) in succulents. In addition, SLA serves to elucidate converging strategies in carbon assimilation and nutrient conservation (Vendramini *et al.*, 2002). The low SLA found across the 14 species of Agave studied in this thesis, may be considered to incur a higher leaf level cost for light interception (Poorter *et al.*, 2009) a strategy that is common in species that inhabit environments where drought and/or nutrient limitation hamper growth (Poorter *et al.*, 2009)

Further results from Chapter 4, showed that increased levels of acid accumulated overnight were accompanied by an increase in leaf osmotic pressure which could expedite osmotic water uptake by cells. Luttge and Nobel (1984) indicated that changes in malate affected the water relations of the succulent stems of *Cereus validus*, which could act as an additional benefit of CAM for nocturnal water storage (Lüttge and Nobel, 1984).

6.2 Fructan content shows a positive link to CAM activity and succulence in *Agave*

*Agave*s display a flexibility of carbohydrate source pools to sustain dark CO_2. Fructans which are stored in the vacuole of the parenchyma of leaves and stems (Black *et al.*, 1996) also increased with leaf succulence in *Agave.* In chapter 2, it was evident that the most succulent *Agave* species under investigation, *A. americana* accumulated larger amounts of fructans than the less succulent species. Thus CAM activity and fructan accumulation appear to be linked traits. This was also evident in chapter 4, which indicated that nocturnal breakdown of fructan content had a positive relationship with the magnitude of CAM across 14 species of *Agave.* However, fructans were not the major substrate for nocturnal CO_2 fixation. In chapter 2, there was no appreciable day/night turnover of fructan in the two most succulent species, but nocturnal fructan depletion was noted in the tip and middle leaf portions of the less succulent species *A. attenuata.* Nocturnal sucrose depletion decreased from tip to base, in line with the decrease in nocturnal accumulation of titratable acids. Data in chapter 2 indicated that sucrose was the major sugar used for nocturnal acid production in *Agave*. This finding is in general agreement with other published data (Reveh *et al.* 1998) showing that diel fluctuations in leaf sucrose which could account for more than 83% of carbon needed for PEP regeneration in *A. americana.* In contrast, fructose and glucose are the major sugars used for nocturnal acid production in *A. comosus* (Carnal and Black, 1989) and *Clusia minor* (Popp *et al.*, 1987). Stoichiometric analyses of sugar breakdown and PEP requirements for CAM indicated that of the 3 *Agave* species studied in this chapter, only *A. americana* showed a shortfall in sugar depletion, implying that some nocturnal fructan depletion may be required in this species to provide PEP. Flexibility of major carbohydrate source used for the sustainability of dark CO_2 uptake is crucial for energy demands and carbon

acquisition for environments with limited precipitation. Together, the findings described above suggest that there may be genotypic variation across *Agave* in the source of carbohydrate used to provide PEP for nocturnal CO_2 uptake. Further results from chapter 4, indicated a positive relationship between the accumulation of soluble sugars with an increase of osmotic pressure. This has also been documented by Olivares and Medina (1990) when observing *Agave humboldiana*.

In this thesis, contrasting physiological roles of different leaf portions of *Agave* leaves, in terms of CAM and fructan accumulation was verified. The highest CAM activity was found in the leaf tip in all 3 *Agave* species varying in succulence, with nocturnal changes in titratable acidity increasing with distance from the leaf base. This is in agreement with published data for other monocot CAM species (Olivares and Medina, 1990; Popp *et al.*, 2003; Freschi *et al.*, 2010). In *Ananas comosus*, an increase of carbohydrate and organic solute from the base to the tip of the leaves was reported (Popp *et al.*, 2003). Olivares and Medina (1990) also showed this physiological gradient in leaves of *Fourcroya humboldtiana*. In the bromeliad *Guzmania monostachia*, there was a significant increase in ΔH^+ exclusively in the tip, where most of the activities of CAM enzymes were detected. On the other hand, little or no changes in ΔH^+ and CAM enzyme activity were detected in the leaf bases of *G. monostachia* (Freschi *et al.*, 2010). The tip, is the part of the leaf that is most t exposed to light whilst the base is shaded by the blades of upper leaves, therefore, a CAM gradient may be expected from the base to the tip (Olivares and Medina, 1990). Also, Borland and Dodd (2002) suggested that the leaf tip portion might be associated with higher availability of carbohydrates at this region and carbohydrates are known to be a key limiting resource for nocturnal CO_2 fixation in CAM plants (Borland and Dodd, 2002).

In contrast, most fructan accumulation occurred in the base of the leaf. Medina et al. (1994) suggested that carbohydrates were translocated to non photosynthetic tissues (leaf bases and stems) in *A .comosus*. The results presented here showing contrasting expression of CAM and fructan accumulation along the leaf indicates that CAM and fructan accumulation are subject to contrasting anatomical and physiological control processes, thereby indicating further complexity in the control of CAM and perhaps other metabolic

pathways. This highlights the importance of further studies regarding the existence of functional gradients along the leaf in CAM expression and establishing potential ecological and mechanistic significance.

6.3 Biochemical determinants of carboxylation process in *Agave*

6.3.1 PEPC

It was hypothesized that the abundance of PEPC will vary between *Agave* species in relation to leaf succulence and as predicted, the most succulent *A*. *americana* showed higher PEPC abundance compared to *A. attenuata*. In terms of leaf age, the abundance of PEPC was the highest in mature leaves of both species of *Agave*, complimenting titratable acidity findings on the magnitude of CAM. PEPC was not detectable in unfolded leaves of *A. attenuata*. This is in agreement with Borland et al. (1998), a study on *Clusia*, where the magnitude of CAM was related to the abundance of PEPC protein. As already discussed above, succulence in *Agave* provides high vacuolar storage capacity for malic acid which was hypothesized to maximize nocturnal PEPC capacity. The potential for high CAM activity in succulent leaved Agave was thus achieved by increased investment of leaf protein into PEPC as observed in the more succulent *Agave* species. Drought conditions intensified the abundance of PEPC in the tip of succulent *A. americana*.

The relationship between magnitude of CAM along the leaf and PEPC abundance in the leaf tip and base in the succulent *A. americana* and less succulent *A. attenuata* was unclear. This finding and others from Chapter 2 suggest that the increasing gradient of CAM activity from base to leaf tip might be due to something other than C4 carboxylase activity. PEPC enzyme activity is regulated via post-translational modification (Nimmo *et al.*, 1984; Honda *et al.*, 1996). At night, PEPC is activated via phosphorylation by a dedicated PEPC kinase which reduces enzyme sensitivity to inhibition by malate. During the day PEPC is dephosphorylated and inactive and sensitive to malate inhibition (Nimmo *et al.*, 1984; Nimmo *et al.*, 1986). However, several attempts to measure PEPC kinase activity using antibodies that recognise phosphorylated residues of PEPC over a diel cycle and in leaf tip versus leaf base by western blotting techniques were unsuccessful. This could mean that the antibodies did not recognise *Agave* PEPC although this was thought unlikely since the maize-

derived antibody recognized PEPC from *Kalanchoë fedschenkoi*, even though this species is less taxonomically related to maize than *Agave* (data not shown). It is possible that there could be a low degree of PEPC phosphorylation in Agave, or PEPC phosphorylation in *Agave* may be more subtle than previously reported for other CAM plants. This lack of detectable PEPC kinase activity could also be due to the effects of other metabolites or proteins present in *Agave* that change PEPC kinase expression or modulate the effects of PEPC phosphorylation. It has been demonstrated previously (Lepiniec *et al.*, 1994),that some C_4 monocots show modifications to the common kinetic and regulatory properties of PEPC. Future research, using molecular techniques to obtain full gene sequences of PEPC in *Agave* could be used to identify phosphorylation sites, and could also be employed to identify genes that encode PEPC kinase in *Agave* (Monocot-ME type CAM plant). Most of the research on CAM PEPC has used dicotyledonous ME-type CAM plants, such as *Mesembryanthemum crystallinum* or *Kalanchoë* species. Perhaps there are differences in regulatory properties of CAM PEPC between monocots and dicots, or PEPCK-type and ME-type CAM plants. Future work could investigate the expression and regulation of key CAM enzymes on a diel basis by employing molecular, proteomic and biochemical techniques, and investigate the possibility that protein turnover plays a role in regulation of enzyme activity, altering substrate affinity or phosphorylation status.

Future studies that consider how the leaf transcriptome and metabolome change from base to tip would be informative in revealing both how leaf development and microclimate along the leaf, influence CAM expression. A recent study (Li *et al.*, 2010) on the maize leaf transcriptome at four regions in the leaf captured a range of anatomical and biochemical states in this C4 plant. The leaf was divided into 3 major biochemical compartments. The basal region was enriched in activities for basic cellular function, the mid-leaf region was enriched in activities involved in transition from sink to source and showed an increase in abundance of transcripts associated with establishing photosynthetic machinery, and finally the leaf tip, which showed exclusive dedication to photosynthesis reactions. This approach could be of future value for identifying candidate genes for functional genomics studies to dissect photosynthetic activities in *Agave.* Such an approach could be used to generate

a transcriptome map to establish a framework for integrating additional physiological and metabolic datasets, and correlating proteomics and transcriptomics when there is low expression or resolution, serving as a foundation for a systems approach in photosynthetic development.

6.3.2 Rubisco & Rubisco activase

In contrast to the situation for PEPC, Rubisco protein abundance was higher in the least succulent leaves of *A. attenuata*, whilst Rubisco activase abundance was comparable in the two *Agave* species. The data presented in this thesis does not support the arguments of Maxwell et al (1997) and Nelson et al (2005) in which they hypothesized that thick leafed CAM plants such as *Agave* might compensate for diffusional limitation in CO_2 uptake by increased investment in Rubisco protein. Increased Rubisco protein might be predicted to enhance photosynthetic carbon gain and overcome anatomical constraints imposed by low intercellular air space to CO_2 diffusion. Leaf tips in *Agave* which are the thinnest part of the leaf had the greatest Rubisco abundance. Also, thinner leafed *A. attenuata* invested more of its leaf protein into Rubisco when compared to the succulent *A. americana*. When looking at leaf age, abundance of Rubisco and Rubisco activase were highest in mature and young leaves of both species and Rubisco abundance was intensified in the tip portion of the leaf, indicating that light intensity regulates Rubisco abundance but not PEPC abundance in *Agave*. The increased availability of light in the tip region of the leaf would help optimise the energetic of CO_2 uptake via Rubisco. However, Rubisco and Rubisco activase were below levels of detection in unfolded leaves of either species. Generally, unfolded leaves have lower chlorophyll content, and have less of an advantage photosynthetically speaking than expanded leaves and this may have influenced Rubisco content. Co-localization of both carboxylation enzymes in the tip region could improve decarboxylation efficiency during the day, allowing direct transfer of CO_2 from acid breakdown to Rubisco, which requires Rubisco activase to promote and maintain the catalytic activity of Rubisco within the same leaf area, overcoming diffusion limitations of CO_2 across the leaf (Griffiths *et al.*, 2008), optimising CO_2 draw-down and uptake. However, this was not supported by the data which showed no difference in overall abundance of Rubisco activase in both *Agave* species.

6.3.3 Rubisco activase abundance changes over a diel cycle

Overall, data in this thesis indicated the regulation of Rubisco activase abundance over a diel (24h) was apparent between the two *Agave* species varying in succulence. The physiological significance of this is unclear but could be related to leaf succulence and the relative magnitude of C3 and C4 carboxylation in the two species. As predicted, the abundance of Rubisco activase varied over the diel cycle particularly in the leaves of the more succulent *A. americana* under well watered conditions. Rubisco activase abundance was the lowest during the middle of the day, which is consistent with the idea of compensating for diffusional resistance to CO_2 (Griffiths *et al.*, 2008). Both (Cockburn W, 1979) and (Spalding MH, 1979) have shown that increasing levels of internal CO_2 within the leaf tend to down regulate the effectiveness of Rubisco activase in C3 plants. This is in agreement with results obtained here for *A. americana*, which would have high levels of internal CO_2 in the middle of the day (Phase III), which could explain the lower abundance of Rubisco activase in the middle of the day. Crafts-Brander and Salvucci (2000) also showed that interactions with high temperatures at midday tend to reduce the effectiveness of Rubisco activase in some C3 plants.

The diel change in Rubisco activase protein abundance in *A. americana* results reported in this thesis was supported by independent studies of the *A. americana* proteome (Plant Systems Biology Group, Oak Ridge National Lab), which also indicated a peak in protein abundance at night. Transcript abundance in *A. americana* however peaked at the start of the day which resulted in no clear correlation between transcript and protein abundances, indicating that Rubisco activase could be subjected to additional layers of control in addition to regulation at the level of transcription. It has been reported in some C3 plants that alternative splicing of Rubisco activase occurs (Zhang and Portis, 1999), giving rise to more than one isoform of Rubisco activase. Findings in this thesis showed several bands were noted for Rubisco activase in the western blots, particularly for *A. attenuata*. A study on rice (Wang *et al.*, 2010) indicated that two Rubisco activase isoforms displayed different roles to photosynthetic heat acclimation. Gene expression of RCA large isoform (RCA_L) and RCA small isoform (RCA_S) were investigated. Heat stress significantly

induced RCA_L expression determined by mRNA and protein levels. RCA_S was significantly related to Rubisco initial activity and net photosynthetic rate under both heat stress and normal conditions. Also the ratio of RCA_L to Rubisco increased in heat acclimated rice leaves, and expressed in enhanced amounts in transgenic rice plants which grew better at high temperatures than the wild type, playing an important role in photosynthetic acclimation to heat stress. It would be very difficult to use a transgenic approach in *Agave*, due to their slow growth; however, future immune-blot western analysis on the RCA complex could investigate the ratios of Rubisco activase isoforms under different environment conditions and their functions in *Agave*.

6.4 Identification of vacuolar sugar transporters in *Agave*

In chapter 5, a method was developed to identify vacuolar sugar transporters in *Agave.* The approach combined biochemical assays to check on the purity of tonoplast membrane isolated via differential centrifugation and this was followed up with a proteomics approach to identify putative sugar transporters which could play key regulatory roles in determining sucrose/hexose turnover for CAM as well as fructan accumulation. Such sugar transporters could represent future targets for genetic engineering of increased sugar content for plants grown for bioenergy.

A combination of tonoplast proteomics alongside the interrogation of diel transcriptome data in chapter 5, led to some 1296 protein identification events (8657 peptides at the peptide level false positive rate of less than 1%) from 6 SDS-PAGE gel bands. Many products from several gene loci which corresponded to 934 gene products were identified.

It is evident that the extraction method used was reliable in obtaining a membrane fraction from the leaves of *Agave* that was enriched in tonoplast proteins, as evidenced by the presence of vacuolar ATPases and several other known tonoplast proteins. However, the presence of heat shock proteins, PEP carboxylase and several mitochondrial proteins shows that this was not a totally pure tonoplast preparation. Treatment with Brij-58 should have reduced the number of contaminating soluble proteins (Alexandersson *et al.*, 2004). This protocol is open for future optimization for tonoplast purity and yield. In a method to isolate intact vacuoles of *A. thaliana* tonoplast (Shimaoka *et al.*,

2004), cells were centrifuged at 120,000g at 4^0C for 75 min to yield a pellet that contained purified tonoplast. In contrast, tonoplast-enriched fractions in this thesis were obtained by centrifuging at 100,000g at 4^0C for 50 min. Also, SDS-PAGE of the tonoplast fraction was performed with 7.5% acrylamide gel. Furthermore, future studies could consider employing western blots using antibodies for the V-ATPase a subunit (Matsuura-Endo *et al.*, 1992), and V-PP_iase (Takasu *et al.*, 1997). Cutting out bands from specific locations in an SDS gel could perhaps further increase purity of the fraction analysed and minimize contaminants (see Figure 6.5). For monosaccharide (hexose) CAM species such as *Agave* it has been reported that V-ATPase has higher activity than V-PP_iase in the tonoplast. This was confirmed here and could explain why it was difficult to measure PP_iase activity in *A. americana*. When transitioning from C3 to CAM photosynthesis in salted *Mesembryanthemum crystallinum* V-ATPase activity increases (Bremberger and Lüttge, 1992) which is due to de novo synthesis of V-ATPase. For the same plant, V-PPiase was highest in young plants and decreased after CAM induction by NaCl treatment. Thus, V-ATPase appears to be the main vacuolar proton pump in the CAM state. It would be interesting to see if this is a trend in several *Agave* species chosen for bioenergy. Reverse genetic techniques could be applied for each vacuolar transporter to establish its physiological role in plants (Maeshima, 2001). Thus, altering the V-ATPase or V-PP_iase activity in *Agave* could reveal their impact on nocturnal and photosynthetic performance. Maeshima (2000) suggested that V-PPiase enzyme is an essential element of giant vacuoles in plant cells. During evolution, plants obtained V-PPiase in addition to V-ATPase perhaps since V-PPiase enables vacuoles to expand (Maeshima, 2000). More studies on this enzyme could provide useful information on general plant metabolism, bioenergetics and photosynthetic specialisation. In the light of this, (Chen and Nose, 2004) demonstrated that starch degrading species such as *Kalanchoe* exhibit a tonoplast V-PPase/V-ATPase activity ratio which is 3 to 4 times higher than that in monosaccharide degrading CAM pineapple. A higher V-PP_iase activity in starch degrading CAM species is employed to generate malate by phosphorylase activity, in which monosaccharide species are unable to do (Holtum *et al.*, 2005). Using V-PP_iase saves energy leading to high nocturnal ATP levels and the release of cytosolic PP_i.

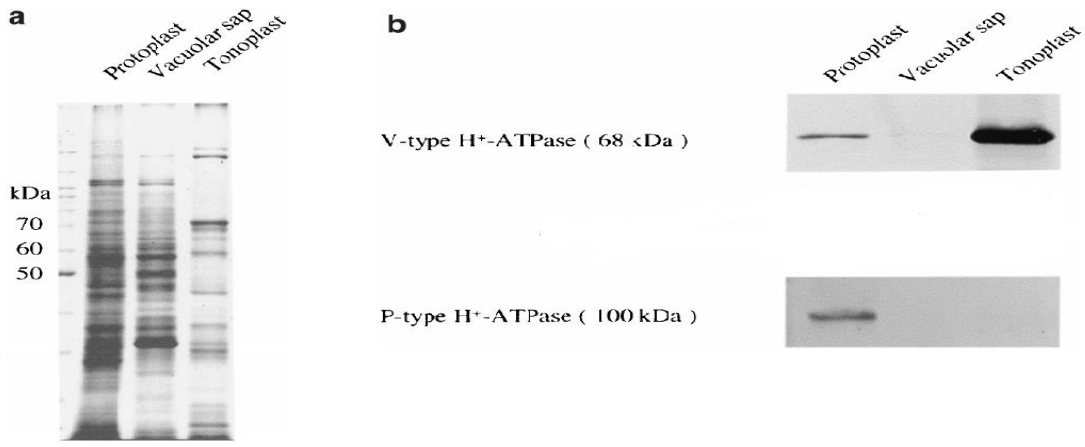


Figure 6.5 (a) SDS-PAGE of protein samples from protoplast, vacuolar sap and tonoplast isolated from suspension-cultured *A. thaliana* cells. Gel was stained with coomassie blue G. **(b)** Western blots of the same samples. Antibodies for V-ATPase and V-PPase. 50 mg tonoplast proteins were loaded in each lane. Image adopted from (Shimaoka *et al.*, 2004).

The focus of the proteomics study in Chapter 5 was to identify vacuolar sugar transporters. Those identified for *Agave* corresponded to 5 different gene loci and to 4 different proteins (3 of TMT2, 1 ERD6-LIKE and 1 inorganic phosphate transporter1-7 TMT). To uncover potential redundancy, the two vacuolar pumps and identified sugar transporter proteins were compared using multiple sequence alignment by Clustal Omega bioinformatics program. The clustal alignment showed Identifiers for 7 different loci of V-PPiases corresponded to only one gene. This was similar to the 8 different loci of V-ATPases. Further investigation into other *Agave* species that are being considered as potential bioenergy feedstocks would be informative to see if all *Agave* species follow a similar trend. In the future, omics data (genomics, transcriptomics, metabolomics, phenomics) will be help to inform genetic improvement in CAM crops and maximize the potential of *Agave* for bioenergy. Future work should help to confirm the function of candidate genes with potential in controlling stress interaction, analysis of co-suppression by overexpression of target genes, loss of function and reduction of mutants and gene silencing by RNA interference.

In addition more knowledge about vacuolar sugar transporters could provide important insights into the regulation of CAM and fructan accumulation, both important traits for bioenergy feedstocks from arid land. This will require carefully laid out quantitative proteomic experiments. Identified protein

sequence of genes for sugar transporters could potentially be used to elucidate the conservation of particular genes between species of *Agave.* Further proteo-genomic analysis of one of the *Agave* species where not only the tonoplast is analysed, but further identification of several proteins to verify the genomic sequences in the databases would be valuable.

6.5 Which Agave, where?

Agave is a promising biofuel feedstock that avoids conflict with current food supply and competition for fertile agricultural lands. *Agave* could also aid in reversing human induced land degradation and desertification by adding organic matter and stabilizing soil surfaces (Davis and Long, 2015). Much of the world's current degraded lands are not suitable for C_3 and C_4 crops without heavy irrigation. Establishment of *Agave* field trials in Kuwait will be critical for quantifying yields under contrasting environmental conditions and for validating existing EPI-based models (Owen and Griffiths, 2013; Owen and Griffiths, 2014). Such field trials are needed to help locate suitable areas for profitable yields, coupling field trials with simulated climatic scenarios. A proposed colocation of *Agave* with solar panels in Kuwait's renewable energy park could prove beneficial in water limited environments providing attractive economic incentives and efficiency in water/land use (Ravi *et al.*, 2012).

6.6 Conclusions

- High leaf succulence is associated with increased magnitude of CAM, manifested as higher ΔH^+ and nocturnal CO_2 uptake. Fructan accumulation also increased with leaf succulence in *Agave*. Sucrose provided most, if not all of the substrate required for dark CO_2 uptake. Lower water availability enhanced the proportion of dark CO_2 uptake but did not influence fructan accumulation. At the leaf level, highest CAM activity was found in the tip region whilst most fructan accumulation occurred in the base of the leaf. These results indicate that CAM and fructan accumulation are subject to contrasting anatomical and physiological control processes.
- Leaf succulence influenced the abundance of PEPC. Thus, the optimal anatomy for nocturnal malic acid accumulation is accompanied by high PEPC abundance in leaves with higher vacuolar storage capacity. In

contrast, the abundances of Rubisco and Rubisco activase showed an inverse relationship to succulence and CAM activity. Thus, in the less succulent *Agave* species which fixes a greater proportion of CO_2 during the day, investment in the C_3 carboxylating system was enhanced compared to the more succulent, strong CAM species. Differences between species in the regulation/activation of Rubisco were also apparent. Ultimately, a systems level of understanding the metabolic pathway of CAM will be required for exploiting and maximizing the potential yield of CAM species for biofuel production in marginal ecosystems.

- Inter-specific variations in the magnitude of expression of CAM in *Agave* are dependent on leaf succulence. The day/night changes in malic acid and soluble sugar contents also affect the cell sap osmotic pressure and water relations of *Agave*. Increasing levels of malic acid uptake facilitate osmotic uptake of water by cells, which is an additional benefit of CAM to nocturnal water storage (Lüttge and Nobel, 1984). Accumulation of osmotically active soluble carbohydrates can contribute to high osmotic pressures (Olivares and Medina, 1990). Soluble sugars serve as the precursors for nocturnal organic acid synthesis (Borland and Griffiths, 1989) and may also contribute to water stress tolerance in *Agave*.
- *Agave* displays flexibility in the use of carbohydrate source pools to sustain dark CO_2 uptake. Some species appear to use fructans and others sucrose as substrate for dark CO_2 uptake.
- Combining tonoplast proteomics with the interrogation of diel transcriptome data is a potentially powerful approach to identify candidate vacuolar sugar transporters in *Agave*. This proof of concept now needs to be developed and a more exhaustive proteomics analyses of the tonoplast membrane should be encouraged in order to identify more candidate sugar transporters which could play key regulatory roles in determining sucrose/hexose turnover for CAM as well as fructan accumulation. Such sugar transporters could represent future targets for genetic engineering of increased sugar content in CAM plants grown for bioenergy.

Appendix A: Image j- How to estimate leaf area measurements

The leaf is scanned on a scanner.

1. Open up Image J
2. Open Jpeg file
3. Image, type 8 bit
4. Process binary, make binary
5. Analyze, set scale, distance in pixels 71, tick global, OK
6. Analyze, analyze particles, Size 0.5-infinity, tick show outlines, display results, summarize (total area), add to manager (number of leaves on scan), OK.

Appendix B: Transcriptome and proteome databases for *Agave americana*

Access to an *Agave americana* database of diel transcript and protein abundances was provided via the CAM Biodesign consortium. The data was collected, processed, annotated and curated by researchers at the Oak Ridge National Laboratory. These databases are not yet publically available. An overview of how the data were collected and analysed by the ORNL team is given below.

Agave americana mRNA and protein profiles were obtained in biological triplicates for mature leaf tissue that was sampled every 3 hours across a 24 hour diel cycle [12-hour light (9 AM, 12 PM, 3 PM, 6 PM)/12-hour dark (9 PM, 12 AM, 3 AM, 6 AM)]. For each sample, RNA sequencing-derived (Illumina) transcript profiles were obtained and the total abundance of each mRNA was assessed by using the number of reads per kilobase and normalizing per million reads (RPKM). The following strict cut-offs were enforced to maintain a low false positive rate and to remove low abundant transcripts for quantification: 1) transcript must be observed in all replicates for at least one sample and 2) an empirically derived threshold was applied to remove low abundant transcripts that had large variance across the entire transcriptomic data set. By enforcing these criteria, a dataset of 37,808 transcripts were identified.

To generate a high-coverage proteome dataset, total protein was extracted from each sample and tryptic peptides from each sample were measured by two-

dimensional liquid chromatography nano-electrospray tandem mass spectrometry (2D-LC-MS/MS). The resulting tandem mass spectra (MS/MS) were searched with MyriMatch against an RNA sequencing-derived (RNA-Seq) proteome database. In total, 32,561 non-redundant distinct peptide sequences were identified across the entire (i.e. 24 h sampling) data set and those peptides mapped to 14,207 *A. americana* proteins sequences. The total abundance of each protein was assessed by adding peptide intensities (i.e., spectral counts) obtained in the MS analysis and normalized to their molecular weight.

Low abundance proteins were removed by enforcing the following for quantification: 1) proteins musts be observed in all replicates for at least one sample and 2) an empirically derived threshold was applied to remove low abundant proteins that had large variance across the entire transcriptomic data set. By enforcing these criteria, a dataset of 5,558 proteins were identified.

Appendix C: Correlation matrix on leaf Area basis

		DAWN acid mmol m-2	DUSK acid mmol m-2	acid accumulation mmol m-2	mmoles fructan/m2 DAWN	mmoles fructan/m2 DUSK	FRUCTAN/m2	DP Fructan Dawn	DP Fructan Dusk	succulence (kg m-2)	DAWN Mosmol	DUSK mOsmol	DAWN mol Glc m-2	DUSK mmol Glc m-2	SLA (area cm2/dry wt g)	mmol m-2 GLUCOSE DAWN	mmol m-2 FRUCTOSE DAWN	mmol m-2 SUCROSE DAWN	mmol m-2 GLUCOSE DUSK	mmol m-2 FRUCTOSE DUSK	mmol m-2 SUCROSE DUSK	GLC depletion Area	FRUC depletion Area	SUC depletion Area	gH2O/m2 leaf
DAWN acid mmol m-2	Pearson Correlation	1	.734**	.824**	-.408*	-.229	.152	-.332*	-.361*	.579**	.292	.414**	.303	.402*	-.486**	.247	.395*	.111	.407**	.466**	.516**	-.317*	-.227	-.487**	.575**
	Sig. (2-tailed)		.000	.000	.007	.150	.343	.031	.022	.000	.060	.008	.051	.008	.001	.114	.010	.485	.003	.002	.001	.041	.149	.001	.000
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
DUSK acid mmol m-2	Pearson Correlation	.734**	1	.468**	-.377*	-.319*	.045	-.377*	-.234	.777**	.063	.466**	.494**	.560**	-.532**	.372*	.482**	-.079	.382*	.527**	.284	-.172	-.246	-.412**	.781**
	Sig. (2-tailed)	.000		.002	.014	.042	.781	.014	.146	.000	.692	.002	.001	.000	.000	.015	.001	.619	.015	.000	.076	.276	.116	.007	.000
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
acid accumulation mmol m-2	Pearson Correlation	.824**	.468**	1	-.353*	-.120	.199	-.247	-.358*	.364*	.323*	.302	.088	.165	-.337*	.159	.298	.233	.279	.308	.559**	-.228	-.103	-.441**	.357*
	Sig. (2-tailed)	.000	.002		.022	.456	.212	.115	.023	.018	.037	.058	.579	.297	.028	.316	.056	.137	.081	.053	.000	.147	.514	.003	.020
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
mmoles fructan/m2 DAWN	Pearson Correlation	-.408**	-.377*	-.353*	1	.338*	-.569**	.732**	.041	-.498**	-.310*	-.510**	-.390*	-.349*	.919**	-.472**	-.429**	-.384*	-.408**	-.416**	-.550**	.071	.114	.320*	-.436**
	Sig. (2-tailed)	.007	.014	.022		.031	.000	.000	.803	.001	.046	.001	.011	.023	.000	.002	.005	.012	.009	.008	.000	.657	.472	.039	.004
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
mmoles fructan/m2 DUSK	Pearson Correlation	-.229	-.319*	-.120	.338*	1	.581**	.399**	.147	-.288	.219	-.268	-.120	-.337*	.282	-.041	-.049	.129	-.167	-.198	.001	.369*	.405**	.349*	-.273
	Sig. (2-tailed)	.150	.042	.456	.031		.000	.010	.367	.068	.170	.099	.454	.031	.074	.798	.762	.423	.302	.222	.993	.018	.009	.025	.085
	N	41	41	41	41	41	41	41	40	41	41	39	41	41	41	41	41	41	40	40	40	41	41	41	41
FRUCTAN/m2	Pearson Correlation	.152	.045	.199	-.569**	.581**	1	-.283	.090	.178	.483**	.305	.232	.007	-.548**	.369*	.325*	.439**	.270	.253	.556**	.269	.261	.024	.137
	Sig. (2-tailed)	.343	.781	.212	.000	.000		.073	.582	.266	.001	.059	.145	.964	.000	.018	.038	.004	.092	.115	.000	.089	.100	.882	.392
	N	41	41	41	41	41	41	41	40	41	41	39	41	41	41	41	41	41	40	40	40	41	41	41	41
DP Fructan Dawn	Pearson Correlation	-.332*	-.377*	-.247	.732**	.399**	-.283	1	.013	-.417**	-.245	-.536**	-.346*	-.336*	.650**	-.332*	-.326*	-.159	-.293	-.328*	-.429**	.103	.146	.408**	-.378*
	Sig. (2-tailed)	.031	.014	.115	.000	.010	.073		.934	.006	.117	.000	.025	.029	.000	.032	.035	.313	.066	.039	.006	.517	.356	.007	.014
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
DP Fructan Dusk	Pearson Correlation	-.361*	-.234	-.358*	.041	.147	.090	.013	1	.067	-.464**	.002	-.041	-.036	.014	.208	.098	-.119	.218	.202	-.130	-.101	-.226	.063	.057
	Sig. (2-tailed)	.022	.146	.023	.803	.367	.582	.934		.680	.002	.992	.799	.826	.930	.198	.547	.466	.177	.211	.423	.536	.160	.701	.726
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40	40	40
succulence (kg m-2)	Pearson Correlation	.579**	.777**	.364*	-.498**	-.288	.178	-.417**	.067	1	-.191	.250	.651**	.660**	-.656**	.705**	.773**	.021	.641**	.801**	.321*	-.224	-.344*	-.436**	.987**
	Sig. (2-tailed)	.000	.000	.018	.001	.068	.266	.006	.680		.226	.120	.000	.000	.000	.000	.000	.894	.000	.000	.043	.154	.028	.004	.000
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
DAWN Mosmol	Pearson Correlation	.292	.063	.323*	-.310*	.219	.483**	-.245	-.464**	-.191	1	.402*	-.006	-.166	-.219	-.007	.023	.472**	.010	-.093	.532**	.241	.399**	.164	-.267
	Sig. (2-tailed)	.060	.692	.037	.046	.170	.001	.117	.003	.226		.010	.968	.294	.163	.965	.884	.002	.952	.568	.000	.124	.009	.299	.188
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
DUSK mOsmol	Pearson Correlation	.414**	.466**	.302	-.510**	-.268	.305	-.536**	.002	.250	.402*	1	.219	.346*	-.516**	.219	.206	.319*	.397*	.366*	.472**	-.344*	-.338*	-.290	.223
	Sig. (2-tailed)	.008	.002	.058	.001	.099	.059	.000	.992	.120	.010		.175	.028	.001	.174	.203	.043	.012	.022	.002	.030	.032	.070	.167
	N	40	40	40	40	39	39	40	39	40	40	40	40	40	40	40	40	40	39	39	39	40	40	40	40
DAWN mol Glc m-2	Pearson Correlation	.303	.494**	.088	-.390*	-.120	.232	-.346*	-.041	.651**	-.006	.219	1	.852**	-.576**	.664**	.716**	.104	.425**	.648**	.261	.075	-.156	-.262	.718**
	Sig. (2-tailed)	.051	.001	.579	.011	.454	.145	.025	.799	.000	.968	.175		.000	.000	.000	.000	.511	.006	.000	.103	.636	.323	.094	.000
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
DUSK mmol Glc m-2	Pearson Correlation	.402**	.560**	.165	-.349*	-.337*	.007	-.336*	-.036	.660**	-.166	.346*	.852**	1	-.489**	.538**	.577**	.101	.524**	.751**	.380*	-.270	-.526**	-.481**	.725**
	Sig. (2-tailed)	.008	.000	.297	.024	.031	.964	.029	.826	.000	.294	.029	.000		.001	.000	.000	.524	.001	.000	.016	.084	.000	.001	.000
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
SLA (area cm2/dry wt g)	Pearson Correlation	-.486**	-.532**	-.337*	.919**	.282	-.548**	.650**	.014	-.656**	-.219	-.516**	-.576**	-.489**	1	-.524**	-.578**	-.261	-.419**	-.536**	-.486**	.014	.100	.325*	-.611**
	Sig. (2-tailed)	.001	.000	.028	.000	.074	.001	.000	.930	.000	.163	.001	.000	.001		.001	.000	.095	.007	.000	.001	.931	.527	.036	.000
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
mmol m-2 GLUCOSE DAWN	Pearson Correlation	.247	.372*	.159	-.472**	-.041	.369*	-.332*	.208	.705**	-.007	.219	.664**	.538**	-.524**	1	.888**	.380*	.796**	.807**	.396*	-.079	-.174	-.183	.698**
	Sig. (2-tailed)	.114	.015	.316	.002	.798	.018	.032	.198	.000	.965	.174	.000	.000	.000		.000	.013	.000	.000	.011	.618	.270	.247	.000
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
mmol m-2 FRUCTOSE DAWN	Pearson Correlation	.395*	.482**	.298	-.429**	-.049	.325*	-.326*	.098	.773**	.023	.206	.716**	.577**	-.578**	.888**	1	.324*	.689**	.839**	.362*	.002	-.038	-.129	.771**
	Sig. (2-tailed)	.010	.001	.056	.005	.762	.038	.035	.547	.000	.884	.203	.000	.000	.000	.000		.038	.000	.000	.022	.989	.812	.416	.000
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
mmol m-2 SUCROSE DAWN	Pearson Correlation	.111	-.079	.233	-.384*	.129	.439**	-.159	-.119	.021	.472**	.319*	.104	.101	-.261	.380*	.324*	1	.403**	.263	.694**	-.166	-.007	.003	-.014
	Sig. (2-tailed)	.485	.619	.137	.012	.423	.004	.313	.466	.894	.002	.045	.511	.524	.095	.013	.036		.010	.101	.000	.294	.965	.983	.929
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
mmol m-2 GLUCOSE DUSK	Pearson Correlation	.407**	.382*	.279	-.408**	-.167	.270	-.293	.218	.641**	.010	.397*	.425**	.524**	-.419**	.796**	.689**	.403**	1	.882**	.419**	-.657**	-.612**	-.199	.611**
	Sig. (2-tailed)	.009	.015	.081	.009	.302	.092	.066	.177	.000	.952	.012	.006	.001	.007	.000	.000	.010		.000	.007	.000	.000	.219	.000
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40	40	40
mmol m-2 FRUCTOSE DUSK	Pearson Correlation	.466**	.527**	.308	-.416**	-.198	.253	-.328*	.202	.801**	-.093	.366*	.648**	.751**	-.536**	.807**	.839**	.263	.882**	1	.408**	-.448**	-.610**	-.317*	.796**
	Sig. (2-tailed)	.002	.000	.053	.008	.222	.115	.038	.211	.000	.568	.022	.000	.000	.000	.000	.000	.101	.000		.009	.004	.000	.046	.000
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40	40	40
mmol m-2 SUCROSE DUSK	Pearson Correlation	.516**	.284	.559**	-.550**	.001	.556**	-.429**	-.130	.321*	.532**	.472**	.261	.380*	-.486**	.396*	.362*	.694**	.419**	.408**	1	-.195	-.222	-.722**	.312*
	Sig. (2-tailed)	.001	.076	.000	.000	.993	.000	.006	.423	.043	.000	.002	.103	.016	.001	.011	.022	.000	.007	.009		.227	.169	.000	.050
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40	40	40
GLC depletion Area	Pearson Correlation	-.317*	-.172	-.228	.071	.369*	.269	.103	-.101	-.224	.241	-.344*	.075	-.270	.014	-.079	.002	-.166	-.657**	-.448**	-.195	1	.831**	.319*	-.190
	Sig. (2-tailed)	.041	.276	.147	.657	.018	.089	.517	.536	.154	.124	.026	.636	.084	.931	.618	.989	.294	.000	.004	.227		.000	.039	.229
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
FRUC depletion Area	Pearson Correlation	-.227	-.246	-.103	.114	.405**	.261	.146	-.226	-.344*	.399**	-.338*	-.156	-.526**	.100	-.174	-.038	-.007	-.612**	-.610**	-.222	.831**	1	.515**	-.339*
	Sig. (2-tailed)	.149	.116	.514	.472	.009	.100	.356	.160	.026	.009	.032	.323	.000	.527	.270	.812	.965	.000	.000	.169	.000		.000	.028
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
SUC depletion Area	Pearson Correlation	-.487**	-.412**	-.441**	.320*	.349*	.024	.408**	.063	-.436**	.164	-.290	-.262	-.481**	.325*	-.183	-.129	.003	-.199	-.317*	-.722**	.319*	.515**	1	-.455**
	Sig. (2-tailed)	.001	.007	.003	.039	.026	.882	.007	.701	.004	.299	.070	.094	.001	.036	.247	.416	.983	.219	.046	.000	.039	.000		.002
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42
gH2O/m2 leaf	Pearson Correlation	.575**	.781**	.357*	-.436**	-.273	.137	-.378*	.057	.987**	-.207	.223	.718**	.725**	-.611**	.698**	.771**	-.014	.611**	.796**	.312*	-.190	-.339*	-.455**	1
	Sig. (2-tailed)	.000	.000	.020	.003	.085	.392	.014	.726	.000	.188	.167	.000	.000	.000	.000	.000	.929	.000	.000	.050	.229	.028	.002	
	N	42	42	42	42	41	41	42	40	42	42	40	42	42	42	42	42	42	40	40	40	42	42	42	42

Appendix D: Correlation Matrix on leaf FWT basis

Correlations

		DAWN acid umol g-1fwt	DUSK acid umol g-1 fwt	acid accumulation umol g 1 fwt	µmoles fructan/g fwt DAWN	µmoles fructan/g fwt DUSK	FRUCTAN /g Fwt	DP Fructan Dawn	DP Fructan Dusk	succulence (kg m-2)	DAWN Mosmol	DUSK mOsmol	SLA (area cm2/dry wt g)	µmoles/g fwt GLUCOSE DAWN	µmoles/g fwt FRUCTOSE DAWN	µmoles/g fwt SUCROSE DAWN	µmoles/g fwt GLUCOSE DUSK	µmoles/g fwt FRUCTOSE DUSK	µmoles/g fwt SUCROSE DUSK	GLC depletion fwt	FRUC depletion fwt	SUC depletion fwt	gH2O/m2 leaf
DAWN acid umol g-1fwt	Pearson Correlation	1	.420**	.828**	.057	.276	.414**	-.088	-.497**	-.291	.614**	.288	-.016	-.151	.001	.293	.013	-.068	.456**	-.021	.157	-.301	-.281
	Sig. (2-tailed)		.006	.000	.718	.085	.008	.578	.001	.061	.000	.071	.918	.338	.997	.060	.938	.676	.003	.897	.320	.052	.071
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
DUSK acid umol g-1 fwt	Pearson Correlation	.420**	1	.174	-.192	-.077	.007	-.246	-.416**	.240	.307*	.489**	-.269	-.201	-.050	-.182	-.127	.034	-.062	.010	-.043	-.074	.241
	Sig. (2-tailed)	.006		.271	.222	.638	.965	.116	.008	.125	.048	.001	.085	.202	.754	.250	.434	.835	.706	.949	.787	.641	.125
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
acid accumulation umol g-1 fwt	Pearson Correlation	.828**	.174	1	.247	.426**	.563**	.008	-.381*	-.449**	.505**	.127	.148	-.016	.054	.448**	.004	-.114	.587**	.076	.233	-.378*	-.434**
	Sig. (2-tailed)	.000	.271		.115	.006	.000	.959	.015	.003	.001	.434	.350	.922	.732	.003	.979	.484	.000	.631	.138	.014	.004
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
µmoles fructan/g fwt DAWN	Pearson Correlation	.057	-.192	.247	1	.881**	.688**	.399*	.145	-.380*	-.011	-.189	.487**	.216	.233	.224	.017	-.013	.201	.200	.240	-.066	-.330*
	Sig. (2-tailed)	.718	.222	.115		.000	.000	.009	.371	.013	.944	.243	.001	.169	.138	.154	.916	.938	.213	.205	.126	.679	.033
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
µmoles fructan/g fwt DUSK	Pearson Correlation	.276	-.077	.426**	.881**	1	.900**	.261	.017	-.339*	.113	-.016	.372*	.278	.440**	.259	.125	.192	.283	.170	.275	-.216	-.300
	Sig. (2-tailed)	.085	.638	.006	.000		.000	.103	.919	.032	.489	.923	.018	.083	.004	.107	.443	.236	.077	.294	.086	.182	.060
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40
FRUCTAN /g Fwt	Pearson Correlation	.414**	.007	.563**	.688**	.900**	1	-.019	-.067	-.150	.293	.195	.028	.433**	.651**	.328*	.286	.401*	.390*	.121	.251	-.318*	-.141
	Sig. (2-tailed)	.008	.965	.000	.000	.000		.910	.683	.356	.066	.233	.862	.005	.000	.039	.074	.010	.013	.458	.118	.045	.387
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40
DP Fructan Dawn	Pearson Correlation	-.088	-.246	.008	.399*	.261	-.019	1	.013	-.417**	-.245	-.536**	.648**	-.248	-.248	-.070	-.228	-.337*	-.247	.083	.147	.362*	-.378*
	Sig. (2-tailed)	.578	.116	.959	.009	.103	.910		.934	.006	.117	.000	.000	.114	.113	.657	.156	.033	.125	.603	.353	.019	.014
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
DP Fructan Dusk	Pearson Correlation	-.497**	-.416**	-.381*	.145	.017	-.067	.013	1	.087	-.464**	.002	.014	.150	.049	-.126	.182	.181	-.152	-.091	-.205	.118	.057
	Sig. (2-tailed)	.001	.008	.015	.371	.919	.683	.934		.680	.003	.992	.930	.355	.763	.437	.260	.263	.349	.576	.205	.469	.726
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40
succulence (kg m-2)	Pearson Correlation	-.291	.240	-.449**	-.380*	-.339*	-.150	-.417**	.087	1	-.191	.250	-.657**	.145	.344*	-.492**	.225	.592**	-.421**	-.209	-.353*	.021	.987**
	Sig. (2-tailed)	.061	.125	.003	.013	.032	.356	.006	.680		.226	.120	.000	.360	.026	.001	.163	.000	.007	.184	.022	.895	.000
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
DAWN Mosmol	Pearson Correlation	.614**	.307*	.505**	-.011	.113	.293	-.245	-.464**	-.191	1	.402*	-.219	.304	.270	.538**	.178	.031	.641**	.373*	.465**	.035	-.207
	Sig. (2-tailed)	.000	.048	.001	.944	.489	.066	.117	.003	.226		.010	.164	.051	.084	.000	.271	.851	.000	.015	.002	.827	.188
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
DUSK mOsmol	Pearson Correlation	.288	.489**	.127	-.189	-.016	.195	-.536**	.002	.250	.402*	1	-.515**	.190	.200	.163	.436**	.456**	.280	-.294*	-.314*	-.278	.223
	Sig. (2-tailed)	.071	.001	.434	.243	.923	.233	.000	.992	.120	.010		.001	.240	.216	.316	.005	.004	.084	.066	.048	.082	.166
	N	40	40	40	40	39	39	40	39	40	40	40	40	40	40	40	39	39	39	40	40	40	40
SLA (area cm2/dry wt g)	Pearson Correlation	-.016	-.269	.148	.487**	.372*	.028	.648**	.014	-.657**	-.219	-.515**	1	-.322*	-.407**	.027	-.213	-.475**	-.067	-.050	.093	.119	-.611**
	Sig. (2-tailed)	.918	.085	.350	.001	.018	.862	.000	.930	.000	.164	.001		.038	.007	.867	.186	.002	.681	.753	.560	.451	.000
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
µmoles/g fwt GLUCOSE DAWN	Pearson Correlation	-.151	-.201	-.016	.216	.278	.433**	-.248	.150	.145	.304	.190	-.322*	1	.724**	.434**	.627**	.576**	.372*	.233	.127	-.027	.131
	Sig. (2-tailed)	.338	.202	.922	.169	.083	.005	.114	.355	.360	.051	.240	.038		.000	.004	.000	.000	.018	.138	.421	.865	.407
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
µmoles/g fwt FRUCTOSE DAWN	Pearson Correlation	.001	-.050	.054	.233	.440**	.651**	-.248	.049	.344*	.270	.200	-.407**	.724**	1	.144	.405**	.712**	.093	.318*	.335*	.173	.331*
	Sig. (2-tailed)	.997	.754	.732	.138	.004	.000	.113	.763	.026	.084	.216	.007	.000		.361	.010	.000	.567	.040	.030	.274	.032
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
µmoles/g fwt SUCROSE DAWN	Pearson Correlation	.293	-.182	.448**	.224	.259	.328*	-.070	-.126	-.492**	.538**	.163	.027	.434**	.144	1	.254	-.047	.905**	.123	.197	-.314*	-.495**
	Sig. (2-tailed)	.060	.250	.003	.154	.107	.039	.657	.437	.001	.000	.316	.867	.004	.361		.113	.772	.000	.438	.212	.043	.001
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
µmoles/g fwt GLUCOSE DUSK	Pearson Correlation	.013	-.127	.004	.017	.125	.286	-.228	.182	.225	.178	.436**	-.213	.627**	.405**	.254	1	.753**	.227	-.618**	-.523**	-.056	.191
	Sig. (2-tailed)	.938	.434	.979	.916	.443	.074	.156	.260	.163	.271	.005	.186	.000	.010	.113		.000	.159	.000	.001	.732	.237
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40
µmoles/g fwt FRUCTOSE DUSK	Pearson Correlation	-.068	.034	-.114	-.013	.192	.401*	-.337*	.181	.592**	.031	.456**	-.475**	.576**	.712**	-.047	.753**	1	-.010	-.368*	-.509**	-.024	.570**
	Sig. (2-tailed)	.676	.835	.484	.938	.236	.010	.033	.263	.000	.851	.004	.002	.000	.000	.772	.000		.950	.019	.001	.883	.000
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40
µmoles/g fwt SUCROSE DUSK	Pearson Correlation	.456**	-.062	.587**	.201	.283	.390*	-.247	-.152	-.421**	.641**	.280	-.067	.372*	.093	.905**	.227	-.010	1	.109	.124	-.753**	-.415**
	Sig. (2-tailed)	.003	.706	.000	.213	.077	.013	.125	.349	.007	.000	.084	.681	.018	.567	.000	.159	.950		.504	.447	.000	.008
	N	40	40	40	40	40	40	40	40	40	40	39	40	40	40	40	40	40	40	40	40	40	40
GLC depletion fwt	Pearson Correlation	-.021	.010	.076	.200	.170	.121	.083	-.091	-.209	.373*	-.294	-.050	.233	.318*	.123	-.618**	-.368*	.109	1	.844**	.311*	-.186
	Sig. (2-tailed)	.897	.949	.631	.205	.294	.458	.603	.576	.184	.015	.066	.753	.138	.040	.438	.000	.019	.504		.000	.045	.237
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
FRUC depletion fwt	Pearson Correlation	.157	-.043	.233	.240	.275	.251	.147	-.205	-.353*	.465**	-.314*	.093	.127	.335*	.197	-.523**	-.509**	.124	.844**	1	.452**	-.344*
	Sig. (2-tailed)	.320	.787	.138	.126	.086	.118	.353	.205	.022	.002	.048	.560	.421	.030	.212	.001	.001	.447	.000		.003	.026
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
SUC depletion fwt	Pearson Correlation	-.301	-.074	-.378*	-.066	-.216	-.318*	.362*	.118	.021	.035	-.278	.119	-.027	.173	-.314*	-.056	-.024	-.753**	.311*	.452**	1	.004
	Sig. (2-tailed)	.052	.641	.014	.679	.182	.045	.019	.469	.895	.827	.082	.451	.865	.274	.043	.732	.883	.000	.045	.003		.981
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42
gH2O/m2 leaf	Pearson Correlation	-.281	.241	-.434**	-.330*	-.300	-.141	-.378*	.057	.987**	-.207	.223	-.611**	.131	.331*	-.495**	.191	.570**	-.415**	-.186	-.344*	.004	1
	Sig. (2-tailed)	.071	.125	.004	.033	.060	.387	.014	.726	.000	.188	.166	.000	.407	.032	.001	.237	.000	.008	.237	.026	.981	
	N	42	42	42	42	40	40	42	40	42	42	40	42	42	42	42	40	40	40	42	42	42	42

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix E: Proteomics analysis of 934 identified protein events

Band	Identifier	log(I)	rI	log(e)	pI	Mr	TotalPep	pfam_description	pfam_start	pfam_end	interpro_description	evalue
4	Locus10407v1rpkm27.57_12	4.6	2	-1.3	9	59	3	Glyco_transf_8	199	477	Glycosyl transferase, family 8	8.90E-74
1	Locus10627v1rpkm26.82_29	4	2	-10	6	184	4	Glyco_transf_8	1328	1540	Glycosyl transferase, family 8	2.40E-06
4	Locus185v1rpkm722.03_12	5.4	15	-127	6	67	39	Glycos_transf_1	384	562	Glycosyl transferase, family 1	5.80E-32
4	Locus185v1rpkm722.03_12	5.4	15	-127	6	67	39	Sucrose_synth	1	379	Sucrose synthase	5.30E-238
3	Locus26662v1rpkm6.36_5	4.7	3	-9.6	7	30	3	Glyco_transf_28	141	276	Glycosyl transferase, family 28	1.50E-31
1	Locus20314v1rpkm10.75_5	4	2	-14	6	35	2	Sugar_tr	96	316	General substrate transporter	3.80E-40
1	Locus3753v1rpkm79.87_8	4.2	6	-44	5	64	6	Sugar_tr	367	586	General substrate transporter	2.60E-42
1	Locus3753v1rpkm79.87_8	4.2	6	-44	5	64	6	Sugar_tr	2	75	General substrate transporter	7.70E-14
1	Locus6095v1rpkm49.88_8	4.6	6	-53	5	56	7	Sugar_tr	7	219	General substrate transporter	1.10E-50
1	Locus7701v1rpkm38.97_6	3.3	1	-3.3	9	43	1	Sugar_tr	57	376	General substrate transporter	2.50E-66
1	Locus834v1rpkm277.18_5	4.5	3	-20	9	55	3	Sugar_tr	2	487	General substrate transporter	1.70E-48
6	Locus12164v1rpkm22.58_15	3.6	1	-2.6	6	92	1	Glycos_transf_1	564	736	Glycosyl transferase, family 1	8.20E-32
6	Locus12164v1rpkm22.58_15	3.6	1	-2.6	6	92	1	Sucrose_synth	8	552	Sucrose synthase	0.00E+00
2	Locus1274v1rpkm203.25_12	4.5	4	-38	6	63	9	PGM_PMM_I	16	163	Alpha-D-phosphohexomutase, alpha/beta/alpha domain I	8.90E-34
2	Locus1274v1rpkm203.25_12	4.5	4	-38	6	63	9	PGM_PMM_II	198	308	Alpha-D-phosphohexomutase, alpha/beta/alpha domain II	6.70E-13
2	Locus1274v1rpkm203.25_12	4.5	4	-38	6	63	9	PGM_PMM_III	316	439	Alpha-D-phosphohexomutase, alpha/beta/alpha domain III	4.90E-26
2	Locus1274v1rpkm203.25_12	4.5	4	-38	6	63	9	PGM_PMM_IV	492	554	Alpha-D-phosphohexomutase, C-terminal	1.10E-10
3	Locus12899v1rpkm21.05_5	4.9	7	-50	6	48	7	DDOST_48kD	31	435	Oligosaccharyl transferase complex, subunit Wbp1	1.00E-129
4	Locus1391v1rpkm189.34_9	4.2	3	-21	6	59	4	PFK	144	382	Phosphofructokinase domain	2.50E-37
4	Locus14119v1rpkm18.72_5	4.7	4	-46	6	48	4	DDOST_48kD	31	435	Oligosaccharyl transferase complex, subunit Wbp1	7.30E-127
4	Locus14564v1rpkm17.93_10	3.5	1	-4.6	7	46	1	PFK	8	159	Phosphofructokinase domain	1.30E-14
1	Locus15508v1rpkm16.44_14	3.8	2	-9.2	9	88	3	STT3	23	716	Oligosaccharyl transferase, STT3 subunit	3.00E-128
2	Locus1569v1rpkm173.19_12	3.3	2	-8.1	7	66	4	PFK	124	331	Phosphofructokinase domain	9.20E-17
3	Locus426v1rpkm432.89_2	4.5	1	-2.2	6	12	1	Sucrose_synth	8	108	Sucrose synthase	1.80E-42
1	Locus45647v1rpkm1.26_16	3.6	2	-10	6	89	2	Glycos_transf_1	568	733	Glycosyl transferase, family 1	2.20E-31
1	Locus45647v1rpkm1.26_16	3.6	2	-10	6	89	2	Sucrose_synth	9	558	Sucrose synthase	1.20E-231
4	Locus517v1rpkm379.21_15	4.8	4	-24	6	77	4	Glycos_transf_1	561	664	Glycosyl transferase, family 1	2.00E-11
4	Locus517v1rpkm379.21_15	4.8	4	-24	6	77	4	Sucrose_synth	7	557	Sucrose synthase	0.00E+00
3	Locus572v1rpkm355.76_5	4.6	2	-12	6	51	3	PFK	90	340	Phosphofructokinase domain	5.90E-25
3	Locus7336v1rpkm41.13_6	5.5	5	-29	5	40	11	Sucrose_synth	7	351	Sucrose synthase	9.20E-197
5	Locus7575v1rpkm39.72_10	4.5	3	-17	9	81	4	STT3	25	669	Oligosaccharyl transferase, STT3 subunit	3.50E-154
4	Locus8777v1rpkm33.78_3	4.2	3	-21	9	29	5	Glycos_transf_1	3	164	Glycosyl transferase, family 1	2.60E-32
3	Locus9434v1rpkm31.13_5	5.1	9	-78	7	48	20	DDOST_48kD	31	435	Oligosaccharyl transferase complex, subunit Wbp1	3.40E-128
1	Locus4200v1rpkm71.65_8	3.6	1	-6.1	7	74	1	Malectin_like	41	411	Malectin-like carbohydrate-binding domain	2.30E-52
4	Locus513v1rpkm381.44_4	4.4	3	-30	5	37	3	PfkB	19	308	Carbohydrate/purine kinase	9.80E-78
5	Locus7602v1rpkm39.56_2	4	1	-3.3	5	35	1	PfkB	16	321	Carbohydrate/purine kinase	1.10E-84
5	Locus1635v1rpkm167.93_3	5.4	9	-74	5	30	13	14-3-3	8	243	14-3-3 domain	1.40E-114
5	Locus280v1rpkm567.27_2	5.1	5	-52	5	26	7	14-3-3	1	215	14-3-3 domain	2.30E-108
6	Locus2892v1rpkm102.68_2	4.7	4	-38	5	28	7	14-3-3	8	246	14-3-3 domain	2.60E-108
5	Locus6243v1rpkm48.84_4	4.9	5	-32	5	30	5	14-3-3	9	244	14-3-3 domain	1.80E-113
3	Locus5428v1rpkm56.19_6	4.2	1	-5.6	7	41	2	2-Hacid_dh	29	346	D-isomer specific 2-hydroxyacid dehydrogenase, catalytic domain	2.60E-19
3	Locus5428v1rpkm56.19_6	4.2	1	-5.6	7	41	2	2-Hacid_dh_C	130	322	D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding	8.00E-46
3	Locus14282v1rpkm18.42_10	4	1	-2.3	8	60	1	2-oxoacid_dh	325	555	2-oxoacid dehydrogenase acyltransferase, catalytic domain	1.60E-78
3	Locus22259v1rpkm9.12_9	4.1	1	-1.6	9	50	1	3Beta_HSD	12	287	3-beta hydroxysteroid dehydrogenase/isomerase	3.50E-76
1	Locus1594v1rpkm170.83_14	4.6	7	-56	5	80	10	AAA	157	286	ATPase, AAA-type, core	2.90E-47
1	Locus1594v1rpkm170.83_14	4.6	7	-56	5	80	10	AAA	430	563	ATPase, AAA-type, core	4.30E-47
5	Locus1697v1rpkm162.63_6	3.5	1	-4.8	6	51	1	AAA	58	189	ATPase, AAA-type, core	1.80E-44
3	Locus2261v1rpkm128.95_7	5	3	-14	6	36	3	AAA	110	243	ATPase, AAA-type, core	2.30E-44
6	Locus2703v1rpkm109.97_12	3.8	1	-4.7	5	92	1	AAA	207	336	ATPase, AAA-type, core	1.80E-15
5	Locus6234v1rpkm48.90_7	3.4	2	-9.7	7	48	4	AAA	162	303	ATPase, AAA-type, core	6.30E-15
3	Locus9724v1rpkm29.90_10	4.9	4	-21	5	38	4	AAA	126	258	ATPase, AAA-type, core	2.00E-42
6	Locus2703v1rpkm109.97_12	3.8	1	-4.7	5	92	1	AAA_2	544	718	ATPase, AAA-2	6.60E-55
6	Locus39826v1rpkm1.99_4	4.1	2	-17	9	32	2	ABC_membrane	4	269	ABC transporter, transmembrane domain	4.80E-33
6	Locus46069v1rpkm1.22_6	4.9	1	-1.1	10	36	1	ABC_membrane	89	324	ABC transporter, transmembrane domain	1.60E-41
3	Locus15197v1rpkm16.92_24	3.9	1	-3.8	8	89	1	ABC_membrane_2	190	460	ABC transporter, N-terminal	5.40E-81
3	Locus15197v1rpkm16.92_24	3.9	1	-3.8	8	89	1	ABC_tran	588	732	ABC transporter-like	5.00E-09
3	Locus15197v1rpkm16.92_24	3.9	1	-3.8	8	89	1	ABC_tran	4	57	ABC transporter-like	3.40E-06
1	Locus40755v1rpkm1.84_25	3.3	1	-8	9	162	1	ABC_tran	190	347	ABC transporter-like	3.80E-06
1	Locus40755v1rpkm1.84_25	3.3	1	-8	9	162	1	ABC_tran	899	1027	ABC transporter-like	6.80E-11
1	Locus40755v1rpkm1.84_25	3.3	1	-8	9	162	1	ABC2_membrane	1172	1386	ABC-2 type transporter	2.30E-54
1	Locus40755v1rpkm1.84_25	3.3	1	-8	9	162	1	ABC2_membrane	502	714	ABC-2 type transporter	6.30E-43
4	Locus12202v1rpkm22.49_5	3.6	1	-7.5	9	40	1	Abhydrolase_6	85	340	NULL	3.80E-24
3	Locus21805v1rpkm9.45_6	4.3	1	-3.7	6	40	1	Abhydrolase_6	67	338	NULL	3.40E-22
3	Locus59537v1rpkm0.63_4	3.7	1	-4.6	8	29	1	Abhydrolase_6	67	181	NULL	1.70E-16

3	Locus7794v1rpkm38.59_4	5	5	-40	8	42	9	Abhydrolase_6	98	355	NULL	7.00E-26
5	Locus3940v1rpkm76.12_5	3.4	1	-1.5	9	43	1	Abi	212	364	CAAX amino terminal protease	6.30E-11
2	Locus3332v1rpkm89.78_14	3.3	1	-2.2	7	56	1	ACOX	348	507	Acyl-CoA oxidase, C-terminal	1.20E-49
3	Locus4251v1rpkm70.82_5	5.1	6	-54	9	52	15	ACP_syn_III_C	339	420	3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C-terminal	2.40E-11
3	Locus4400v1rpkm68.84_2	5.3	5	-28	9	56	5	ACP_syn_III_C	386	466	3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C-terminal	2.60E-13
1	Locus7676v1rpkm39.08_6	4	3	-27	9	59	13	ACP_syn_III_C	415	495	3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C-terminal	4.60E-11
3	Locus8644v1rpkm34.26_6	5.2	4	-21	9	48	5	ACP_syn_III_C	322	403	3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C-terminal	3.30E-10
4	Locus140v1rpkm817.16_3	3.9	2	-12	6	8.6	2	Actin	4	80	Actin-like	2.70E-24
5	Locus300v1rpkm553.21_7	5.2	7	-73	5	42	10	Actin	5	377	Actin-like	1.40E-159
3	Locus6034v1rpkm50.34_7	5.5	8	-64	5	42	13	Actin	5	377	Actin-like	2.90E-159
3	Locus10021v1rpkm28.89_17	3.5	1	-2	7	92	1	Acyl-CoA_dh_1	674	822	Acyl-CoA oxidase/dehydrogenase, type 1	5.10E-34
3	Locus10021v1rpkm28.89_17	3.5	1	-2	7	92	1	Acyl-CoA_dh_M	561	615	Acyl-CoA oxidase/dehydrogenase, central	3.20E-16
2	Locus3332v1rpkm89.78_14	3.3	1	-2.2	7	56	1	Acyl-CoA_dh_M	6	64	Acyl-CoA oxidase/dehydrogenase, central domain	6.90E-11
3	Locus10021v1rpkm28.89_17	3.5	1	-2	7	92	1	Acyl-CoA_dh_N	414	557	Acyl-CoA dehydrogenase, N-terminal	7.60E-10
4	Locus3225v1rpkm92.43_10	3.6	1	-2.2	6	38	1	ADH_N	32	145	Alcohol dehydrogenase GroES-like	2.90E-26
4	Locus8579v1rpkm34.57_8	3.4	1	-1.2	6	35	1	ADH_N	29	87	Alcohol dehydrogenase GroES-like	8.30E-12
4	Locus16115v1rpkm15.52_1	3.2	1	-5.6	11	12	1	adh_short	39	109	Short-chain dehydrogenase/reductase SDR	1.30E-12
4	Locus11886v1rpkm23.31_4	3.5	1	-2.7	9	20	1	adh_short	1	67	Short-chain dehydrogenase/reductase SDR	1.10E-08
4	Locus13197v1rpkm20.43_6	4.4	4	-25	6	24	4	adh_short	12	121	Short-chain dehydrogenase/reductase SDR	1.20E-18
5	Locus38297v1rpkm2.27_5	4.7	2	-9.1	9	37	3	adh_short	73	244	Short-chain dehydrogenase/reductase SDR	6.30E-21
5	Locus5091v1rpkm59.77_6	3.5	1	-2	6	21	1	adh_short	29	182	Short-chain dehydrogenase/reductase SDR	2.50E-31
2	Locus75876v1rpkm0.40_6	4	1	-1.8	9	24	1	adh_short	14	62	Short-chain dehydrogenase/reductase SDR	5.10E-06
4	Locus1042v1rpkm234.65_3	4.5	2	-16	8	24	2	ADH_zinc_N	71	134	Alcohol dehydrogenase, C-terminal	5.50E-09
4	Locus3225v1rpkm92.43_10	3.6	1	-2.2	6	38	1	ADH_zinc_N	188	310	Alcohol dehydrogenase, C-terminal	3.10E-20
4	Locus8579v1rpkm34.57_8	3.4	1	-1.2	6	35	1	ADH_zinc_N	151	266	Alcohol dehydrogenase, C-terminal	6.70E-33
5	Locus14229v1rpkm18.51_3	5.5	7	-84	7	20	12	ADK	1	157	Adenylate kinase	2.80E-43
5	Locus15653v1rpkm16.22_4	5.5	8	-83	8	25	8	ADK	23	208	Adenylate kinase	1.40E-58
5	Locus14229v1rpkm18.51_3	5.5	7	-84	7	20	12	ADK_lid	94	129	Adenylate kinase, active site lid domain	1.80E-17
5	Locus15653v1rpkm16.22_4	5.5	8	-83	8	25	8	ADK_lid	145	180	Adenylate kinase, active site lid domain	2.60E-17
3	Locus246v1rpkm610.10_9	4.6	3	-15	6	51	3	AdoHcyase	1	466	Adenosylhomocysteinase	1.80E-139
3	Locus246v1rpkm610.10_9	4.6	3	-15	6	51	3	AdoHcyase_NAD	222	385	S-adenosyl-L-homocysteine hydrolase, NAD binding	6.10E-84
1	Locus4551v1rpkm66.45_11	4.8	6	-52	8	70	9	AIG1	29	163	AIG1	2.30E-26
4	Locus8600v1rpkm34.45_9	4.2	1	-1.9	9	35	1	AIG1	38	199	AIG1	2.90E-34
3	Locus7053v1rpkm42.76_7	3.9	1	-3.5	6	47	1	ALAD	102	421	Tetrapyrrole biosynthesis, porphobilinogen synthase	1.30E-138
3	Locus30465v1rpkm4.55_7	5.8	14	-95	9	54	26	Aldedh	23	445	Aldehyde dehydrogenase domain	1.50E-81
1	Locus5285v1rpkm57.56_2	3.7	1	-3.8	9	31	1	Aldedh	1	237	Aldehyde dehydrogenase domain	1.80E-49
3	Locus6303v1rpkm48.31_5	6	19	-139	8	42	38	Aldedh	9	389	Aldehyde dehydrogenase domain	2.20E-78
4	Locus6333v1rpkm47.95_8	3.1	1	-14	7	66	1	Aldedh	60	526	Aldehyde dehydrogenase domain	4.20E-125
3	Locus8237v1rpkm36.42_3	5	3	-16	9	35	6	Aldedh	5	255	Aldehyde dehydrogenase domain	2.00E-46
1	Locus18087v1rpkm13.05_9	4.3	3	-21	6	59	3	Alpha-mann_mid	297	381	Glycoside hydrolase, family 38, central domain	2.60E-19
2	Locus15811v1rpkm16.00_8	4.9	8	-70	6	69	8	Amidase	187	598	Amidase	3.60E-83
2	Locus17811v1rpkm13.36_12	4.2	4	-26	6	57	4	Amidohydro_1	89	434	Amidohydrolase 1	1.20E-15
2	Locus17168v1rpkm14.13_10	3.3	1	-5.8	9	56	4	Amino_oxidase	27	501	Amine oxidase	6.40E-47
2	Locus7712v1rpkm38.92_10	4.6	8	-73	6	56	9	Amino_oxidase	67	105	Amine oxidase	1.80E-05
3	Locus1885v1rpkm148.74_10	4.4	2	-7.1	6	39	2	Aminotran_5	12	324	Aminotransferase, class V/Cysteine desulfurase	1.80E-31
3	Locus314v1rpkm527.22_10	4.6	2	-9	9	27	2	Aminotran_5	7	204	Aminotransferase, class V/Cysteine desulfurase	4.00E-23
1	Locus15904v1rpkm15.85_14	5.1	12	-93	9	83	12	AMP-binding	120	634	AMP-dependent synthetase/ligase	1.10E-97
6	Locus18516v1rpkm12.57_9	4.8	2	-11	6	35	2	AMP-binding	107	320	AMP-dependent synthetase/ligase	8.30E-48
1	Locus4901v1rpkm61.89_7	2.8	1	-2.9	6	39	2	AMP-binding	107	362	AMP-dependent synthetase/ligase	1.50E-52
6	Locus5588v1rpkm54.58_15	4.9	8	-74	7	76	20	AMP-binding	107	588	AMP-dependent synthetase/ligase	8.70E-109

5	Locus5604v1rpkm54.42_6	3.7	1	-1.3	6	38	1	Ank_2	69	179	Ankyrin repeat-containing domain	1.50E-15
5	Locus5604v1rpkm54.42_6	3.7	1	-1.3	6	38	1	Ank_2	190	252	Ankyrin repeat-containing domain	2.80E-13
5	Locus5604v1rpkm54.42_6	3.7	1	-1.3	6	38	1	Ank_2	258	314	Ankyrin repeat-containing domain	2.50E-10
5	Locus10566v1rpkm27.04_8	4.3	2	-9.7	8	35	3	Annexin	253	308	Annexin repeat	1.30E-09
5	Locus10566v1rpkm27.04_8	4.3	2	-9.7	8	35	3	Annexin	109	156	Annexin repeat	2.60E-07
5	Locus10566v1rpkm27.04_8	4.3	2	-9.7	8	35	3	Annexin	176	233	Annexin repeat	5.70E-12
4	Locus5131v1rpkm59.31_5	4.9	5	-41	9	35	7	Annexin	171	236	Annexin repeat	1.10E-20
4	Locus5131v1rpkm59.31_5	4.9	5	-41	9	35	7	Annexin	87	151	Annexin repeat	1.70E-10
4	Locus5131v1rpkm59.31_5	4.9	5	-41	9	35	7	Annexin	15	79	Annexin repeat	2.30E-17
4	Locus5131v1rpkm59.31_5	4.9	5	-41	9	35	7	Annexin	246	311	Annexin repeat	6.90E-25
4	Locus54257v1rpkm0.78_8	4.4	3	-26	7	31	3	Annexin	112	158	Annexin repeat	1.10E-07
4	Locus54257v1rpkm0.78_8	4.4	3	-26	7	31	3	Annexin	177	233	Annexin repeat	3.40E-12
3	Locus5905v1rpkm51.42_6	5.9	16	-114	7	35	21	AP_endonuc_2	130	300	Xylose isomerase, TIM barrel domain	3.50E-17
3	Locus6036v1rpkm50.33_8	5.7	11	-47	5	32	20	AP_endonuc_2	12	166	Xylose isomerase, TIM barrel domain	2.50E-19
3	Locus10021v1rpkm28.89_17	3.5	1	-2	7	92	1	APH	43	278	Aminoglycoside phosphotransferase	1.70E-41
6	Locus1595v1rpkm170.76_3	5.6	9	-86	6	21	9	Arf	7	177	Small GTPase superfamily, ARF/SAR type	1.00E-79
6	Locus2665v1rpkm111.45_4	5.1	6	-55	7	22	6	Arf	10	192	Small GTPase superfamily, ARF/SAR type	6.50E-65
6	Locus6977v1rpkm43.27_2	5.4	10	-91	8	21	10	Arf	12	178	Small GTPase superfamily, ARF/SAR type	8.70E-43
6	Locus7659v1rpkm39.18_2	5.1	9	-73	8	21	9	Arf	12	178	Small GTPase superfamily, ARF/SAR type	9.90E-43
6	Locus36053v1rpkm2.76_2	5.1	1	-1.3	6	18	1	ARPC4	1	149	ARP23 complex 20kDa subunit	2.10E-67
5	Locus2922v1rpkm101.38_11	5.6	7	-76	6	44	7	Asp	82	321	Peptidase A1	1.20E-90
5	Locus5529v1rpkm55.05_4	5.1	2	-17	9	20	2	Asp	82	177	Peptidase A1	4.00E-38
4	Locus5615v1rpkm54.33_3	3.5	1	-9.1	9	48	1	Asp	270	434	Peptidase A1	9.10E-09
4	Locus5615v1rpkm54.33_3	3.5	1	-9.1	9	48	1	Asp	72	122	Peptidase A1	1.20E-06
6	Locus5869v1rpkm51.81_10	4.5	1	-4.3	5	33	1	Asp	1	297	Peptidase A1	6.20E-66
4	Locus711v1rpkm308.68_7	4.8	3	-16	5	36	3	Asp	4	248	Peptidase A1	1.40E-100
6	Locus5445v1rpkm55.95_4	4.9	5	-47	7	40	5	ATP-synt	45	365	ATPase, F1 complex, gamma subunit	1.20E-96
4	Locus8537v1rpkm34.85_4	5.5	13	-128	10	27	13	ATP-synt	43	224	ATPase, F1 complex, gamma subunit	5.70E-39
6	Locus9562v1rpkm30.59_6	4.9	3	-29	10	20	4	ATP-synt	1	169	ATPase, F1 complex, gamma subunit	2.90E-36
4	Locus17835v1rpkm13.33_5	5	3	-21	9	60	3	ATP-synt_ab	190	414	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	7.60E-70
4	Locus18159v1rpkm12.96_9	6.7	86	-433	6	51	306	ATP-synt_ab	68	295	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	5.00E-111
3	Locus1176v1rpkm213.98_7	7.1	56	-220	5	29	200	ATP-synt_ab	147	267	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	1.10E-18
3	Locus15609v1rpkm16.29_10	6.7	42	-243	6	59	143	ATP-synt_ab	206	428	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	3.20E-61
4	Locus22887v1rpkm8.69_7	6.6	75	-412	5	48	169	ATP-synt_ab	229	443	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	1.00E-101
4	Locus353v1rpkm492.72_9	7.1	170	-870	6	62	599	ATP-synt_ab	229	456	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	1.10E-110
3	Locus3732v1rpkm80.28_11	7.2	94	-368	5	51	396	ATP-synt_ab	146	377	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	2.40E-61
2	Locus56806v1rpkm0.70_7	6.1	7	-47	5	57	11	ATP-synt_ab	143	374	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	8.90E-60
2	Locus59514v1rpkm0.63_4	4.2	1	-7.9	8	27	2	ATP-synt_ab	1	120	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	7.60E-40
5	Locus60839v1rpkm0.60_5	5.1	2	-16	6	35	2	ATP-synt_ab	173	332	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	1.10E-42
5	Locus62532v1rpkm0.58_5	4.8	2	-12	9	51	2	ATP-synt_ab	105	329	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	2.80E-71
5	Locus993v1rpkm242.42_6	6.2	21	-224	5	34	38	ATP-synt_ab	3	191	ATPase, F1/V1/A1 complex, alpha/beta subunit, nucleotide-binding domain	5.60E-47
4	Locus17835v1rpkm13.33_5	5	3	-21	9	60	3	ATP-synt_ab_C	426	525	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	5.20E-26
4	Locus18159v1rpkm12.96_9	6.7	86	-433	6	51	306	ATP-synt_ab_C	315	456	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	4.90E-28
3	Locus15609v1rpkm16.29_10	6.7	42	-243	6	59	143	ATP-synt_ab_C	442	545	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	8.70E-26
4	Locus353v1rpkm492.72_9	7.1	170	-870	6	62	599	ATP-synt_ab_C	476	555	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	1.60E-20

3	Locus3732v1rpkm80.28_11	7.2	94	-368	5	51	396	ATP-synt_ab_C	395	449	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	1.40E-11
2	Locus56806v1rpkm0.70_7	6.1	7	-47	5	57	11	ATP-synt_ab_C	392	486	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	1.10E-17
2	Locus59514v1rpkm0.63_4	4.2	1	-7.9	8	27	2	ATP-synt_ab_C	132	232	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	3.80E-28
5	Locus62532v1rpkm0.58_5	4.8	2	-12	9	51	2	ATP-synt_ab_C	341	439	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	6.20E-26
5	Locus993v1rpkm242.42_6	6.2	21	-224	5	34	38	ATP-synt_ab_C	205	308	ATPase, F1/V1/A1 complex, alpha/beta subunit, C-terminal	1.60E-25
2	Locus17463v1rpkm13.77_6	6.2	32	-201	5	26	32	ATP-synt_ab_N	23	83	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	1.80E-14
4	Locus17835v1rpkm13.33_5	5	3	-21	9	60	3	ATP-synt_ab_N	69	134	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	2.80E-15
3	Locus1176v1rpkm213.98_7	7.1	56	-220	5	29	200	ATP-synt_ab_N	25	91	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	3.50E-13
3	Locus15609v1rpkm16.29_10	6.7	42	-243	6	59	143	ATP-synt_ab_N	84	150	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	2.70E-21
4	Locus22887v1rpkm8.69_7	6.6	75	-412	5	48	169	ATP-synt_ab_N	23	83	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	4.80E-14
3	Locus2786v1rpkm106.74_2	4.3	2	-14	10	14	4	ATP-synt_ab_N	88	131	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	9.40E-10
4	Locus353v1rpkm492.72_9	7.1	170	-870	6	62	599	ATP-synt_ab_N	23	83	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	3.50E-13
3	Locus3732v1rpkm80.28_11	7.2	94	-368	5	51	396	ATP-synt_ab_N	24	90	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	8.30E-13
2	Locus56806v1rpkm0.70_7	6.1	7	-47	5	57	11	ATP-synt_ab_N	24	87	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	6.20E-12
5	Locus60839v1rpkm0.60_5	5.1	2	-16	6	35	2	ATP-synt_ab_N	50	117	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	1.80E-24
5	Locus62532v1rpkm0.58_5	4.8	2	-12	9	51	2	ATP-synt_ab_N	1	49	ATPase, F1/V1/A1 complex, alpha/beta subunit, N-terminal	8.30E-09
1	Locus1862v1rpkm150.03_2	5.4	3	-7.6	8	9.8	9	ATP-synt_C	13	77	ATPase, F0/V0 complex, subunit C	2.80E-14
6	Locus1902v1rpkm147.70_4	5.4	4	-16	9	14	4	ATP-synt_C	66	130	ATPase, F0/V0 complex, subunit C	1.10E-18
6	Locus1902v1rpkm147.70_4	5.4	4	-16	9	14	4	ATP-synt_C	1	51	ATPase, F0/V0 complex, subunit C	9.40E-11
5	Locus820v1rpkm280.10_4	6.7	34	-210	10	29	79	ATP-synt_D	17	213	ATPase, V1/A1 complex, subunit D	5.40E-65
6	Locus11016v1rpkm25.61_3	5.1	4	-28	6	22	4	ATP-synt_DE_N	73	151	ATPase, F1 complex, delta/epsilon subunit, N-terminal	1.30E-18
6	Locus5680v1rpkm53.67_1	4.5	3	-14	6	22	3	ATP-synt_DE_N	76	156	ATPase, F1 complex, delta/epsilon subunit, N-terminal	8.70E-20
6	Locus2724v1rpkm108.98_2	6.3	22	-171	6	15	22	ATP-synt_F	15	115	ATPase, V1/A1 complex, subunit F	2.30E-29
2	Locus32140v1rpkm3.94_26	4.3	2	-7.5	6	89	2	B_lectin	80	190	Bulb-type lectin domain	2.50E-26
4	Locus33011v1rpkm3.65_14	2.9	1	-1.2	6	90	1	B_lectin	68	179	Bulb-type lectin domain	2.90E-30
1	Locus983v1rpkm243.95_3	4.9	4	-24	5	23	6	B_lectin	59	134	Bulb-type lectin domain	6.00E-14
5	Locus49385v1rpkm1.00_10	3.9	1	-1.3	9	49	1	B3	319	408	Transcriptional factor B3	1.90E-17
2	Locus10465v1rpkm27.38_9	4	3	-24	7	30	8	Bac_surface_Ag	18	266	Bacterial surface antigen (D15)	3.80E-23
3	Locus17336v1rpkm13.92_5	5.3	4	-23	6	41	4	Band_7	55	237	Band 7 protein	5.20E-19
3	Locus10949v1rpkm25.81_4	5.4	6	-37	6	40	12	Band_7	55	249	Band 7 protein	3.60E-18
5	Locus3381v1rpkm88.63_6	5.4	10	-85	6	32	11	Band_7	9	182	Band 7 protein	1.50E-32
5	Locus4248v1rpkm70.84_6	4.8	5	-43	5	32	5	Band_7	9	182	Band 7 protein	1.80E-33
5	Locus7543v1rpkm39.91_3	4.8	4	-41	9	31	4	Band_7	34	212	Band 7 protein	3.10E-23
5	Locus8742v1rpkm33.92_6	4.5	2	-12	6	33	2	Band_7	12	183	Band 7 protein	1.10E-26
6	Locus32306v1rpkm3.87_4	3.5	1	-7	9	15	1	Barwin	22	140	Barwin	2.50E-62
3	Locus14282v1rpkm18.42_10	4	1	-2.3	8	60	1	Biotin_lipoyl	133	205	Biotin/lipoyl attachment	1.80E-19
1	Locus12276v1rpkm22.35_7	3.9	3	-19	8	40	3	C2	208	289	C2 calcium-dependent membrane targeting	7.80E-26
3	Locus27574v1rpkm5.85_7	4.5	1	-3.5	5	42	1	C2	242	323	C2 calcium-dependent membrane targeting	5.80E-25
3	Locus27574v1rpkm5.85_7	4.5	1	-3.5	5	42	1	C2	66	149	C2 calcium-dependent membrane targeting	4.70E-22
2	Locus28456v1rpkm5.42_11	4.2	2	-9.1	6	68	2	C2	463	555	C2 calcium-dependent membrane targeting	6.30E-15
2	Locus30619v1rpkm4.49_8	3.8	2	-8.6	6	67	2	C2	474	553	C2 calcium-dependent membrane targeting	1.70E-13
1	Locus44623v1rpkm1.36_16	3.8	1	-8	9	90	1	C2	60	139	C2 calcium-dependent membrane targeting	4.70E-20
1	Locus44623v1rpkm1.36_16	3.8	1	-8	9	90	1	C2	225	305	C2 calcium-dependent membrane targeting	1.20E-11
1	Locus44623v1rpkm1.36_16	3.8	1	-8	9	90	1	C2	386	475	C2 calcium-dependent membrane targeting	9.80E-14
3	Locus6944v1rpkm43.47_7	5.1	4	-25	8	54	6	C2	265	346	C2 calcium-dependent membrane targeting	1.30E-22
3	Locus7595v1rpkm39.60_16	4.7	2	-13	9	105	5	C2	7	91	C2 calcium-dependent membrane targeting	1.10E-14
3	Locus7595v1rpkm39.60_16	4.7	2	-13	9	105	5	C2	593	682	C2 calcium-dependent membrane targeting	7.30E-14
3	Locus7595v1rpkm39.60_16	4.7	2	-13	9	105	5	C2	270	349	C2 calcium-dependent membrane targeting	2.80E-18

3	Locus7595v1rpkm39.60_16	4.7	2	-13	9	105	5	C2	432	512	C2 calcium-dependent membrane targeting	1.60E-15
1	Locus18098v1rpkm13.03_21	4	1	-11	7	114	1	CaATP_NAI	5	51	Calcium-transporting ATPase, N-terminal autoinhibitory domain	1.00E-19
6	Locus2690v1rpkm110.38_2	4.4	2	-20	7	14	2	Caleosin	29	125	Caleosin	5.30E-41
6	Locus29743v1rpkm4.84_4	4.6	1	-13	9	16	1	Caleosin	2	131	Caleosin	6.80E-56
3	Locus14079v1rpkm18.78_6	5.2	7	-43	6	37	7	Calreticulin	6	221	Calreticulin/calnexin	9.40E-77
3	Locus2021v1rpkm140.83_7	6.3	45	-279	5	48	52	Calreticulin	25	337	Calreticulin/calnexin	6.30E-117
2	Locus2157v1rpkm133.47_5	6.1	18	-136	5	45	21	Calreticulin	1	250	Calreticulin/calnexin	1.80E-108
3	Locus22451v1rpkm8.98_4	5.2	8	-33	8	18	8	Calreticulin	37	154	Calreticulin/calnexin	4.30E-36
3	Locus2472v1rpkm119.13_8	5.8	18	-116	6	40	18	Calreticulin	26	337	Calreticulin/calnexin	1.70E-109
2	Locus3478v1rpkm86.23_4	6	12	-104	6	23	19	Calreticulin	28	197	Calreticulin/calnexin	2.80E-50
2	Locus5480v1rpkm55.55_3	5.9	10	-70	6	24	13	Calreticulin	33	205	Calreticulin/calnexin	3.40E-51
3	Locus8919v1rpkm33.16_4	5.8	14	-76	6	27	14	Calreticulin	2	138	Calreticulin/calnexin	1.80E-44
2	Locus3481v1rpkm86.15_8	5.6	8	-53	7	57	12	Catalase	18	400	Catalase, N-terminal	8.80E-181
2	Locus78v1rpkm1091.78_6	5.4	10	-69	7	34	18	Catalase	1	198	Catalase, N-terminal	1.30E-87
2	Locus3481v1rpkm86.15_8	5.6	8	-53	7	57	12	Catalase-rel	421	487	Catalase-related immune responsive	4.50E-22
2	Locus78v1rpkm1091.78_6	5.4	10	-69	7	34	18	Catalase-rel	221	286	Catalase-related immune responsive	1.70E-19
1	Locus18098v1rpkm13.03_21	4	1	-11	7	114	1	Cation_ATPase_C	837	1010	ATPase, P-type cation-transporter, C-terminal	2.90E-42
1	Locus18098v1rpkm13.03_21	4	1	-11	7	114	1	Cation_ATPase_N	118	183	ATPase, P-type cation-transporter, N-terminal	4.30E-14
3	Locus11745v1rpkm23.64_6	4.2	1	-1.2	5	34	1	Cation_ATPase_N	24	91	ATPase, P-type cation-transporter, N-terminal	1.00E-20
1	Locus12344v1rpkm22.20_6	4.5	6	-56	5	37	6	Cation_ATPase_N	20	83	ATPase, P-type cation-transporter, N-terminal	1.70E-13
1	Locus24602v1rpkm7.55_4	4.5	5	-42	5	26	8	Cation_ATPase_N	19	82	ATPase, P-type cation-transporter, N-terminal	4.00E-15
6	Locus3426v1rpkm87.37_6	5.2	5	-41	5	28	5	Cation_ATPase_N	20	83	ATPase, P-type cation-transporter, N-terminal	1.40E-13
1	Locus1594v1rpkm170.83_14	4.6	7	-56	5	80	10	CDC48_2	44	107	Cell division protein 48, Cdc48, domain 2	8.10E-12
3	Locus1855v1rpkm150.41_6	6.1	27	-215	9	54	73	Chal_sti_synt_C	423	477	Chalcone/stilbene synthase, C-terminal	1.10E-06
3	Locus33313v1rpkm3.54_4	5.9	9	-42	8	49	9	Chal_sti_synt_C	328	384	Chalcone/stilbene synthase, C-terminal	2.50E-09
3	Locus944v1rpkm252.21_6	6.1	17	-94	9	48	28	Chal_sti_synt_C	328	386	Chalcone/stilbene synthase, C-terminal	4.40E-10
6	Locus22322v1rpkm9.07_4	5	6	-43	10	30	6	Chalcone	100	275	Chalcone isomerase, subgroup	4.30E-32
5	Locus1000v1rpkm241.44_3	5.6	6	-68	6	28	6	Chloroa_b-bind	65	231	Chlorophyll A-B binding protein	1.40E-50
5	Locus104v1rpkm930.88_3	5.7	9	-93	5	31	9	Chloroa_b-bind	92	253	Chlorophyll A-B binding protein	5.10E-46
6	Locus2182v1rpkm132.22_3	5.4	6	-66	6	16	6	Chloroa_b-bind	1	114	Chlorophyll A-B binding protein	6.40E-23
6	Locus227v1rpkm649.00_3	5.4	4	-42	9	30	6	Chloroa_b-bind	65	243	Chlorophyll A-B binding protein	3.80E-50
6	Locus2331v1rpkm125.32_2	4.1	1	-11	6	23	1	Chloroa_b-bind	71	207	Chlorophyll A-B binding protein	3.60E-36
6	Locus2826v1rpkm105.19_3	3.4	1	-1.9	10	21	1	Chloroa_b-bind	90	187	Chlorophyll A-B binding protein	3.50E-05
5	Locus368v1rpkm472.83_3	5.6	7	-76	5	26	10	Chloroa_b-bind	38	204	Chlorophyll A-B binding protein	7.10E-51
6	Locus469v1rpkm404.33_4	5.2	4	-38	5	22	4	Chloroa_b-bind	12	166	Chlorophyll A-B binding protein	2.30E-47
6	Locus708v1rpkm309.43_6	4	1	-9.4	5	21	1	Chloroa_b-bind	67	197	Chlorophyll A-B binding protein	6.10E-35
5	Locus76v1rpkm1101.34_2	5.4	6	-65	6	25	8	Chloroa_b-bind	65	231	Chlorophyll A-B binding protein	7.00E-51
4	Locus21574v1rpkm9.65_33	4.1	1	-1.4	8	220	1	CLASP_N	299	477	CLASP N-terminal domain	2.00E-11
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin	557	690	Clathrin, heavy chain/VPS, 7-fold repeat	4.00E-20
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin	701	840	Clathrin, heavy chain/VPS, 7-fold repeat	4.20E-21
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin	850	983	Clathrin, heavy chain/VPS, 7-fold repeat	5.80E-30
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin	993	1133	Clathrin, heavy chain/VPS, 7-fold repeat	1.70E-30
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin	1145	1281	Clathrin, heavy chain/VPS, 7-fold repeat	6.70E-27
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin	1288	1431	Clathrin, heavy chain/VPS, 7-fold repeat	1.30E-33
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin	1437	1511	Clathrin, heavy chain/VPS, 7-fold repeat	6.60E-13
2	Locus9208v1rpkm32.01_7	4.1	3	-14	5	29	4	Clathrin	2	127	Clathrin, heavy chain/VPS, 7-fold repeat	4.70E-27
6	Locus18806v1rpkm12.26_5	3.7	2	-9	5	40	2	Clathrin_lg_ch	106	273	Clathrin light chain	2.70E-12
5	Locus3194v1rpkm93.28_3	4.2	1	-5.1	10	21	1	Clathrin_lg_ch	9	139	Clathrin light chain	1.50E-07
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin_propel	154	197	Clathrin, heavy chain, propeller repeat	2.00E-08
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin_propel	20	55	Clathrin, heavy chain, propeller repeat	3.50E-05
2	Locus2100v1rpkm136.32_28	5.7	26	-281	5	171	100	Clathrin-link	344	367	Clathrin, heavy chain, linker, core motif	7.70E-10
6	Locus2703v1rpkm109.97_12	3.8	1	-4.7	5	92	1	Clp_N	93	143	Clp, N-terminal	8.90E-19
6	Locus2703v1rpkm109.97_12	3.8	1	-4.7	5	92	1	Clp_N	16	68	Clp, N-terminal	3.30E-16
6	Locus2703v1rpkm109.97_12	3.8	1	-4.7	5	92	1	ClpB_D2-small	725	810	Clp ATPase, C-terminal	3.20E-23
3	Locus8746v1rpkm33.91_6	4.5	2	-6.9	5	41	3	CLPTM1	2	353	Cleft lip and palate transmembrane 1	4.80E-112
4	Locus18233v1rpkm12.87_8	5.4	2	-7.2	5	38	2	CN_hydrolase	35	220	Nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase	2.00E-39

2	Locus14187v1rpkm18.59_9	4.3	1	-2.2	9	37	1	cNMP_binding	105	200	Cyclic nucleotide-binding domain	1.20E-09
5	Locus44368v1rpkm1.39_6	4.2	1	-4.1	7	53	1	Complex1_49kDa	198	468	NADH-quinone oxidoreductase, subunit D	2.30E-133
3	Locus4274v1rpkm70.56_11	4.6	1	-4.1	8	27	1	Complex1_51K	115	163	NADH:ubiquinone oxidoreductase, 51kDa subunit	5.30E-08
3	Locus4274v1rpkm70.56_11	4.6	1	-4.1	8	27	1	Complex1_51K	1	109	NADH:ubiquinone oxidoreductase, 51kDa subunit	1.70E-19
3	Locus4274v1rpkm70.56_11	4.6	1	-4.1	8	27	1	Complex1_51K	169	227	NADH:ubiquinone oxidoreductase, 51kDa subunit	8.80E-11
3	Locus7086v1rpkm42.54_9	4.5	1	-2.4	7	40	1	COPIIcoated_ERV	154	328	Domain of unknown function DUF1692	7.00E-46
6	Locus77v1rpkm1096.15_1	3.4	1	-3.2	5	17	1	Copper-bind	70	167	Blue (type 1) copper domain	1.60E-35
6	Locus11090v1rpkm25.35_4	4.9	3	-23	6	19	3	COX5B	65	150	Cytochrome c oxidase, subunit Vb	1.00E-22
5	Locus8689v1rpkm34.13_4	3.3	1	-6.7	4	21	1	COX6B	128	188	Cytochrome c oxidase, subunit VIb	3.80E-24
3	Locus245v1rpkm613.45_12	5.4	15	-91	9	48	17	CP12	431	452	Domain of unknown function CP12	6.70E-07
2	Locus1982v1rpkm143.62_9	4.7	6	-66	5	44	7	Cpn60_TCP1	4	393	Chaperonin Cpn60/TCP-1	2.50E-80
2	Locus12068v1rpkm22.82_7	3.5	1	-3.4	5	25	1	Cpn60_TCP1	2	218	Chaperonin Cpn60/TCP-1	1.80E-50
2	Locus2496v1rpkm117.85_10	4.7	7	-58	5	61	14	Cpn60_TCP1	63	566	Chaperonin Cpn60/TCP-1	1.90E-120
2	Locus2759v1rpkm107.56_7	4	2	-14	6	49	2	Cpn60_TCP1	31	453	Chaperonin Cpn60/TCP-1	8.20E-126
2	Locus3428v1rpkm87.26_7	4.3	5	-48	6	51	8	Cpn60_TCP1	86	481	Chaperonin Cpn60/TCP-1	6.20E-90
6	Locus34356v1rpkm3.22_39	3.8	1	-1.1	6	200	1	Cpn60_TCP1	404	640	Chaperonin Cpn60/TCP-1	1.00E-29
3	Locus6941v1rpkm43.48_8	3.7	1	-1.1	5	46	2	Cpn60_TCP1	63	438	Chaperonin Cpn60/TCP-1	9.70E-79
4	Locus6995v1rpkm43.14_10	3.8	1	-8.1	6	59	3	Cpn60_TCP1	40	532	Chaperonin Cpn60/TCP-1	1.30E-148
2	Locus7248v1rpkm41.60_10	4.4	2	-9	5	59	2	Cpn60_TCP1	38	529	Chaperonin Cpn60/TCP-1	1.50E-134
3	Locus19028v1rpkm12.03_7	3.6	1	-1.3	9	53	1	CRAL_TRIO	271	388	Cellular retinaldehyde-binding/triple function, C-terminal	1.30E-17
3	Locus3435v1rpkm87.12_4	5.9	13	-85	5	57	36	CRAL_TRIO	263	392	Cellular retinaldehyde-binding/triple function, C-terminal	5.40E-19
3	Locus5469v1rpkm55.69_2	5.8	10	-64	9	24	17	CRAL_TRIO	2	95	Cellular retinaldehyde-binding/triple function, C-terminal	1.00E-16
3	Locus3435v1rpkm87.12_4	5.9	13	-85	5	57	36	CRAL_TRIO_N	178	206	Cellular retinaldehyde-binding/triple function, N-terminal	5.20E-08
1	Locus4370v1rpkm69.32_2	4.6	4	-22	4	47	4	CRAL_TRIO_N	280	309	Cellular retinaldehyde-binding/triple function, N-terminal	3.00E-08
4	Locus15048v1rpkm17.16_7	3.6	1	-3.1	5	25	1	CTP_transf_3	1	186	Acylneuraminate cytidylyltransferase	6.80E-40
1	Locus5191v1rpkm58.52_6	3.5	1	-1.3	6	55	1	Cu_amine_oxid	57	482	Copper amine oxidase, C-terminal	7.20E-144
1	Locus28768v1rpkm5.27_1	3.5	1	-1.3	10	20	1	Cu_amine_oxidN2	64	142	Copper amine oxidase, N2-terminal	1.80E-20
1	Locus5191v1rpkm58.52_6	3.5	1	-1.3	6	55	1	Cu_amine_oxidN3	1	31	Copper amine oxidase, N3-terminal	2.10E-06
1	Locus1513v1rpkm178.74_3	4.9	7	-84	9	23	7	Cu_bind_like	35	113	Plastocyanin-like	5.60E-19
1	Locus15089v1rpkm17.10_13	4.7	5	-56	9	45	5	Cu-oxidase	1	126	Multicopper oxidase, type 1	5.10E-33
1	Locus24383v1rpkm7.69_5	4.3	2	-21	5	22	2	Cu-oxidase	161	201	Multicopper oxidase, type 1	4.30E-06
1	Locus15089v1rpkm17.10_13	4.7	5	-56	9	45	5	Cu-oxidase_2	227	346	Multicopper oxidase, type 2	3.20E-25
1	Locus24383v1rpkm7.69_5	4.3	2	-21	5	22	2	Cu-oxidase_3	34	148	Multicopper oxidase, type 3	1.40E-40
4	Locus7424v1rpkm40.55_9	3.9	1	-4.3	5	31	1	Cupin_2	227	273	Cupin 2, conserved barrel	4.70E-05
3	Locus20070v1rpkm10.97_10	3.7	1	-1.3	7	66	1	CwfJ_C_1	380	493	Cwf19-like, C-terminal domain-1	8.40E-35
3	Locus20070v1rpkm10.97_10	3.7	1	-1.3	7	66	1	CwfJ_C_2	512	598	Cwf19-like protein, C-terminal domain-2	1.40E-20
6	Locus17793v1rpkm13.38_2	5.5	5	-59	5	15	5	Cyt-b5	8	81	Cytochrome b5	2.10E-28
6	Locus2673v1rpkm111.04_4	5.9	9	-57	5	15	9	Cyt-b5	7	81	Cytochrome b5	4.60E-28
6	Locus4298v1rpkm70.22_5	5.7	8	-52	5	15	8	Cyt-b5	7	81	Cytochrome b5	1.50E-28
5	Locus812v1rpkm284.46_2	4.1	2	-15	5	22	4	Cyt-b5	72	167	Cytochrome b5	3.00E-17
4	Locus855v1rpkm273.32_3	4.5	2	-13	5	25	2	Cyt-b5	72	167	Cytochrome b5	1.70E-17
6	Locus4450v1rpkm68.08_5	4.9	2	-7.9	9	24	2	CytB6-F_Fe-S	57	95	Cytochrome b6-f complex Fe-S subunit	5.00E-19
1	Locus2947v1rpkm100.72_4	3.7	1	-4.5	6	20	1	Cytochrom_B561	1	128	Cytochrome b561, eukaryote	2.20E-46
5	Locus6285v1rpkm48.52_6	4.5	4	-32	6	34	5	Cytochrom_C1	77	293	Cytochrome c1	2.20E-85
5	Locus6807v1rpkm44.51_5	4.4	3	-22	7	22	3	Cytochrom_C1	76	198	Cytochrome c1	5.60E-50
3	Locus6903v1rpkm43.81_9	3.8	1	-1.4	7	54	1	DAGK_acc	289	466	Diacylglycerol kinase, accessory domain	7.00E-48
3	Locus6903v1rpkm43.81_9	3.8	1	-1.4	7	54	1	DAGK_cat	99	243	Diacylglycerol kinase, catalytic domain	1.10E-25
3	Locus704v1rpkm311.52_7	4.8	2	-13	6	43	2	DEAD	65	226	DNA/RNA helicase, DEAD/DEAH box type, N-terminal	5.40E-43
4	Locus2635v1rpkm112.64_8	3.9	2	-7.5	6	40	2	Dimerisation	32	79	Plant methyltransferase dimerisation	2.40E-11
3	Locus19955v1rpkm11.08_8	4.8	1	-1.1	6	42	1	DJ-1_PfpI	250	380	ThiJ/PfpI	7.80E-33
3	Locus19955v1rpkm11.08_8	4.8	1	-1.1	6	42	1	DJ-1_PfpI	57	188	ThiJ/PfpI	2.10E-32
4	Locus10462v1rpkm27.39_6	3.7	1	-6.8	5	39	1	DnaJ	6	68	Heat shock protein DnaJ, N-terminal	3.80E-22
4	Locus26901v1rpkm6.23_7	4.3	2	-21	5	29	2	DnaJ	26	88	Heat shock protein DnaJ, N-terminal	9.50E-30
4	Locus37074v1rpkm2.53_5	4.5	1	-2.2	10	35	1	DnaJ	66	127	Heat shock protein DnaJ, N-terminal	2.90E-21

4	Locus46454v1rpkm1.19_6	4.1	1	-1.3	9	32	1	DnaJ	217	271	Heat shock protein DnaJ, N-terminal	1.20E-09
6	Locus16277v1rpkm15.29_17	3.3	1	-2.4	9	95	1	DUF1012	385	578	CASTOR/POLLUX/SYM8 ion channels	3.20E-71
6	Locus13758v1rpkm19.36_1	4.5	1	-4.1	10	25	1	DUF106	47	219	Protein of unknown function DUF106, transmembrane	4.50E-42
5	Locus8808v1rpkm33.67_3	4.5	3	-12	9	24	5	DUF106	8	156	Protein of unknown function DUF106, transmembrane	2.10E-39
2	Locus48459v1rpkm1.05_5	3.8	1	-5	10	34	2	DUF1191	30	307	Protein of unknown function DUF1191	6.60E-100
3	Locus5950v1rpkm50.97_13	3.8	1	-1.1	7	67	3	DUF1620	374	592	Domain of unknown function DUF1620	4.30E-65
1	Locus5359v1rpkm56.83_7	4.5	6	-59	6	58	9	DUF1982	489	525	NADH-quinone oxidoreductase, chain G, C-terminal	3.10E-13
1	Locus12729v1rpkm21.39_14	3.4	1	-4.2	9	77	1	DUF221	302	618	Domain of unknown function DUF221	5.50E-87
1	Locus4330v1rpkm69.89_13	4.1	3	-17	9	82	5	DUF221	318	639	Domain of unknown function DUF221	1.30E-97
1	Locus5639v1rpkm54.11_12	4.2	4	-28	10	68	6	DUF221	317	607	Domain of unknown function DUF221	2.10E-87
	Locus10050v1rpkm28.78_8	4.9	4	-24	9	63	5	DUF2359	298	562	Protein of unknown function DUF2359, TMEM214	5.70E-18
3	Locus7163v1rpkm42.10_10	4.8	3	-17	9	46	3	DUF2359	107	399	Protein of unknown function DUF2359, TMEM214	2.60E-19
1	Locus3906v1rpkm76.98_5	3.5	1	-1.1	9	28	1	DUF300	1	244	Protein of unknown function DUF300	5.40E-75
6	Locus15838v1rpkm15.95_9	3.9	1	-9.8	10	32	1	DUF3353	78	152	Protein of unknown function DUF3353	1.70E-07
2	Locus32140v1rpkm3.94_26	4.3	2	-7.5	6	89	2	DUF3403	766	811	S-locus receptor kinase, C-terminal	8.90E-07
1	Locus4551v1rpkm66.45_11	4.8	6	-52	8	70	9	DUF3406	385	639	Domain of unknown function DUF3406, chloroplast translocase	4.10E-129
5	Locus4269v1rpkm70.60_5	3.9	1	-1.2	6	27	1	DUF3700	1	119	Domain of unknown function DUF3700	3.70E-48
5	Locus4269v1rpkm70.60_5	3.9	1	-1.2	6	27	1	DUF3711	170	226	Domain of unknown function DUF3711	7.60E-34
6	Locus5534v1rpkm55.03_2	4.7	1	-2.5	5	19	1	DUF538	28	137	Protein of unknown function DUF538	2.10E-28
1	Locus817v1rpkm282.29_1	3.4	1	-3.9	7	18	1	DUF538	26	134	Protein of unknown function DUF538	1.80E-29
2	Locus27146v1rpkm6.09_18	5.7	1	-1.1	6	111	2	DUF863	140	984	Protein of unknown function DUF863, plant	1.20E-97
3	Locus6210v1rpkm49.00_4	4.2	2	-9.1	7	47	2	Dynamin_M	221	428	Dynamin central domain	8.50E-57
3	Locus6210v1rpkm49.00_4	4.2	2	-9.1	7	47	2	Dynamin_N	37	212	Dynamin, GTPase domain	3.20E-54
3	Locus5372v1rpkm56.69_10	4	1	-3.5	7	43	1	E1_dh	68	363	Dehydrogenase, E1 component	1.20E-114
1	Locus18098v1rpkm13.03_21	4	1	-11	7	114	1	E1-E2_ATPase	203	443	ATPase, P-type, ATPase-associated domain	9.20E-62
3	Locus11745v1rpkm23.64_6	4.2	1	-1.2	5	34	1	E1-E2_ATPase	115	299	ATPase, P-type, ATPase-associated domain	7.00E-46
1	Locus12344v1rpkm22.20_6	4.5	6	-56	5	37	6	E1-E2_ATPase	102	323	ATPase, P-type, ATPase-associated domain	2.40E-58
1	Locus24602v1rpkm7.55_4	4.5	5	-42	5	26	8	E1-E2_ATPase	101	243	ATPase, P-type, ATPase-associated domain	7.20E-42
1	Locus3118v1rpkm95.27_15	5.2	14	-125	7	90	17	E1-E2_ATPase	34	255	ATPase, P-type, ATPase-associated domain	1.10E-57
6	Locus3426v1rpkm87.37_6	5.2	5	-41	5	28	5	E1-E2_ATPase	102	257	ATPase, P-type, ATPase-associated domain	2.50E-44
1	Locus4144v1rpkm72.42_13	4.7	9	-62	6	59	9	E1-E2_ATPase	34	255	ATPase, P-type, ATPase-associated domain	2.00E-58
1	Locus9810v1rpkm29.54_11	4.7	7	-49	9	78	7	E1-E2_ATPase	8	65	ATPase, P-type, ATPase-associated domain	6.10E-08
3	Locus14282v1rpkm18.42_10	4	1	-2.3	8	60	1	E3_binding	266	302	E3 binding	1.80E-13
1	Locus1201v1rpkm210.03_5	5.7	5	-48	9	33	5	EamA	8	128	Drug/metabolite transporter	5.70E-05
1	Locus8441v1rpkm35.36_3	2.7	1	-4.5	10	30	1	EamA	2	105	Drug/metabolite transporter	1.60E-13
4	Locus41929v1rpkm1.68_4	3.6	1	-1.2	6	46	1	Ebp2	131	406	Eukaryotic rRNA processing	2.80E-92
3	Locus4728v1rpkm64.14_8	4.4	1	-5.7	6	48	1	EF1G	258	366	Translation elongation factor EF1B, gamma chain, conserved	1.90E-42
4	Locus5013v1rpkm60.63_4	3.1	1	-7.4	6	37	1	EF1G	161	269	Translation elongation factor EF1B, gamma chain, conserved	1.10E-42
6	Locus1777v1rpkm156.02_4	5	6	-59	4	17	6	efhand	48	75	EF-hand	1.30E-08
6	Locus1777v1rpkm156.02_4	5	6	-59	4	17	6	efhand	121	148	EF-hand	6.20E-10
6	Locus1777v1rpkm156.02_4	5	6	-59	4	17	6	efhand	85	113	EF-hand	1.30E-09
6	Locus1777v1rpkm156.02_4	5	6	-59	4	17	6	efhand	12	40	EF-hand	4.20E-09
6	Locus12228v1rpkm22.43_6	3.3	1	-4.4	5	24	1	efhand	140	166	EF-hand	6.90E-08
6	Locus12228v1rpkm22.43_6	3.3	1	-4.4	5	24	1	efhand	175	201	EF-hand	4.70E-09
6	Locus12228v1rpkm22.43_6	3.3	1	-4.4	5	24	1	efhand	68	94	EF-hand	7.80E-08
5	Locus27363v1rpkm5.96_7	4	1	-2.8	5	59	1	efhand	487	513	EF-hand	3.90E-07
5	Locus27363v1rpkm5.96_7	4	1	-2.8	5	59	1	efhand	454	478	EF-hand	8.50E-07
5	Locus27363v1rpkm5.96_7	4	1	-2.8	5	59	1	efhand	380	406	EF-hand	5.40E-08
2	Locus28456v1rpkm5.42_11	4.2	2	-9.1	6	68	2	efhand_like	24	102	Phospholipase C, phosphoinositol-specific, EF-hand-like	3.20E-11
2	Locus30619v1rpkm4.49_8	3.8	2	-8.6	6	67	2	efhand_like	25	99	Phospholipase C, phosphoinositol-specific, EF-hand-like	2.20E-16
4	Locus7955v1rpkm37.72_5	3.2	1	-3.2	5	39	1	EIF_2_alpha	130	261	Translation initiation factor 2, alpha subunit	5.70E-45
3	Locus24032v1rpkm7.92_10	4.7	1	-1.2	5	68	1	EIN3	41	411	Ethylene insensitive 3	2.50E-124

6	Locus18223v1rpkm12.88_5	4.9	2	-11	6	24	2	EMP24_GP25L	31	203	GOLD	1.80E-32
6	Locus11188v1rpkm25.04_2	3.3	1	-1.2	8	23	1	EMP24_GP25L	23	204	GOLD	1.10E-52
6	Locus7092v1rpkm42.51_4	4.7	2	-13	8	25	2	EMP24_GP25L	35	207	GOLD	3.10E-42
6	Locus8132v1rpkm36.94_2	4.9	4	-16	6	24	4	EMP24_GP25L	23	208	GOLD	1.20E-57
1	Locus4891v1rpkm61.99_14	3.9	2	-9.5	7	72	2	EMP70	55	589	Nonaspanin (TM9SF)	2.70E-225
3	Locus171v1rpkm738.63_7	6.5	42	-200	6	38	52	Enolase_C	56	349	Enolase, C-terminal	1.30E-163
3	Locus171v1rpkm738.63_7	6.5	42	-200	6	38	52	Enolase_N	3	47	Enolase, N-terminal	1.80E-10
4	Locus17454v1rpkm13.78_5	3.4	1	-1.2	9	35	1	Epimerase	6	236	NAD-dependent epimerase/dehydratase	1.10E-34
1	Locus2115v1rpkm135.58_5	4.3	1	-5	9	28	1	ERG4_ERG24	32	186	Ergosterol biosynthesis ERG4/ERG24	1.30E-25
6	Locus6886v1rpkm43.91_3	2.8	1	-4.9	5	19	1	ETC_C1_NDUFA5	32	88	ETC complex I subunit	1.20E-23
4	Locus74055v1rpkm0.42_6	4.7	2	-2.3	5	31	4	F_actin_cap_B	6	254	WASH complex, F-actin capping protein, beta subunit	5.60E-109
5	Locus2071v1rpkm137.76_11	4.3	2	-13	5	58	3	FAD_binding_1	298	521	FAD-binding, type 1	2.20E-80
4	Locus3301v1rpkm90.74_11	4.1	2	-13	6	69	2	FAD_binding_2	44	440	Fumarate reductase/succinate dehydrogenase flavoprotein, N-terminal	5.70E-124
2	Locus36448v1rpkm2.67_10	3.9	2	-13	7	71	3	FAD_binding_2	65	460	Fumarate reductase/succinate dehydrogenase flavoprotein, N-terminal	2.40E-123
3	Locus8608v1rpkm34.42_6	4.2	2	-9.8	8	44	2	FAD_binding_3	6	353	Monooxygenase, FAD-binding	8.10E-29
5	Locus9589v1rpkm30.46_3	5.2	5	-38	6	54	6	FAD_binding_4	14	103	FAD linked oxidase, N-terminal	5.20E-17
5	Locus16161v1rpkm15.46_3	3.3	1	-1.9	9	36	1	FAD_binding_6	74	177	Oxidoreductase, FAD-binding domain	9.60E-21
5	Locus12223v1rpkm22.44_5	4.1	1	-5.9	9	31	1	FAD_binding_6	48	146	Oxidoreductase, FAD-binding domain	2.10E-30
4	Locus4035v1rpkm74.07_7	4.3	3	-22	8	41	3	FAD_binding_6	141	204	Oxidoreductase, FAD-binding domain	5.40E-05
1	Locus1624v1rpkm168.72_4	4.6	6	-48	8	21	10	FAE1_CUT1_RppA	1	166	FAE1/Type III polyketide synthase-like protein	5.40E-72
3	Locus1855v1rpkm150.41_6	6.1	27	-215	9	54	73	FAE1_CUT1_RppA	116	404	FAE1/Type III polyketide synthase-like protein	1.10E-135
3	Locus33313v1rpkm3.54_4	5.9	9	-42	8	49	9	FAE1_CUT1_RppA	28	311	FAE1/Type III polyketide synthase-like protein	2.40E-112
3	Locus4251v1rpkm70.82_5	5.1	6	-54	9	52	15	FAE1_CUT1_RppA	27	311	FAE1/Type III polyketide synthase-like protein	4.70E-117
3	Locus4400v1rpkm68.84_2	5.3	5	-28	9	56	5	FAE1_CUT1_RppA	81	369	FAE1/Type III polyketide synthase-like protein	1.00E-139
1	Locus7676v1rpkm39.08_6	4	3	-27	9	59	13	FAE1_CUT1_RppA	109	398	FAE1/Type III polyketide synthase-like protein	1.30E-144
3	Locus8644v1rpkm34.26_6	5.2	4	-21	9	48	5	FAE1_CUT1_RppA	21	307	FAE1/Type III polyketide synthase-like protein	6.80E-109
3	Locus944v1rpkm252.21_6	6.1	17	-94	9	48	28	FAE1_CUT1_RppA	28	311	FAE1/Type III polyketide synthase-like protein	9.80E-108
1	Locus1680v1rpkm164.45_1	5.7	12	-101	6	35	16	Fasciclin	127	257	FAS1 domain	8.20E-20
2	Locus1884v1rpkm148.77_4	5.9	21	-148	7	44	36	Fasciclin	206	337	FAS1 domain	3.50E-14
2	Locus1884v1rpkm148.77_4	5.9	21	-148	7	44	36	Fasciclin	42	136	FAS1 domain	9.80E-06
2	Locus10967v1rpkm25.74_3	5.4	11	-94	7	35	16	Fasciclin	132	263	FAS1 domain	5.00E-14
2	Locus12608v1rpkm21.62_1	4.8	4	-28	9	12	4	Fasciclin	39	104	FAS1 domain	6.60E-05
2	Locus23460v1rpkm8.30_3	3.6	1	-3.8	5	17	1	Fasciclin	5	137	FAS1 domain	1.40E-19
1	Locus27096v1rpkm6.11_3	4	2	-19	8	27	2	Fasciclin	54	185	FAS1 domain	1.20E-22
2	Locus45277v1rpkm1.29_7	4.3	1	-1.2	9	46	1	FBA_3	236	323	F-box associated domain, type 3	4.30E-05
2	Locus45277v1rpkm1.29_7	4.3	1	-1.2	9	46	1	F-box	10	45	F-box domain, cyclin-like	2.30E-06
5	Locus8201v1rpkm36.59_4	4.2	1	-1.4	8	46	1	F-box	31	64	F-box domain, cyclin-like	1.60E-06
6	Locus1615v1rpkm169.12_8	3.7	1	-3	6	37	1	FBPase	13	335	Fructose-1,6-bisphosphatase class 1/Sedoheptulose-1,7-bisphosphatase	1.80E-135
4	Locus2610v1rpkm113.54_8	2.7	1	-4	8	42	1	FBPase	70	378	Fructose-1,6-bisphosphatase class 1/Sedoheptulose-1,7-bisphosphatase	8.90E-104
3	Locus8138v1rpkm36.91_9	4.3	1	-1.1	6	78	1	Fer2	74	134	Ferredoxin	1.70E-08
3	Locus17972v1rpkm13.18_16	4.2	1	-1.6	5	97	1	FG-GAP	490	516	FG-GAP	7.40E-05
3	Locus17972v1rpkm13.18_16	4.2	1	-1.6	5	97	1	FG-GAP	582	609	FG-GAP	5.00E-05
5	Locus21255v1rpkm9.90_5	4.2	1	-9.9	9	27	2	Fibrillarin	9	236	Fibrillarin	4.70E-114
6	Locus8304v1rpkm36.10_4	4.9	2	-15	6	16	2	FKBP_C	40	132	Peptidyl-prolyl cis-trans isomerase, FKBP-type, domain	6.70E-36
5	Locus2071v1rpkm137.76_11	4.3	2	-13	5	58	3	Flavodoxin_1	96	239	Flavodoxin/nitric oxide synthase	3.10E-37
5	Locus17149v1rpkm14.16_5	5.4	4	-33	9	36	4	FMN_dh	13	327	FMN-dependent dehydrogenase	4.60E-120
5	Locus27761v1rpkm5.76_7	3.7	1	-3.6	8	40	1	FMN_dh	15	353	FMN-dependent dehydrogenase	2.40E-134
6	Locus11295v1rpkm24.77_3	4.2	1	-9.4	6	20	1	FMN_red	55	133	NADPH-dependent FMN reductase	1.80E-12
5	Locus9675v1rpkm30.10_4	4.9	3	-20	6	28	4	FMN_red	122	200	NADPH-dependent FMN reductase	4.30E-12
6	Locus34356v1rpkm3.22_39	3.8	1	-1.1	6	200	1	FYVE	33	101	Zinc finger, FYVE-type	9.80E-18
6	Locus2257v1rpkm129.02_4	3.5	1	-5.1	5	18	1	GCV_H	42	161	Glycine cleavage H-protein	4.80E-50
1	Locus1267v1rpkm203.77_2	2.5	1	-1.3	6	29	1	GDC-P	1	272	Glycine cleavage system P-protein, N-terminal	3.90E-126
1	Locus17144v1rpkm14.16_5	5.2	10	-68	5	28	10	GDPD	7	152	Glycerophosphoryl diester phosphodiesterase	1.60E-10
1	Locus19233v1rpkm11.81_9	5.5	21	-162	5	62	24	GDPD	177	468	Glycerophosphoryl diester phosphodiesterase	2.40E-39
1	Locus24140v1rpkm7.84_14	5.7	22	-217	6	83	24	GDPD	367	658	Glycerophosphoryl diester phosphodiesterase	1.40E-38

1	Locus6907v1rpkm43.79_13	5.5	15	-116	5	63	15	GDPD	186	477	Glycerophosphoryl diester phosphodiesterase	1.70E-39
5	Locus2899v1rpkm102.49_3	2.6	1	-3.4	5	18	1	Gln-synt_C	104	165	Glutamine synthetase, catalytic domain	3.40E-06
3	Locus772v1rpkm294.37_7	3.8	1	-7	6	31	2	Gln-synt_C	74	278	Glutamine synthetase, catalytic domain	7.40E-48
3	Locus948v1rpkm251.59_4	4	1	-4.3	6	13	1	Gln-synt_C	1	121	Glutamine synthetase, catalytic domain	1.70E-30
5	Locus2899v1rpkm102.49_3	2.6	1	-3.4	5	18	1	Gln-synt_N	20	97	Glutamine synthetase, beta-Grasp	2.40E-20
3	Locus772v1rpkm294.37_7	3.8	1	-7	6	31	2	Gln-synt_N	3	67	Glutamine synthetase, beta-Grasp	2.60E-16
2	Locus29312v1rpkm5.02_3	3.1	1	-1.3	5	19	1	GLTP	1	118	Glycolipid transfer protein domain	1.10E-31
5	Locus178648v1rpkm0.00_3	4.3	1	-1.2	10	9.6	1	Glu_syn_central	39	71	Glutamate synthase, central-N	1.30E-13
2	Locus129v1rpkm840.43_5	5.9	13	-108	5	25	16	Glyco_hydro_1	2	208	Glycoside hydrolase, family 1	3.80E-60
2	Locus20518v1rpkm10.55_16	5.4	9	-38	6	63	9	Glyco_hydro_1	78	551	Glycoside hydrolase, family 1	1.90E-169
2	Locus21538v1rpkm9.68_12	5.7	15	-97	7	63	18	Glyco_hydro_1	81	551	Glycoside hydrolase, family 1	4.20E-176
2	Locus23527v1rpkm8.26_5	5.7	15	-83	5	20	15	Glyco_hydro_1	40	178	Glycoside hydrolase, family 1	2.30E-70
2	Locus29434v1rpkm4.97_5	4.9	3	-19	7	16	3	Glyco_hydro_1	15	136	Glycoside hydrolase, family 1	8.40E-44
3	Locus5251v1rpkm57.88_11	5.4	10	-63	6	73	13	Glyco_hydro_1	191	376	Glycoside hydrolase, family 1	7.40E-31
3	Locus5251v1rpkm57.88_11	5.4	10	-63	6	73	13	Glyco_hydro_1	397	543	Glycoside hydrolase, family 1	2.80E-25
3	Locus10440v1rpkm27.48_8	4.9	3	-12	5	47	3	Glyco_hydro_17	2	298	Glycoside hydrolase, family 17	6.90E-76
3	Locus12436v1rpkm22.01_10	4.9	4	-40	5	53	6	Glyco_hydro_17	35	352	Glycoside hydrolase, family 17	6.20E-80
2	Locus12970v1rpkm20.88_3	4.1	1	-11	10	42	2	Glyco_hydro_17	24	347	Glycoside hydrolase, family 17	2.00E-92
1	Locus46503v1rpkm1.19_9	3.1	1	-1.9	5	45	1	Glyco_hydro_17	1	272	Glycoside hydrolase, family 17	3.40E-72
3	Locus9481v1rpkm30.96_11	4.5	2	-12	7	49	2	Glyco_hydro_18	109	423	Glycoside hydrolase, family 18, catalytic domain	1.10E-16
3	Locus31971v1rpkm4.00_9	4.4	1	-2.3	8	36	1	Glyco_hydro_28	5	292	Glycoside hydrolase, family 28	8.30E-94
1	Locus56602v1rpkm0.71_12	2.9	1	-2.4	9	67	1	Glyco_hydro_3	107	338	Glycoside hydrolase, family 3, N-terminal	4.80E-71
1	Locus56602v1rpkm0.71_12	2.9	1	-2.4	9	67	1	Glyco_hydro_3_C	411	618	Glycoside hydrolase, family 3, C-terminal	1.80E-43
1	Locus7908v1rpkm37.99_15	5.4	20	-176	6	105	20	Glyco_hydro_31	346	790	Glycoside hydrolase, family 31	8.80E-163
1	Locus18087v1rpkm13.05_9	4.3	3	-21	6	59	3	Glyco_hydro_38	1	291	Glycoside hydrolase, family 38, core	3.10E-75
1	Locus28054v1rpkm5.61_4	3.6	1	-2.7	10	32	1	Glyco_hydro_38	160	280	Glycoside hydrolase, family 38, core	8.30E-46
2	Locus8039v1rpkm37.37_10	3.5	2	-20	6	65	2	Glyco_hydro_47	101	533	Glycoside hydrolase, family 47	7.00E-155
3	Locus13344v1rpkm20.12_13	4.7	2	-11	7	45	2	Glyco_hydro_63	1	387	Glycoside hydrolase, family 63	5.80E-186
3	Locus1069v1rpkm231.18_2	5.2	3	-20	6	15	3	Glycolytic	14	137	Fructose-bisphosphate aldolase, class-I	5.40E-35
3	Locus138v1rpkm827.02_3	5.9	16	-117	7	39	28	Glycolytic	11	358	Fructose-bisphosphate aldolase, class-I	4.10E-170
4	Locus449v1rpkm416.92_6	3.9	1	-10	6	30	1	Glycolytic	44	281	Fructose-bisphosphate aldolase, class-I	1.60E-118
4	Locus456v1rpkm412.62_4	5	4	-39	7	38	6	Glycolytic	11	357	Fructose-bisphosphate aldolase, class-I	7.30E-173
3	Locus618v1rpkm335.46_3	4.8	2	-15	6	10	4	Glycolytic	27	93	Fructose-bisphosphate aldolase, class-I	1.10E-33
3	Locus73v1rpkm1109.52_6	6.4	31	-201	7	34	66	Glycolytic	11	315	Fructose-bisphosphate aldolase, class-I	1.60E-158
4	Locus1139v1rpkm219.77_4	5.8	11	-60	5	16	31	Gp_dh_C	1	140	Glyceraldehyde 3-phosphate dehydrogenase, catalytic domain	6.60E-64
3	Locus121v1rpkm854.87_6	5.7	17	-119	7	37	55	Gp_dh_C	159	316	Glyceraldehyde 3-phosphate dehydrogenase, catalytic domain	7.50E-73
4	Locus15237v1rpkm16.86_3	4.6	1	-15	6	20	2	Gp_dh_C	34	178	Glyceraldehyde 3-phosphate dehydrogenase, catalytic domain	1.40E-54
4	Locus1526v1rpkm177.40_7	5	3	-27	8	37	6	Gp_dh_C	161	318	Glyceraldehyde 3-phosphate dehydrogenase, catalytic domain	1.80E-74
3	Locus245v1rpkm613.45_12	5.4	15	-91	9	48	17	Gp_dh_C	245	402	Glyceraldehyde 3-phosphate dehydrogenase, catalytic domain	3.30E-67
4	Locus3287v1rpkm91.18_5	5.4	6	-61	7	22	15	Gp_dh_C	28	185	Glyceraldehyde 3-phosphate dehydrogenase, catalytic domain	2.70E-75
3	Locus121v1rpkm854.87_6	5.7	17	-119	7	37	55	Gp_dh_N	4	154	Glyceraldehyde 3-phosphate dehydrogenase, NAD(P) binding domain	1.40E-55
4	Locus1526v1rpkm177.40_7	5	3	-27	8	37	6	Gp_dh_N	6	156	Glyceraldehyde 3-phosphate dehydrogenase, NAD(P) binding domain	1.50E-56
3	Locus245v1rpkm613.45_12	5.4	15	-91	9	48	17	Gp_dh_N	88	240	Glyceraldehyde 3-phosphate dehydrogenase, NAD(P) binding domain	1.20E-52
4	Locus7117v1rpkm42.35_10	3.8	1	-8.2	7	49	1	Granulin	367	415	Granulin	5.70E-10
3	Locus4728v1rpkm64.14_8	4.4	1	-5.7	6	48	1	GST_C	122	197	Glutathione S-transferase, C-terminal	1.90E-10

4	Locus5013v1rpkm60.63_4	3.1	1	-7.4	6	37	1	GST_C	35	100	Glutathione S-transferase, C-terminal	7.80E-10
3	Locus4728v1rpkm64.14_8	4.4	1	-5.7	6	48	1	GST_N	13	75	Glutathione S-transferase, N-terminal	1.30E-12
3	Locus193v1rpkm709.59_6	5.3	5	-30	8	33	5	GTP_EFTU	6	222	Protein synthesis factor, GTP-binding	2.40E-56
5	Locus37161v1rpkm2.51_3	5.1	3	-23	9	50	3	GTP_EFTU	7	233	Protein synthesis factor, GTP-binding	6.00E-57
3	Locus384v1rpkm465.72_11	3.9	1	-1.5	6	60	1	GTP_EFTU	17	341	Protein synthesis factor, GTP-binding	1.00E-59
3	Locus42159v1rpkm1.64_4	5.3	3	-18	9	51	3	GTP_EFTU	5	222	Protein synthesis factor, GTP-binding	1.10E-58
1	Locus43681v1rpkm1.46_4	3.8	2	-9.6	9	45	2	GTP_EFTU	9	174	Protein synthesis factor, GTP-binding	5.40E-42
5	Locus72v1rpkm1116.00_9	5.1	3	-22	9	44	3	GTP_EFTU	7	174	Protein synthesis factor, GTP-binding	1.60E-41
3	Locus193v1rpkm709.59_6	5.3	5	-30	8	33	5	GTP_EFTU_D2	248	297	Translation elongation factor EFTu/EF1A, domain 2	1.70E-12
5	Locus37161v1rpkm2.51_3	5.1	3	-23	9	50	3	GTP_EFTU_D2	259	325	Translation elongation factor EFTu/EF1A, domain 2	5.30E-15
3	Locus384v1rpkm465.72_11	3.9	1	-1.5	6	60	1	GTP_EFTU_D2	393	468	Translation elongation factor EFTu/EF1A, domain 2	4.00E-14
4	Locus41069v1rpkm1.79_1	3	1	-2.5	9	20	1	GTP_EFTU_D2	104	170	Translation elongation factor EFTu/EF1A, domain 2	2.60E-17
3	Locus42159v1rpkm1.64_4	5.3	3	-18	9	51	3	GTP_EFTU_D2	248	314	Translation elongation factor EFTu/EF1A, domain 2	5.50E-16
1	Locus43681v1rpkm1.46_4	3.8	2	-9.6	9	45	2	GTP_EFTU_D2	200	266	Translation elongation factor EFTu/EF1A, domain 2	1.80E-17
3	Locus497v1rpkm389.97_5	5.2	3	-18	9	27	3	GTP_EFTU_D2	48	114	Translation elongation factor EFTu/EF1A, domain 2	2.00E-17
5	Locus72v1rpkm1116.00_9	5.1	3	-22	9	44	3	GTP_EFTU_D2	200	266	Translation elongation factor EFTu/EF1A, domain 2	5.40E-17
5	Locus37161v1rpkm2.51_3	5.1	3	-23	9	50	3	GTP_EFTU_D3	335	440	Translation elongation factor EFTu/EF1A, C-terminal	6.40E-36
3	Locus42159v1rpkm1.64_4	5.3	3	-18	9	51	3	GTP_EFTU_D3	324	429	Translation elongation factor EFTu/EF1A, C-terminal	1.60E-35
1	Locus43681v1rpkm1.46_4	3.8	2	-9.6	9	45	2	GTP_EFTU_D3	276	381	Translation elongation factor EFTu/EF1A, C-terminal	2.50E-35
3	Locus497v1rpkm389.97_5	5.2	3	-18	9	27	3	GTP_EFTU_D3	122	229	Translation elongation factor EFTu/EF1A, C-terminal	1.30E-33
5	Locus72v1rpkm1116.00_9	5.1	3	-22	9	44	3	GTP_EFTU_D3	274	381	Translation elongation factor EFTu/EF1A, C-terminal	4.80E-33
1	Locus14772v1rpkm17.58_22	3.1	1	-1.5	6	205	1	GYF	617	666	GYF	7.80E-13
1	Locus18589v1rpkm12.49_1	4.4	4	-23	5	13	4	H_PPase	20	111	Pyrophosphate-energised proton pump	2.10E-08
1	Locus195v1rpkm706.72_5	6.3	27	-188	5	44	51	H_PPase	20	415	Pyrophosphate-energised proton pump	2.90E-107
6	Locus106v1rpkm921.90_2	6.2	30	-222	6	15	42	H_PPase	1	134	Pyrophosphate-energised proton pump	1.90E-59
6	Locus2238v1rpkm129.65_6	6.2	29	-227	6	41	56	H_PPase	20	381	Pyrophosphate-energised proton pump	1.80E-163
1	Locus2512v1rpkm117.37_7	6	20	-168	5	56	30	H_PPase	20	532	Pyrophosphate-energised proton pump	2.70E-160
1	Locus3621v1rpkm82.92_6	5.8	10	-80	5	43	10	H_PPase	1	412	Pyrophosphate-energised proton pump	2.70E-155
6	Locus847v1rpkm274.83_10	6.2	29	-248	5	80	88	H_PPase	20	751	Pyrophosphate-energised proton pump	8.00E-268
1	Locus18098v1rpkm13.03_21	4	1	-11	7	114	1	HAD	450	762	NULL	2.80E-16
3	Locus163v1rpkm761.84_8	4.9	3	-20	5	80	8	HATPase_c	30	179	ATPase-like, ATP-binding domain	3.70E-10
1	Locus2554v1rpkm115.94_14	5.6	28	-299	5	83	97	HATPase_c	16	165	ATPase-like, ATP-binding domain	4.90E-11
1	Locus3407v1rpkm87.91_4	5.2	15	-144	5	37	36	HATPase_c	106	255	ATPase-like, ATP-binding domain	1.00E-11
4	Locus32686v1rpkm3.75_27	5.9	2	-2.3	5	146	2	Helicase_C	931	1009	Helicase, C-terminal	8.00E-16
3	Locus704v1rpkm311.52_7	4.8	2	-13	6	43	2	Helicase_C	299	374	Helicase, C-terminal	2.30E-25
2	Locus306v1rpkm543.59_10	5.4	17	-157	5	56	22	Hemopexin	380	429	Hemopexin/matrixin, repeat	3.80E-07
5	Locus5690v1rpkm53.55_5	4.8	2	-10	6	30	2	Hexapep	120	151	Bacterial transferase hexapeptide repeat	7.20E-07
5	Locus5690v1rpkm53.55_5	4.8	2	-10	6	30	2	Hexapep	136	170	Bacterial transferase hexapeptide repeat	3.40E-04
5	Locus5690v1rpkm53.55_5	4.8	2	-10	6	30	2	Hexapep	54	87	Bacterial transferase hexapeptide repeat	1.70E-05
3	Locus5521v1rpkm55.11_8	5.1	7	-56	7	54	15	Hexokinase_1	44	240	Hexokinase, N-terminal	3.20E-57
5	Locus12135v1rpkm22.65_4	4.9	4	-32	6	31	4	Hexokinase_2	32	272	Hexokinase, C-terminal	1.10E-67
3	Locus5521v1rpkm55.11_8	5.1	7	-56	7	54	15	Hexokinase_2	246	486	Hexokinase, C-terminal	7.50E-68
2	Locus24382v1rpkm7.69_12	4.3	1	-1.6	5	64	1	Hist_deacetyl	6	209	Histone deacetylase domain	2.10E-70
6	Locus16424v1rpkm15.09_1	6.4	13	-100	12	11	13	Histone	28	94	Histone core	2.60E-14
6	Locus17585v1rpkm13.62_2	6.1	5	-31	11	16	8	Histone	25	98	Histone core	7.80E-25
6	Locus1941v1rpkm145.73_2	5.4	2	-9.1	11	18	3	Histone	76	150	Histone core	6.30E-33
6	Locus12752v1rpkm21.35_1	5.7	3	-17	10	14	3	Histone	29	103	Histone core	2.30E-23
6	Locus29883v1rpkm4.78_1	5.8	15	-39	10	17	19	Histone	69	134	Histone core	1.10E-22
3	Locus33727v1rpkm3.40_1	4.5	1	-4.6	11	15	2	Histone	27	99	Histone core	1.10E-26
6	Locus4589v1rpkm65.96_2	5.7	5	-11	10	14	5	Histone	19	92	Histone core	8.20E-27
6	Locus9358v1rpkm31.45_1	5.6	2	-8.6	11	15	2	Histone	58	132	Histone core	7.50E-33
6	Locus18499v1rpkm12.59_3	3.8	1	-9.5	8	19	1	HMA	102	168	Heavy metal-associated domain, HMA	1.00E-10
4	Locus29131v1rpkm5.10_5	4.5	1	-1.6	8	34	2	HSF_DNA-bind	26	117	Heat shock factor (HSF)-type, DNA-binding	3.20E-34
6	Locus1765v1rpkm156.49_2	3.4	1	-3.9	10	12	1	HSP20	1	53	Heat shock protein Hsp20	3.00E-17
3	Locus43822v1rpkm1.44_1	4.7	1	-1.7	6	21	1	HSP20	125	189	Heat shock protein Hsp20	6.60E-14

6	Locus691v1rpkm317.48_3	4.9	6	-52	6	19	6	HSP20	58	161	Heat shock protein Hsp20	2.10E-33
2	Locus16044v1rpkm15.62_12	5.9	21	-194	5	61	37	HSP70	1	527	Heat shock protein 70	1.60E-224
1	Locus17843v1rpkm13.32_7	5	7	-90	5	50	8	HSP70	145	278	Heat shock protein 70	2.20E-05
2	Locus1014v1rpkm239.05_9	3.8	2	-8.3	5	62	2	HSP70	1	573	Heat shock protein 70	6.10E-254
2	Locus1199v1rpkm210.54_5	5.3	9	-79	6	61	9	HSP70	9	557	Heat shock protein 70	8.50E-267
2	Locus2449v1rpkm120.26_9	6.2	31	-326	5	61	71	HSP70	1	527	Heat shock protein 70	1.30E-225
1	Locus30283v1rpkm4.61_13	5	7	-67	5	85	7	HSP70	1	593	Heat shock protein 70	3.40E-78
4	Locus31265v1rpkm4.25_3	4.7	5	-56	7	30	5	HSP70	6	268	Heat shock protein 70	3.30E-140
2	Locus35v1rpkm1508.57_6	5.5	18	-156	6	42	36	HSP70	1	385	Heat shock protein 70	1.10E-184
2	Locus3651v1rpkm81.96_6	6.2	30	-294	5	61	74	HSP70	1	527	Heat shock protein 70	2.20E-225
3	Locus379v1rpkm467.40_3	4.6	1	-1.9	5	15	1	HSP70	1	95	Heat shock protein 70	4.10E-13
3	Locus52408v1rpkm0.85_6	4.9	2	-7.1	6	74	2	HSP70	48	654	Heat shock protein 70	5.50E-266
2	Locus620v1rpkm335.15_8	5.4	14	-104	6	49	20	HSP70	84	447	Heat shock protein 70	1.70E-166
2	Locus620v1rpkm335.15_8	5.4	14	-104	6	49	20	HSP70	2	84	Heat shock protein 70	6.90E-42
2	Locus6330v1rpkm47.98_4	5.3	5	-36	9	17	12	HSP70	38	157	Heat shock protein 70	1.50E-49
2	Locus6755v1rpkm44.94_2	5	6	-62	5	14	10	HSP70	9	124	Heat shock protein 70	1.10E-49
2	Locus839v1rpkm275.85_11	5.7	21	-214	5	71	32	HSP70	10	618	Heat shock protein 70	9.00E-274
2	Locus9432v1rpkm31.14_6	5.6	8	-83	5	74	23	HSP70	41	647	Heat shock protein 70	1.70E-272
2	Locus985v1rpkm243.83_10	5.7	25	-251	5	71	45	HSP70	9	618	Heat shock protein 70	1.60E-277
3	Locus163v1rpkm761.84_8	4.9	3	-20	5	80	8	HSP90	184	701	Heat shock protein Hsp90	6.20E-265
1	Locus2554v1rpkm115.94_14	5.6	28	-299	5	83	97	HSP90	170	710	Heat shock protein Hsp90	3.90E-241
1	Locus3407v1rpkm87.91_4	5.2	15	-144	5	37	36	HSP90	260	328	Heat shock protein Hsp90	7.30E-18
1	Locus3118v1rpkm95.27_15	5.2	14	-125	7	90	17	Hydrolase	259	536	Haloacid dehalogenase-like hydrolase	5.50E-18
1	Locus3125v1rpkm95.04_10	4.9	6	-51	9	53	6	Hydrolase	5	118	Haloacid dehalogenase-like hydrolase	4.20E-14
1	Locus4144v1rpkm72.42_13	4.7	9	-62	6	59	9	Hydrolase	259	536	Haloacid dehalogenase-like hydrolase	2.10E-19
1	Locus9810v1rpkm29.54_11	4.7	7	-49	9	78	7	Hydrolase	69	346	Haloacid dehalogenase-like hydrolase	8.50E-20
3	Locus5917v1rpkm51.29_5	5.8	15	-78	5	37	24	Inhibitor_I29	26	82	Proteinase inhibitor I29, cathepsin propeptide	5.60E-17
4	Locus7117v1rpkm42.35_10	3.8	1	-8.2	7	49	1	Inhibitor_I29	36	93	Proteinase inhibitor I29, cathepsin propeptide	2.30E-18
3	Locus79396v1rpkm0.38_6	5.8	4	-11	7	34	6	Inhibitor_I29	1	54	Proteinase inhibitor I29, cathepsin propeptide	5.60E-13
4	Locus88119v1rpkm0.32_2	4.9	1	-1.7	6	11	2	Inhibitor_I9	11	92	Proteinase inhibitor I9, subtilisin propeptide	5.40E-06
2	Locus1426v1rpkm185.95_7	4.4	3	-16	5	29	3	iPGM_N	1	42	BPG-independent PGAM, N-terminal	3.80E-05
2	Locus48643v1rpkm1.04_13	4.3	3	-22	6	61	4	iPGM_N	102	332	BPG-independent PGAM, N-terminal	1.50E-67
6	Locus43762v1rpkm1.45_16	3.2	1	-4.2	9	88	1	K_trans	29	603	K+ potassium transporter	3.50E-177
5	Locus8400v1rpkm35.60_4	6	15	-123	10	26	15	KH_2	45	102	K Homology, type 2	2.00E-08
5	Locus99v1rpkm945.26_6	6	18	-151	10	30	25	KH_2	79	136	K Homology, type 2	2.60E-08
5	Locus2338v1rpkm125.15_4	6.3	27	-172	10	28	36	KOW	159	192	KOW	8.60E-07
5	Locus2707v1rpkm109.74_7	6.2	19	-121	10	23	19	KOW	112	145	KOW	6.20E-07
6	Locus271v1rpkm581.53_1	5.7	7	-55	11	14	9	KOW	51	82	KOW	1.00E-08
6	Locus8863v1rpkm33.43_2	3.8	1	-5.4	10	21	1	KOW	61	92	KOW	2.00E-09
3	Locus35126v1rpkm3.00_9	5.5	1	-1.3	9	54	1	Lactamase_B	89	249	Beta-lactamase-like	7.20E-21
4	Locus20361v1rpkm10.70_6	4.8	3	-31	7	36	4	Ldh_1_C	157	325	Lactate/malate dehydrogenase, C-terminal	1.90E-42
4	Locus2679v1rpkm110.76_7	5.7	14	-113	6	32	14	Ldh_1_C	117	285	Lactate/malate dehydrogenase, C-terminal	4.00E-42
4	Locus4501v1rpkm67.17_8	5.5	12	-111	9	43	12	Ldh_1_C	238	404	Lactate/malate dehydrogenase, C-terminal	4.90E-43
4	Locus783v1rpkm290.40_6	5.7	13	-115	5	32	21	Ldh_1_C	134	299	Lactate/malate dehydrogenase, C-terminal	3.20E-40
4	Locus9331v1rpkm31.53_6	5.2	4	-52	9	37	6	Ldh_1_C	179	343	Lactate/malate dehydrogenase, C-terminal	1.60E-45
4	Locus20361v1rpkm10.70_6	4.8	3	-31	7	36	4	Ldh_1_N	6	154	Lactate/malate dehydrogenase, N-terminal	3.10E-33
4	Locus2679v1rpkm110.76_7	5.7	14	-113	6	32	14	Ldh_1_N	2	114	Lactate/malate dehydrogenase, N-terminal	3.70E-26
4	Locus4501v1rpkm67.17_8	5.5	12	-111	9	43	12	Ldh_1_N	94	235	Lactate/malate dehydrogenase, N-terminal	1.70E-46
4	Locus783v1rpkm290.40_6	5.7	13	-115	5	32	21	Ldh_1_N	11	131	Lactate/malate dehydrogenase, N-terminal	1.80E-26
4	Locus9331v1rpkm31.53_6	5.2	4	-52	9	37	6	Ldh_1_N	35	177	Lactate/malate dehydrogenase, N-terminal	4.90E-50
4	Locus5306v1rpkm57.34_7	3.5	1	-8.9	5	27	2	LEA_2	8	103	Late embryogenesis abundant protein, LEA-14	4.40E-19
4	Locus5306v1rpkm57.34_7	3.5	1	-8.9	5	27	2	LEA_2	139	227	Late embryogenesis abundant protein, LEA-14	2.30E-10
4	Locus8036v1rpkm37.40_9	4	1	-5.6	5	35	2	LEA_2	203	297	Late embryogenesis abundant protein, LEA-14	5.10E-14
4	Locus8036v1rpkm37.40_9	4	1	-5.6	5	35	2	LEA_2	78	173	Late embryogenesis abundant protein, LEA-14	1.80E-18
6	Locus8940v1rpkm33.07_2	4.3	1	-3	11	27	1	Linker_histone	60	129	Histone H1/H5	1.00E-21

5	Locus5127v1rpkm59.36_2	4.2	1	-5.9	6	12	1	Lipase_3	35	80	Lipase, class 3	2.00E-06
4	Locus6265v1rpkm48.68_9	3.4	1	-1.5	9	54	2	Lipase_3	215	376	Lipase, class 3	4.00E-34
4	Locus640v1rpkm331.49_2	4.7	3	-27	8	17	7	Lipase_3	57	148	Lipase, class 3	5.40E-15
6	Locus423v1rpkm434.36_4	5.8	12	-92	6	22	12	Lipocalin_2	13	161	Lipocalin/cytosolic fatty-acid binding protein domain	8.30E-51
6	Locus4481v1rpkm67.50_2	5.5	6	-50	6	22	7	Lipocalin_2	13	161	Lipocalin/cytosolic fatty-acid binding protein domain	7.40E-51
3	Locus11164v1rpkm25.12_13	5.2	3	-19	6	99	3	Lipoxygenase	178	854	Lipoxygenase, C-terminal	0.00E+00
1	Locus14630v1rpkm17.84_14	3.2	1	-3.5	5	98	1	Lipoxygenase	171	848	Lipoxygenase, C-terminal	8.90E-299
1	Locus2057v1rpkm138.73_10	2.8	1	-1.6	6	97	1	Lipoxygenase	176	850	Lipoxygenase, C-terminal	3.50E-299
3	Locus2130v1rpkm134.82_6	5.1	4	-30	6	35	5	Lipoxygenase	1	286	Lipoxygenase, C-terminal	4.10E-134
3	Locus24327v1rpkm7.72_17	4.8	2	-12	5	97	4	Lipoxygenase	166	844	Lipoxygenase, C-terminal	7.70E-302
1	Locus3132v1rpkm94.95_5	4	2	-11	5	48	4	Lipoxygenase	174	431	Lipoxygenase, C-terminal	1.50E-86
4	Locus38409v1rpkm2.25_16	4.5	2	-14	6	98	2	Lipoxygenase	168	844	Lipoxygenase, C-terminal	0.00E+00
4	Locus507v1rpkm383.58_5	4.7	5	-36	9	20	6	Lipoxygenase	3	153	Lipoxygenase, C-terminal	4.00E-66
3	Locus557v1rpkm362.05_5	4.9	4	-22	6	29	7	Lipoxygenase	1	256	Lipoxygenase, C-terminal	1.20E-125
1	Locus40029v1rpkm1.95_7	4.2	2	-23	9	31	2	LrgB	119	281	LrgB-like protein	1.30E-34
1	Locus19862v1rpkm11.17_2	4.3	3	-16	10	22	5	LRR_1	163	180	Leucine-rich repeat	2.20E-01
6	Locus30264v1rpkm4.62_4	3.4	1	-3.2	6	67	1	LRR_1	506	524	Leucine-rich repeat	6.20E-01
6	Locus30264v1rpkm4.62_4	3.4	1	-3.2	6	67	1	LRR_1	263	285	Leucine-rich repeat	6.80E-03
6	Locus30264v1rpkm4.62_4	3.4	1	-3.2	6	67	1	LRR_1	166	188	Leucine-rich repeat	5.00E-01
6	Locus30264v1rpkm4.62_4	3.4	1	-3.2	6	67	1	LRR_1	335	356	Leucine-rich repeat	1.50E-01
6	Locus30264v1rpkm4.62_4	3.4	1	-3.2	6	67	1	LRR_1	530	550	Leucine-rich repeat	2.80E-01
1	Locus30301v1rpkm4.60_15	2.9	1	-1.5	6	104	1	LRR_1	319	341	Leucine-rich repeat	9.90E-02
1	Locus30301v1rpkm4.60_15	2.9	1	-1.5	6	104	1	LRR_1	273	294	Leucine-rich repeat	5.40E-01
4	Locus38580v1rpkm2.21_20	4.6	1	-1.1	5	113	1	LRR_1	217	237	Leucine-rich repeat	6.40E-01
4	Locus38580v1rpkm2.21_20	4.6	1	-1.1	5	113	1	LRR_1	591	613	Leucine-rich repeat	2.20E-02
4	Locus38580v1rpkm2.21_20	4.6	1	-1.1	5	113	1	LRR_1	568	587	Leucine-rich repeat	9.40E-01
4	Locus38580v1rpkm2.21_20	4.6	1	-1.1	5	113	1	LRR_1	483	504	Leucine-rich repeat	1.90E-02
4	Locus38580v1rpkm2.21_20	4.6	1	-1.1	5	113	1	LRR_1	459	479	Leucine-rich repeat	1.50E-02
5	Locus49472v1rpkm0.99_13	4.4	1	-1.4	6	84	1	LRR_1	533	555	Leucine-rich repeat	3.80E-01
1	Locus19862v1rpkm11.17_2	4.3	3	-16	10	22	5	LRR_4	90	127	NULL	7.20E-08
2	Locus11800v1rpkm23.51_14	4	2	-8.6	6	95	2	LRR_4	137	174	NULL	1.20E-06
6	Locus30264v1rpkm4.62_4	3.4	1	-3.2	6	67	1	LRR_4	119	154	NULL	5.70E-07
4	Locus38580v1rpkm2.21_20	4.6	1	-1.1	5	113	1	LRR_4	361	400	NULL	7.50E-07
1	Locus19862v1rpkm11.17_2	4.3	3	-16	10	22	5	LRRNT_2	26	61	Leucine-rich repeat-containing N-terminal, type 2	1.00E-07
6	Locus30264v1rpkm4.62_4	3.4	1	-3.2	6	67	1	LRRNT_2	30	66	Leucine-rich repeat-containing N-terminal, type 2	2.60E-07
1	Locus30301v1rpkm4.60_15	2.9	1	-1.5	6	104	1	LRRNT_2	31	69	Leucine-rich repeat-containing N-terminal, type 2	1.70E-08
4	Locus38580v1rpkm2.21_20	4.6	1	-1.1	5	113	1	LRRNT_2	29	66	Leucine-rich repeat-containing N-terminal, type 2	4.00E-08
3	Locus4348v1rpkm69.66_2	4.7	3	-23	7	29	5	LysM	174	217	Peptidoglycan-binding lysin domain	2.00E-08
3	Locus4348v1rpkm69.66_2	4.7	3	-23	7	29	5	LysM	109	156	Peptidoglycan-binding lysin domain	3.20E-03
3	Locus24189v1rpkm7.81_5	4.6	2	-13	6	46	2	M20_dimer	211	306	Peptidase M20, dimerisation	4.50E-10
3	Locus3471v1rpkm86.45_7	5.1	7	-58	6	47	15	M20_dimer	212	308	Peptidase M20, dimerisation	2.10E-11
3	Locus5055v1rpkm60.09_7	5.6	10	-87	6	50	21	M20_dimer	212	321	Peptidase M20, dimerisation	1.20E-05
3	Locus15310v1rpkm16.72_10	4.1	1	-1.5	6	32	1	Macro	98	210	Appr-1-p processing	4.70E-25
2	Locus11800v1rpkm23.51_14	4	2	-8.6	6	95	2	Malectin	267	450	Malectin	1.10E-52
3	Locus1989v1rpkm143.06_9	4.8	4	-27	8	71	6	malic	166	346	Malic enzyme, N-terminal	3.50E-78
5	Locus2108v1rpkm135.81_3	3.3	1	-5.5	6	33	2	malic	124	292	Malic enzyme, N-terminal	3.80E-70
3	Locus1989v1rpkm143.06_9	4.8	4	-27	8	71	6	Malic_M	357	609	Malic enzyme, NAD-binding	8.60E-92
6	Locus52803v1rpkm0.83_3	3.4	1	-1.8	9	17	1	MAPEG	19	138	Membrane-associated, eicosanoid/glutathione metabolism (MAPEG) protein	1.60E-15
4	Locus22464v1rpkm8.97_6	4.4	4	-34	9	42	5	MCE	128	203	Mammalian cell entry-related	1.80E-14
6	Locus29773v1rpkm4.83_14	4.4	1	-1.3	6	92	1	MCM	409	734	Mini-chromosome maintenance, DNA-dependent ATPase	1.60E-136
2	Locus1426v1rpkm185.95_7	4.4	3	-16	5	29	3	Metalloenzyme	44	262	Metalloenzyme	1.90E-72
2	Locus48643v1rpkm1.04_13	4.3	3	-22	6	61	4	Metalloenzyme	21	552	Metalloenzyme	2.20E-99
1	Locus16209v1rpkm15.39_13	3.3	1	-1.4	6	74	1	Metallophos	299	490	Metallophosphoesterase domain	4.50E-16
4	Locus14114v1rpkm18.73_7	5.3	6	-73	9	38	12	Methyltransf_11	116	210	Methyltransferase type 11	3.20E-20
6	Locus2042v1rpkm139.44_3	2.9	1	-5	6	26	1	Methyltransf_2	17	224	O-methyltransferase, family 2	4.60E-58
4	Locus2635v1rpkm112.64_8	3.9	2	-7.5	6	40	2	Methyltransf_2	91	331	O-methyltransferase, family 2	2.20E-69
1	Locus2220v1rpkm130.45_4	4.8	2	-9	10	21	2	MIP	1	176	Major intrinsic protein	3.10E-69
1	Locus2455v1rpkm120.01_2	4.9	2	-23	10	21	3	MIP	1	177	Major intrinsic protein	6.60E-70
1	Locus2478v1rpkm119.00_2	3.6	1	-10	7	24	1	MIP	14	232	Major intrinsic protein	2.30E-78
1	Locus2795v1rpkm106.32_1	4.9	3	-22	10	21	3	MIP	1	177	Major intrinsic protein	8.80E-69
1	Locus2948v1rpkm100.60_1	4.9	2	-18	7	18	2	MIP	26	172	Major intrinsic protein	1.20E-46
1	Locus3413v1rpkm87.75_2	5.2	3	-22	9	20	3	MIP	28	174	Major intrinsic protein	6.30E-54
1	Locus3611v1rpkm83.12_3	5.1	2	-8.8	6	25	2	MIP	14	231	Major intrinsic protein	9.00E-77
1	Locus575v1rpkm354.58_2	5.1	4	-33	9	30	4	MIP	31	267	Major intrinsic protein	1.00E-85
1	Locus617v1rpkm335.84_4	5	2	-13	9	31	4	MIP	44	273	Major intrinsic protein	6.30E-85
1	Locus6515v1rpkm46.60_3	3.5	1	-3.8	10	21	1	MIP	1	177	Major intrinsic protein	1.50E-70
5	Locus16872v1rpkm14.51_10	4.4	1	-1.6	9	47	1	Mito_carr	223	309	Mitochondrial substrate/solute carrier	1.00E-19
5	Locus16872v1rpkm14.51_10	4.4	1	-1.6	9	47	1	Mito_carr	129	215	Mitochondrial substrate/solute carrier	7.70E-19
5	Locus16872v1rpkm14.51_10	4.4	1	-1.6	9	47	1	Mito_carr	336	422	Mitochondrial substrate/solute carrier	7.50E-19
6	Locus1701v1rpkm162.26_8	4.6	5	-43	9	28	12	Mito_carr	103	190	Mitochondrial substrate/solute carrier	7.90E-18

6	Locus1701v1rpkm162.26_8	4.6	5	-43	9	28	12	Mito_carr	8	92	Mitochondrial substrate/solute carrier	5.30E-22
6	Locus11156v1rpkm25.15_8	4.7	3	-32	10	19	3	Mito_carr	1	66	Mitochondrial substrate/solute carrier	3.80E-12
5	Locus1202v1rpkm209.91_7	5.9	9	-68	10	26	16	Mito_carr	79	176	Mitochondrial substrate/solute carrier	3.30E-26
5	Locus1202v1rpkm209.91_7	5.9	9	-68	10	26	16	Mito_carr	184	221	Mitochondrial substrate/solute carrier	1.20E-04
4	Locus22990v1rpkm8.63_4	3.6	1	-4.1	10	16	1	Mito_carr	34	123	Mitochondrial substrate/solute carrier	1.30E-20
5	Locus240v1rpkm621.32_2	4.3	2	-10	11	28	2	Mito_carr	230	259	Mitochondrial substrate/solute carrier	1.50E-05
5	Locus240v1rpkm621.32_2	4.3	2	-10	11	28	2	Mito_carr	129	221	Mitochondrial substrate/solute carrier	2.30E-20
5	Locus240v1rpkm621.32_2	4.3	2	-10	11	28	2	Mito_carr	4	122	Mitochondrial substrate/solute carrier	1.20E-16
5	Locus29369v1rpkm5.00_4	4.5	3	-22	9	35	3	Mito_carr	57	127	Mitochondrial substrate/solute carrier	5.10E-12
5	Locus29369v1rpkm5.00_4	4.5	3	-22	9	35	3	Mito_carr	133	213	Mitochondrial substrate/solute carrier	4.70E-19
5	Locus29369v1rpkm5.00_4	4.5	3	-22	9	35	3	Mito_carr	228	314	Mitochondrial substrate/solute carrier	6.00E-24
5	Locus3539v1rpkm84.59_6	5.4	13	-135	10	32	18	Mito_carr	209	293	Mitochondrial substrate/solute carrier	1.30E-16
5	Locus3539v1rpkm84.59_6	5.4	13	-135	10	32	18	Mito_carr	14	91	Mitochondrial substrate/solute carrier	2.80E-15
5	Locus3539v1rpkm84.59_6	5.4	13	-135	10	32	18	Mito_carr	103	198	Mitochondrial substrate/solute carrier	3.90E-20
5	Locus5463v1rpkm55.77_7	5.4	13	-130	10	32	13	Mito_carr	103	198	Mitochondrial substrate/solute carrier	4.60E-20
5	Locus5463v1rpkm55.77_7	5.4	13	-130	10	32	13	Mito_carr	210	293	Mitochondrial substrate/solute carrier	1.70E-16
5	Locus5463v1rpkm55.77_7	5.4	13	-130	10	32	13	Mito_carr	15	91	Mitochondrial substrate/solute carrier	1.90E-15
5	Locus592v1rpkm346.40_9	6.2	22	-149	10	36	33	Mito_carr	230	318	Mitochondrial substrate/solute carrier	2.70E-16
5	Locus592v1rpkm346.40_9	6.2	22	-149	10	36	33	Mito_carr	25	122	Mitochondrial substrate/solute carrier	1.50E-25
5	Locus592v1rpkm346.40_9	6.2	22	-149	10	36	33	Mito_carr	130	223	Mitochondrial substrate/solute carrier	1.90E-21
1	Locus5359v1rpkm56.83_7	4.5	6	-59	6	58	9	Molybdopterin	129	448	Molybdopterin oxidoreductase	1.70E-68
3	Locus8138v1rpkm36.91_9	4.3	1	-1.1	6	78	1	Molybdopterin	340	659	Molybdopterin oxidoreductase	4.60E-67
5	Locus2760v1rpkm107.50_5	3.2	1	-1.7	10	31	1	Motile_Sperm	94	205	Major sperm protein	8.30E-28
5	Locus1320v1rpkm198.10_3	5	5	-43	8	23	7	MSP	91	217	Photosystem II PsbO, manganese-stabilising	4.60E-54
5	Locus309v1rpkm541.69_2	4.8	6	-50	5	25	9	MSP	2	230	Photosystem II PsbO, manganese-stabilising	2.20E-108
6	Locus3705v1rpkm80.66_2	4.1	2	-16	5	18	2	Mt_ATP-synt_D	12	154	ATPase, F0 complex, subunit D, mitochondrial	6.60E-13
1	Locus881v1rpkm266.16_10	3	1	-4.2	6	51	2	Na_Ca_ex	318	447	Sodium/calcium exchanger membrane region	6.00E-18
1	Locus881v1rpkm266.16_10	3	1	-4.2	6	51	2	Na_Ca_ex	126	275	Sodium/calcium exchanger membrane region	3.30E-23
1	Locus16823v1rpkm14.57_6	4.4	2	-8.6	9	44	2	Na_sulph_symp	5	413	Sodium/sulphate symporter	2.70E-150
1	Locus12222v1rpkm22.44_6	4.4	2	-16	10	60	2	Na_sulph_symp	94	558	Sodium/sulphate symporter	2.70E-133
5	Locus19057v1rpkm12.00_4	4.3	1	-6.2	4	23	1	NAC	70	127	Nascent polypeptide-associated complex NAC	3.50E-24
6	Locus10426v1rpkm27.53_1	5.1	7	-66	6	16	7	NAC	28	84	Nascent polypeptide-associated complex NAC	1.40E-19
6	Locus3834v1rpkm78.15_2	5	9	-79	5	19	9	NAC	36	91	Nascent polypeptide-associated complex NAC	9.00E-19
5	Locus16161v1rpkm15.46_3	3.3	1	-1.9	9	36	1	NAD_binding_1	187	293	Oxidoreductase FAD/NAD(P)-binding	1.10E-27
5	Locus12223v1rpkm22.44_5	4.1	1	-5.9	9	31	1	NAD_binding_1	156	262	Oxidoreductase FAD/NAD(P)-binding	8.10E-28
4	Locus4035v1rpkm74.07_7	4.3	3	-22	8	41	3	NAD_binding_1	218	333	Oxidoreductase FAD/NAD(P)-binding	6.30E-27
3	Locus8138v1rpkm36.91_9	4.3	1	-1.1	6	78	1	NADH-G_4Fe-4S_3	152	192	NADH:ubiquinone oxidoreductase, subunit G, iron-sulphur binding	4.80E-17
2	Locus12054v1rpkm22.87_2	3.4	1	-3.5	5	11	1	NAP	61	96	Nucleosome assembly protein (NAP)	4.00E-05
5	Locus49472v1rpkm0.99_13	4.4	1	-1.4	6	84	1	NB-ARC	1	44	NB-ARC	5.30E-09
6	Locus25518v1rpkm7.04_5	4.1	1	-6.1	9	26	1	NDK	86	219	Nucleoside diphosphate kinase	2.50E-53
1	Locus2367v1rpkm124.06_7	4	1	-4.9	9	59	1	Nodulin-like	14	263	Nodulin-like	3.10E-87
3	Locus9992v1rpkm28.97_9	4.4	1	-2.3	9	72	2	Nop	251	397	Pre-mRNA processing ribonucleoprotein, snoRNA-binding domain	5.30E-61
3	Locus9992v1rpkm28.97_9	4.4	1	-2.3	9	72	2	NOP5NT	1	65	NOP5, N-terminal	9.80E-23
3	Locus9992v1rpkm28.97_9	4.4	1	-2.3	9	72	2	NOSIC	159	211	NOSIC	3.20E-30
1	Locus8559v1rpkm34.70_9	3.1	1	-1.4	6	55	1	Nramp	77	437	Natural resistance-associated macrophage protein	1.10E-107
6	Locus2114v1rpkm135.60_1	4.9	3	-29	9	27	3	OSCP	65	237	ATPase, F1 complex, OSCP/delta subunit	4.90E-44

6	Locus2244v1rpkm129.45_1	4.2	3	-27	10	23	3	Oxidored_q6	84	193	NADH:ubiquinone oxidoreductase-like, 20kDa subunit	3.30E-22
3	Locus1651v1rpkm166.61_11	5.4	10	-69	8	65	21	p450	311	425	Cytochrome P450	2.20E-17
3	Locus17626v1rpkm13.57_8	4.3	1	-1.8	9	57	1	p450	29	496	Cytochrome P450	3.70E-84
5	Locus19733v1rpkm11.30_2	3.8	1	-4.2	9	18	1	p450	25	155	Cytochrome P450	1.90E-14
3	Locus11591v1rpkm24.00_6	5.2	6	-51	8	57	16	p450	39	479	Cytochrome P450	1.50E-55
5	Locus13195v1rpkm20.44_7	5.9	11	-87	7	51	28	p450	38	451	Cytochrome P450	5.00E-52
4	Locus1350v1rpkm194.03_1	4.7	4	-34	8	55	4	p450	323	431	Cytochrome P450	2.50E-14
5	Locus15856v1rpkm15.93_5	5.2	3	-16	8	56	3	p450	55	480	Cytochrome P450	7.70E-56
3	Locus22182v1rpkm9.17_4	4.8	4	-30	8	49	4	p450	5	413	Cytochrome P450	4.50E-59
4	Locus28020v1rpkm5.63_6	2.8	1	-1.8	6	22	1	p450	42	194	Cytochrome P450	4.40E-09
3	Locus288v1rpkm560.32_8	4.2	1	-7.3	9	56	2	p450	39	459	Cytochrome P450	5.90E-62
6	Locus3248v1rpkm91.95_9	4.8	2	-12	9	57	2	p450	84	491	Cytochrome P450	2.90E-81
4	Locus4259v1rpkm70.75_5	3.8	2	-12	6	40	2	p450	3	325	Cytochrome P450	4.60E-71
3	Locus4441v1rpkm68.22_7	4.5	3	-15	7	57	3	p450	36	485	Cytochrome P450	1.20E-105
6	Locus45206v1rpkm1.30_3	3.9	2	-16	9	56	2	p450	325	433	Cytochrome P450	6.20E-14
4	Locus4905v1rpkm61.86_3	3.3	1	-2.5	9	62	1	p450	60	525	Cytochrome P450	4.30E-95
1	Locus51763v1rpkm0.87_3	4.3	1	-1.4	7	34	3	p450	68	301	Cytochrome P450	2.20E-20
3	Locus5982v1rpkm50.70_6	4.4	2	-12	5	36	2	p450	8	300	Cytochrome P450	9.50E-64
3	Locus6771v1rpkm44.79_7	4.3	3	-22	7	40	10	p450	7	331	Cytochrome P450	1.20E-46
3	Locus850v1rpkm274.35_4	5.1	5	-44	8	24	8	p450	40	189	Cytochrome P450	1.20E-17
3	Locus8556v1rpkm34.73_5	3.8	1	-1.9	8	58	1	p450	30	501	Cytochrome P450	1.20E-96
3	Locus9167v1rpkm32.16_14	4.3	1	-3	9	58	2	p450	81	488	Cytochrome P450	1.60E-78
3	Locus3317v1rpkm90.27_8	4.4	3	-10	6	45	3	PA	91	164	Protease-associated domain, PA	6.20E-12
2	Locus7768v1rpkm38.68_9	3.2	1	-2.7	5	46	1	PA	87	163	Protease-associated domain, PA	4.30E-14
4	Locus8355v1rpkm35.87_6	3.7	2	-9.6	6	34	2	PALP	13	299	Pyridoxal phosphate-dependent enzyme, beta subunit	2.10E-58
2	Locus32140v1rpkm3.94_26	4.3	2	-7.5	6	89	2	PAN_2	345	407	PAN-2 domain	5.10E-19
4	Locus33011v1rpkm3.65_14	2.9	1	-1.2	6	90	1	PAN_2	348	406	PAN-2 domain	2.70E-17
3	Locus13894v1rpkm19.10_11	4.4	1	-6.3	7	63	1	PaO	319	413	Pheophorbide a oxygenase	1.90E-26
5	Locus22716v1rpkm8.80_5	4.6	2	-15	6	30	2	PAP_fibrillin	77	272	Plastid lipid-associated protein/fibrillin conserved domain	7.70E-49
1	Locus15627v1rpkm16.25_12	4.2	3	-18	7	72	3	PAS	179	268	PAS fold	6.50E-08
1	Locus3933v1rpkm76.24_14	3.6	1	-3.1	6	50	1	PD40	292	330	WD40-like Beta Propeller	1.50E-11
1	Locus3933v1rpkm76.24_14	3.6	1	-3.1	6	50	1	PD40	244	279	WD40-like Beta Propeller	1.50E-06
1	Locus3933v1rpkm76.24_14	3.6	1	-3.1	6	50	1	PD40	102	137	WD40-like Beta Propeller	1.10E-02
1	Locus3933v1rpkm76.24_14	3.6	1	-3.1	6	50	1	PD40	355	370	WD40-like Beta Propeller	1.30E-03
1	Locus40755v1rpkm1.84_25	3.3	1	-8	9	162	1	PDR_assoc	719	782	Plant PDR ABC transporter associated	3.90E-26
4	Locus2958v1rpkm100.32_9	4.1	2	-8.7	9	61	2	Pectinesterase	243	540	Pectinesterase, catalytic	2.80E-141
3	Locus12291v1rpkm22.33_7	5.5	6	-30	6	32	6	PEPcase	169	271	Phosphoenolpyruvate carboxylase	8.90E-09
1	Locus4288v1rpkm70.34_18	5.7	16	-126	7	77	16	PEPcase	26	660	Phosphoenolpyruvate carboxylase	1.40E-172
3	Locus5233v1rpkm58.01_11	4.8	3	-13	6	56	3	PEPcase	26	492	Phosphoenolpyruvate carboxylase	4.00E-114
1	Locus59v1rpkm1185.23_16	6.4	74	-576	6	105	342	PEPcase	169	893	Phosphoenolpyruvate carboxylase	2.00E-170
1	Locus7058v1rpkm42.73_6	4.7	2	-15	7	20	9	PEPcase	11	166	Phosphoenolpyruvate carboxylase	6.50E-25
4	Locus28v1rpkm1805.41_6	3.8	1	-1.1	5	20	1	Peptidase_C1	6	158	Peptidase C1A, papain C-terminal	1.70E-37
3	Locus2909v1rpkm102.00_2	5	3	-19	10	8.5	4	Peptidase_C1	3	76	Peptidase C1A, papain C-terminal	8.60E-26
3	Locus5917v1rpkm51.29_5	5.8	15	-78	5	37	24	Peptidase_C1	113	324	Peptidase C1A, papain C-terminal	7.30E-72
4	Locus7117v1rpkm42.35_10	3.8	1	-8.2	7	49	1	Peptidase_C1	124	339	Peptidase C1A, papain C-terminal	2.00E-82
3	Locus79396v1rpkm0.38_6	5.8	4	-11	7	34	6	Peptidase_C1	89	301	Peptidase C1A, papain C-terminal	3.40E-68
3	Locus36341v1rpkm2.70_9	4.2	2	-11	6	55	2	Peptidase_M16	91	236	Peptidase M16, N-terminal	2.00E-37
4	Locus4046v1rpkm73.87_9	5.2	10	-90	6	43	27	Peptidase_M16	1	109	Peptidase M16, N-terminal	7.10E-34
5	Locus4322v1rpkm69.96_9	4	1	-1.3	6	55	1	Peptidase_M16	91	235	Peptidase M16, N-terminal	8.60E-39
6	Locus5385v1rpkm56.60_2	3.4	1	-1.3	6	24	1	Peptidase_M16	106	221	Peptidase M16, N-terminal	1.70E-44
3	Locus36341v1rpkm2.70_9	4.2	2	-11	6	55	2	Peptidase_M16_C	246	425	Peptidase M16, C-terminal	1.30E-35
4	Locus4046v1rpkm73.87_9	5.2	10	-90	6	43	27	Peptidase_M16_C	116	301	Peptidase M16, C-terminal	8.10E-40
5	Locus4322v1rpkm69.96_9	4	1	-1.3	6	55	1	Peptidase_M16_C	243	426	Peptidase M16, C-terminal	2.70E-36
2	Locus4242v1rpkm71.02_2	3.6	2	-7.5	6	30	2	Peptidase_M17	6	282	Peptidase M17, leucyl aminopeptidase, C-terminal	2.00E-124
3	Locus14205v1rpkm18.56_7	4.5	3	-15	6	49	3	Peptidase_M20	97	431	Peptidase M20	1.10E-28
3	Locus14416v1rpkm18.18_7	5.3	8	-47	6	40	8	Peptidase_M20	32	365	Peptidase M20	3.20E-27
3	Locus24189v1rpkm7.81_5	4.6	2	-13	6	46	2	Peptidase_M20	100	412	Peptidase M20	4.00E-30
3	Locus3471v1rpkm86.45_7	5.1	7	-58	6	47	15	Peptidase_M20	103	419	Peptidase M20	3.80E-33
3	Locus40765v1rpkm1.84_8	5	6	-32	5	51	7	Peptidase_M20	133	467	Peptidase M20	3.30E-26
3	Locus5055v1rpkm60.09_7	5.6	10	-87	6	50	21	Peptidase_M20	98	435	Peptidase M20	9.20E-31
5	Locus1697v1rpkm162.63_6	3.5	1	-4.8	6	51	1	Peptidase_M41	251	464	Peptidase M41	5.70E-83
6	Locus11676v1rpkm23.80_4	4.6	4	-26	7	20	4	Peptidase_S24	53	104	Peptidase S24/S26A/S26B	1.70E-11
6	Locus23328v1rpkm8.39_4	4.7	4	-28	7	20	4	Peptidase_S24	53	105	Peptidase S24/S26A/S26B	1.50E-11
3	Locus14674v1rpkm17.75_17	5	6	-37	6	109	10	Peptidase_S8	3	181	Peptidase S8/S53, subtilisin/kexin/sedolisin	4.60E-31
1	Locus221v1rpkm662.77_3	4.5	2	-18	5	12	11	PEP-utilizers_C	4	113	PEP-utilising enzyme, C-terminal	1.00E-36

4	Locus13112v1rpkm20.61_4	5.3	10	-76	10	34	16	peroxidase	43	280	Haem peroxidase, plant/fungal/bacterial	3.00E-77
5	Locus22591v1rpkm8.89_3	5.7	10	-77	8	32	11	peroxidase	18	224	Haem peroxidase, plant/fungal/bacterial	3.80E-52
5	Locus2668v1rpkm111.34_3	5.9	14	-116	8	32	19	peroxidase	18	224	Haem peroxidase, plant/fungal/bacterial	6.00E-52
3	Locus5602v1rpkm54.46_3	5	6	-43	8	38	6	peroxidase	53	295	Haem peroxidase, plant/fungal/bacterial	2.00E-72
5	Locus63v1rpkm1179.83_5	4.3	3	-22	5	27	3	peroxidase	23	226	Haem peroxidase, plant/fungal/bacterial	1.20E-47
4	Locus9698v1rpkm29.98_4	3.7	1	-1.5	7	38	1	peroxidase	51	293	Haem peroxidase, plant/fungal/bacterial	1.10E-72
5	Locus13595v1rpkm19.64_3	4.3	3	-18	10	18	3	PEX11	12	158	NULL	1.10E-32
2	Locus1306v1rpkm200.01_14	5.1	7	-64	7	63	7	PGI	51	539	Phosphoglucose isomerase (PGI)	2.40E-204
3	Locus1355v1rpkm193.28_8	5	6	-51	9	50	11	PGK	85	464	Phosphoglycerate kinase	1.40E-163
3	Locus470v1rpkm403.18_5	5.6	14	-119	6	42	29	PGK	10	390	Phosphoglycerate kinase	5.00E-166
2	Locus10009v1rpkm28.92_3	4.7	6	-57	5	58	6	Phosphoesterase	7	371	Phosphoesterase	8.40E-102
2	Locus913v1rpkm259.97_9	4.4	4	-29	9	62	6	Phosphoesterase	52	415	Phosphoesterase	3.70E-99
6	Locus34356v1rpkm3.22_39	3.8	1	-1.1	6	200	1	PIP5K	1533	1758	Phosphatidylinositol-4-phosphate 5-kinase, core	1.50E-65
4	Locus10414v1rpkm27.54_3	5.3	8	-56	7	35	8	PI-PLC-X	68	166	Phospholipase C, phosphatidylinositol-specific , X domain	4.30E-09
2	Locus28456v1rpkm5.42_11	4.2	2	-9.1	6	68	2	PI-PLC-X	112	254	Phospholipase C, phosphatidylinositol-specific , X domain	1.10E-46
2	Locus30619v1rpkm4.49_8	3.8	2	-8.6	6	67	2	PI-PLC-X	108	250	Phospholipase C, phosphatidylinositol-specific , X domain	8.50E-47
2	Locus28456v1rpkm5.42_11	4.2	2	-9.1	6	68	2	PI-PLC-Y	353	441	Phospholipase C, phosphatidylinositol-specific, Y domain	6.00E-28
2	Locus30619v1rpkm4.49_8	3.8	2	-8.6	6	67	2	PI-PLC-Y	348	439	Phospholipase C, phosphatidylinositol-specific, Y domain	1.40E-28
1	Locus17788v1rpkm13.38_9	4.3	3	-20	6	49	4	Pkinase	106	390	Serine/threonine-protein kinase-like domain	1.10E-63
1	Locus15627v1rpkm16.25_12	4.2	3	-18	7	72	3	Pkinase	357	626	Serine/threonine-protein kinase-like domain	6.60E-61
6	Locus20689v1rpkm10.39_6	3.2	1	-2	7	43	1	Pkinase	56	328	Serine/threonine-protein kinase-like domain	1.50E-47
5	Locus27363v1rpkm5.96_7	4	1	-2.8	5	59	1	Pkinase	74	332	Serine/threonine-protein kinase-like domain	2.90E-69
4	Locus33011v1rpkm3.65_14	2.9	1	-1.2	6	90	1	Pkinase	496	762	Serine/threonine-protein kinase-like domain	5.30E-48
4	Locus38580v1rpkm2.21_20	4.6	1	-1.1	5	113	1	Pkinase	762	1027	Serine/threonine-protein kinase-like domain	3.80E-43
2	Locus8722v1rpkm33.98_10	4.2	5	-33	6	43	8	Pkinase	50	259	Serine/threonine-protein kinase-like domain	1.90E-48
6	Locus10842v1rpkm26.15_6	4.7	1	-7.8	6	31	1	Pkinase_Tyr	94	273	Serine-threonine/tyrosine-protein kinase	6.90E-23
2	Locus11800v1rpkm23.51_14	4	2	-8.6	6	95	2	Pkinase_Tyr	532	798	Serine-threonine/tyrosine-protein kinase	5.40E-49
4	Locus13588v1rpkm19.66_2	4.3	3	-28	5	36	4	Pkinase_Tyr	34	295	Serine-threonine/tyrosine-protein kinase	3.40E-32
5	Locus15211v1rpkm16.89_2	4.4	2	-17	9	22	2	Pkinase_Tyr	2	175	Serine-threonine/tyrosine-protein kinase	2.40E-28
1	Locus30301v1rpkm4.60_15	2.9	1	-1.5	6	104	1	Pkinase_Tyr	619	885	Serine-threonine/tyrosine-protein kinase	4.40E-46
2	Locus32140v1rpkm3.94_26	4.3	2	-7.5	6	89	2	Pkinase_Tyr	497	764	Serine-threonine/tyrosine-protein kinase	4.30E-47
1	Locus4200v1rpkm71.65_8	3.6	1	-6.1	7	74	1	Pkinase_Tyr	547	679	Serine-threonine/tyrosine-protein kinase	1.20E-28
3	Locus5028v1rpkm60.44_6	4.4	2	-6.8	7	41	2	Pkinase_Tyr	76	351	Serine-threonine/tyrosine-protein kinase	2.30E-47
4	Locus5126v1rpkm59.36_11	3	1	-2.4	7	39	1	Pkinase_Tyr	73	345	Serine-threonine/tyrosine-protein kinase	2.60E-48
6	Locus7071v1rpkm42.64_6	3.1	1	-5.8	5	29	1	Pkinase_Tyr	3	162	Serine-threonine/tyrosine-protein kinase	4.80E-23
3	Locus11164v1rpkm25.12_13	5.2	3	-19	6	99	3	PLAT	61	166	Lipoxygenase, LH2	1.20E-24
1	Locus14630v1rpkm17.84_14	3.2	1	-3.5	5	98	1	PLAT	54	159	Lipoxygenase, LH2	1.90E-17
1	Locus2057v1rpkm138.73_10	2.8	1	-1.6	6	97	1	PLAT	60	164	Lipoxygenase, LH2	1.30E-18
3	Locus24327v1rpkm7.72_17	4.8	2	-12	5	97	4	PLAT	49	154	Lipoxygenase, LH2	3.80E-17
1	Locus3132v1rpkm94.95_5	4	2	-11	5	48	4	PLAT	58	162	Lipoxygenase, LH2	1.50E-21
4	Locus38409v1rpkm2.25_16	4.5	2	-14	6	98	2	PLAT	52	156	Lipoxygenase, LH2	2.30E-19
5	Locus848v1rpkm274.78_3	4.6	2	-12	6	18	2	PLAT	58	162	Lipoxygenase, LH2	1.50E-20
1	Locus1103v1rpkm225.40_14	3.9	1	-4	6	68	1	PLD_C	551	593	Phospholipase D, C-terminal	7.30E-15
1	Locus12724v1rpkm21.40_12	4.4	2	-17	7	65	2	PLD_C	487	556	Phospholipase D, C-terminal	1.80E-29
1	Locus1103v1rpkm225.40_14	3.9	1	-4	6	68	1	PLDc	481	507	Phospholipase D/Transphosphatidylase	3.10E-08
1	Locus1103v1rpkm225.40_14	3.9	1	-4	6	68	1	PLDc	152	190	Phospholipase D/Transphosphatidylase	8.20E-13
1	Locus12724v1rpkm21.40_12	4.4	2	-17	7	65	2	PLDc	48	83	Phospholipase D/Transphosphatidylase	2.40E-05
1	Locus12724v1rpkm21.40_12	4.4	2	-17	7	65	2	PLDc	413	439	Phospholipase D/Transphosphatidylase	5.70E-08
5	Locus4554v1rpkm66.42_5	5.1	4	-32	9	38	4	Plug_translocon	42	76	Translocon Sec61/SecY, plug domain	1.30E-19

4	Locus2958v1rpkm100.32_9	4.1	2	-8.7	9	61	2	PMEI	78	195	Pectinesterase inhibitor	1.10E-17
4	Locus5361v1rpkm56.82_4	4	1	-5.6	5	28	1	PNP_UDP_1	27	255	Nucleoside phosphorylase domain	6.90E-26
4	Locus19121v1rpkm11.95_7	5.5	4	-19	9	22	4	Porin_3	3	202	Porin, eukaryotic type	4.00E-42
4	Locus21887v1rpkm9.39_5	3.9	1	-4.4	9	29	1	Porin_3	5	260	Porin, eukaryotic type	1.20E-58
4	Locus22868v1rpkm8.70_5	3.1	1	-2.2	7	36	1	Porin_3	48	320	Porin, eukaryotic type	6.00E-83
4	Locus5084v1rpkm59.81_5	5.4	10	-95	9	30	10	Porin_3	5	269	Porin, eukaryotic type	1.50E-58
4	Locus19v1rpkm2035.53_4	5.2	12	-104	6	39	42	PPDK_N	30	336	Pyruvate phosphate dikinase, PEP/pyruvate-binding	2.00E-49
4	Locus26811v1rpkm6.28_10	4.4	1	-1.2	8	61	1	PPR	208	232	Pentatricopeptide repeat	4.20E-01
4	Locus26811v1rpkm6.28_10	4.4	1	-1.2	8	61	1	PPR	381	411	Pentatricopeptide repeat	5.60E-06
4	Locus26811v1rpkm6.28_10	4.4	1	-1.2	8	61	1	PPR	347	367	Pentatricopeptide repeat	5.00E-02
4	Locus63029v1rpkm0.57_8	3.8	1	-1.4	9	66	1	PPR	421	446	Pentatricopeptide repeat	3.10E-01
4	Locus63029v1rpkm0.57_8	3.8	1	-1.4	9	66	1	PPR	49	73	Pentatricopeptide repeat	1.30E-02
4	Locus63029v1rpkm0.57_8	3.8	1	-1.4	9	66	1	PPR	79	109	Pentatricopeptide repeat	1.20E-05
4	Locus63029v1rpkm0.57_8	3.8	1	-1.4	9	66	1	PPR	149	175	Pentatricopeptide repeat	1.50E-03
4	Locus63029v1rpkm0.57_8	3.8	1	-1.4	9	66	1	PPR	315	341	Pentatricopeptide repeat	2.50E-02
4	Locus63029v1rpkm0.57_8	3.8	1	-1.4	9	66	1	PPR	280	307	Pentatricopeptide repeat	1.30E-07
4	Locus63029v1rpkm0.57_8	3.8	1	-1.4	9	66	1	PPR	177	207	Pentatricopeptide repeat	2.90E-08
4	Locus76842v1rpkm0.40_4	3.7	1	-1.1	9	24	1	PPR	73	88	Pentatricopeptide repeat	4.60E-01
4	Locus76842v1rpkm0.40_4	3.7	1	-1.1	9	24	1	PPR	197	217	Pentatricopeptide repeat	6.40E-01
4	Locus76842v1rpkm0.40_4	3.7	1	-1.1	9	24	1	PPR	167	193	Pentatricopeptide repeat	2.50E-02
4	Locus76842v1rpkm0.40_4	3.7	1	-1.1	9	24	1	PPR	93	119	Pentatricopeptide repeat	1.30E-05
4	Locus26811v1rpkm6.28_10	4.4	1	-1.2	8	61	1	PPR_1	451	478	NULL	3.70E-07
4	Locus26811v1rpkm6.28_10	4.4	1	-1.2	8	61	1	PPR_1	303	336	NULL	3.80E-08
4	Locus26811v1rpkm6.28_10	4.4	1	-1.2	8	61	1	PPR_1	269	301	NULL	1.10E-06
4	Locus26811v1rpkm6.28_10	4.4	1	-1.2	8	61	1	PPR_1	235	266	NULL	1.40E-07
4	Locus7712v1rpkm38.92_10	4.6	8	-73	6	56	9	Prenylcys_lyase	158	485	Prenylcysteine lyase	8.20E-105
4	Locus25306v1rpkm7.15_20	4.6	3	-25	6	87	3	Prenyltrans	589	626	Prenyltransferase/squalene oxidase	6.40E-10
4	Locus25306v1rpkm7.15_20	4.6	3	-25	6	87	3	Prenyltrans	638	688	Prenyltransferase/squalene oxidase	1.80E-10
4	Locus3586v1rpkm83.60_14	5	10	-110	6	71	13	Prenyltrans	146	188	Prenyltransferase/squalene oxidase	8.30E-09
4	Locus3586v1rpkm83.60_14	5	10	-110	6	71	13	Prenyltrans	589	620	Prenyltransferase/squalene oxidase	1.60E-07
4	Locus20124v1rpkm10.92_16	3.1	1	-5.1	5	74	1	PRKCSH	535	586	Glucosidase II beta subunit-like	1.80E-05
4	Locus25814v1rpkm6.85_16	3.1	1	-4.9	5	74	1	PRKCSH	531	583	Glucosidase II beta subunit-like	1.70E-05
4	Locus20124v1rpkm10.92_16	3.1	1	-5.1	5	74	1	PRKCSH-like	32	176	NULL	1.10E-35
4	Locus25814v1rpkm6.85_16	3.1	1	-4.9	5	74	1	PRKCSH-like	31	176	NULL	3.40E-36
4	Locus1240v1rpkm205.56_3	4.7	4	-31	5	18	7	Pro_CA	1	154	Carbonic anhydrase	8.10E-41
4	Locus1329v1rpkm197.00_5	5.3	3	-21	5	10	6	Pro_CA	1	92	Carbonic anhydrase	7.00E-15
4	Locus5202v1rpkm58.39_3	4.9	4	-28	6	22	4	Pro_CA	38	191	Carbonic anhydrase	1.50E-41
4	Locus25559v1rpkm7.01_1	4.2	1	-4.2	6	15	1	Pro_isomerase	9	138	Peptidyl-prolyl cis-trans isomerase, cyclophilin-type	6.10E-37
4	Locus4071v1rpkm73.51_4	5.3	8	-52	9	22	8	Pro_isomerase	42	204	Peptidyl-prolyl cis-trans isomerase, cyclophilin-type	7.30E-45
4	Locus33686v1rpkm3.42_3	3.6	1	-6.3	5	14	1	Profilin	2	126	Profilin/allergen	1.60E-47
4	Locus13940v1rpkm19.02_6	2.8	1	-1.9	5	19	1	Proteasome	1	150	Proteasome, subunit alpha/beta	2.90E-48
4	Locus2678v1rpkm110.77_5	4.4	3	-17	6	27	3	Proteasome	31	216	Proteasome, subunit alpha/beta	5.90E-51
4	Locus3035v1rpkm97.98_2	4.4	5	-37	7	27	5	Proteasome	27	210	Proteasome, subunit alpha/beta	1.90E-58
4	Locus4808v1rpkm63.09_3	4	2	-14	5	20	2	Proteasome	2	150	Proteasome, subunit alpha/beta	3.90E-50
4	Locus5371v1rpkm56.71_6	4.3	3	-19	6	27	3	Proteasome	35	220	Proteasome, subunit alpha/beta	1.10E-52
4	Locus6143v1rpkm49.41_2	4.1	1	-1.6	6	22	1	Proteasome	3	182	Proteasome, subunit alpha/beta	1.40E-39
4	Locus7067v1rpkm42.67_2	4.5	2	-16	5	23	2	Proteasome	7	186	Proteasome, subunit alpha/beta	2.50E-41
4	Locus8034v1rpkm37.40_5	4.4	2	-2.3	5	19	2	Proteasome	3	146	Proteasome, subunit alpha/beta	4.20E-48
4	Locus8797v1rpkm33.72_5	4.1	2	-8.5	7	28	2	Proteasome	30	213	Proteasome, subunit alpha/beta	3.30E-36
4	Locus2678v1rpkm110.77_5	4.4	3	-17	6	27	3	Proteasome_A_N	8	30	Proteasome, alpha-subunit, conserved site	1.00E-13
4	Locus3035v1rpkm97.98_2	4.4	5	-37	7	27	5	Proteasome_A_N	4	26	Proteasome, alpha-subunit, conserved site	7.20E-15
4	Locus5371v1rpkm56.71_6	4.3	3	-19	6	27	3	Proteasome_A_N	9	31	Proteasome, alpha-subunit, conserved site	1.40E-15
4	Locus8025v1rpkm37.44_6	3	1	-1.3	6	21	1	Prp19	65	130	Pre-mRNA-splicing factor 19	1.60E-32
4	Locus44623v1rpkm1.36_16	3.8	1	-8	9	90	1	PRT_C	639	794	Phosphoribosyltransferase C-terminal	1.20E-81
4	Locus7595v1rpkm39.60_16	4.7	2	-13	9	105	5	PRT_C	846	925	Phosphoribosyltransferase C-terminal	5.30E-38
4	Locus1290v1rpkm201.92_6	4.3	1	-2.2	9	28	1	PsbP	71	261	Photosystem II PsbP, oxygen evolving complex	1.70E-55
4	Locus459v1rpkm410.42_3	3.4	1	-2.2	9	26	1	PsbQ	45	242	Photosystem II PsbQ, oxygen evolving complex	5.20E-87
4	Locus4125v1rpkm72.67_2	5.3	3	-28	10	15	3	PSI_PsaE	81	142	Photosystem I PsaE, reaction centre subunit IV	2.40E-31
4	Locus5392v1rpkm56.56_1	5.6	6	-64	10	26	6	PSI_PsaF	55	231	Photosystem I PsaF, reaction centre subunit III	1.30E-84
4	Locus4621v1rpkm65.49_11	2.6	1	-2.5	6	65	1	PTR2	113	516	Oligopeptide transporter	1.50E-110
4	Locus18592v1rpkm12.49_4	4.5	2	-19	6	23	2	Pyrophosphatase	43	195	Inorganic pyrophosphatase	2.80E-55
4	Locus3473v1rpkm86.41_7	4.4	2	-17	5	23	2	Pyrophosphatase	43	195	Inorganic pyrophosphatase	2.90E-56
4	Locus9516v1rpkm30.80_6	4.5	1	-5.6	6	25	1	Pyrophosphatase	61	214	Inorganic pyrophosphatase	1.90E-58
4	Locus5488v1rpkm55.47_2	3.6	2	-6.9	9	15	4	Rad17	77	126	NULL	5.80E-06
4	Locus19521v1rpkm11.52_3	3.9	1	-2.8	6	24	1	Ras	16	177	Small GTPase superfamily	3.80E-56

4	Locus10332v1rpkm27.83_5	5	3	-22	7	22	3	Ras	35	195	Small GTPase superfamily	1.20E-58
4	Locus13503v1rpkm19.82_3	5.4	6	-64	5	23	6	Ras	10	176	Small GTPase superfamily	2.70E-55
4	Locus31894v1rpkm4.03_3	5.1	2	-9.8	7	23	2	Ras	11	169	Small GTPase superfamily	1.60E-56
4	Locus3645v1rpkm82.17_4	3.9	3	-30	9	21	3	Ras	17	178	Small GTPase superfamily	3.40E-66
4	Locus46679v1rpkm1.17_5	5.3	4	-42	7	23	4	Ras	10	170	Small GTPase superfamily	5.60E-67
4	Locus5659v1rpkm53.87_4	5.4	5	-50	5	23	5	Ras	10	177	Small GTPase superfamily	3.80E-55
4	Locus8267v1rpkm36.25_2	3.9	1	-7.2	8	24	1	Ras	13	173	Small GTPase superfamily	3.50E-61
4	Locus8534v1rpkm34.86_3	5.1	5	-38	6	24	5	Ras	15	175	Small GTPase superfamily	3.50E-60
4	Locus9845v1rpkm29.46_2	5.3	8	-82	7	23	8	Ras	8	168	Small GTPase superfamily	1.70E-62
4	Locus1612v1rpkm169.43_1	6.6	31	-135	9	20	31	RbcS	2	45	Ribulose-1,5-bisphosphate carboxylase small subunit, N-terminal	4.60E-19
4	Locus201v1rpkm695.83_1	6.6	38	-164	9	20	41	RbcS	2	45	Ribulose-1,5-bisphosphate carboxylase small subunit, N-terminal	7.90E-19
4	Locus5223v1rpkm58.19_4	3.5	1	-1.2	9	25	1	Rdx	71	210	Selenoprotein, Rdx type	1.20E-18
4	Locus31079v1rpkm4.32_8	3.6	1	-3.3	7	46	1	Redoxin	230	321	Redoxin	1.70E-05
4	Locus27608v1rpkm5.83_7	5.7	1	-1.5	8	54	1	Remorin_C	379	486	Remorin, C-terminal	1.90E-35
4	Locus7230v1rpkm41.70_2	5.1	3	-18	6	22	3	Remorin_C	87	196	Remorin, C-terminal	2.60E-35
4	Locus7230v1rpkm41.70_2	5.1	3	-18	6	22	3	Remorin_N	29	85	Remorin, N-terminal	1.80E-16
4	Locus11744v1rpkm23.64_3	3.9	1	-2	10	22	1	Rer1	17	182	Retrieval of early ER protein Rer1	3.30E-71
4	Locus409v1rpkm446.22_3	5.3	5	-32	9	29	8	Reticulon	71	232	Reticulon	2.70E-57
4	Locus4910v1rpkm61.79_1	3.6	1	-7.7	5	13	1	RGP	12	109	Alpha-1,4-glucan-protein synthase, UDP-forming	1.50E-42
4	Locus6751v1rpkm44.98_4	3.8	2	-18	6	15	2	RGP	1	132	Alpha-1,4-glucan-protein synthase, UDP-forming	1.40E-77
4	Locus5509v1rpkm55.21_14	4.9	3	-21	6	90	6	RHD3	44	766	Root hair defective 3 GTP-binding	2.60E-296
4	Locus10187v1rpkm28.29_3	4.1	3	-20	7	18	4	Ribophorin_I	52	161	Ribophorin I	2.90E-11
4	Locus12483v1rpkm21.94_8	5.1	4	-22	8	53	4	Ribophorin_I	37	464	Ribophorin I	1.10E-133
4	Locus48812v1rpkm1.03_3	4.3	2	-10	7	28	4	Ribophorin_I	2	135	Ribophorin I	2.60E-40
4	Locus5453v1rpkm55.85_9	5.4	8	-49	8	53	8	Ribophorin_I	37	464	Ribophorin I	8.60E-135
4	Locus14666v1rpkm17.77_9	3.4	1	-5.1	5	51	1	Ribophorin_II	10	476	Ribophorin II	2.30E-110
4	Locus21619v1rpkm9.61_11	4.3	2	-16	6	75	7	Ribophorin_II	9	685	Ribophorin II	2.20E-205
4	Locus1695v1rpkm162.69_4	5.7	11	-80	5	34	16	Ribosomal_60s	234	319	Ribosomal protein 60S	3.90E-23
4	Locus2517v1rpkm117.17_2	5.1	7	-64	10	37	7	Ribosomal_L1	128	329	Ribosomal protein L1	3.60E-45
4	Locus2990v1rpkm99.34_6	6.1	13	-88	10	25	18	Ribosomal_L1	14	211	Ribosomal protein L1	1.00E-48
4	Locus5216v1rpkm58.28_4	5.9	11	-71	10	25	11	Ribosomal_L1	16	211	Ribosomal protein L1	4.50E-48
4	Locus1695v1rpkm162.69_4	5.7	11	-80	5	34	16	Ribosomal_L10	7	108	Ribosomal protein L10/acidic P0	4.00E-23
4	Locus11162v1rpkm25.13_2	4.1	3	-35	10	24	3	Ribosomal_L10	38	134	Ribosomal protein L10/acidic P0	1.60E-25
4	Locus17035v1rpkm14.30_2	3.9	1	-7.1	10	21	1	Ribosomal_L11	142	195	Ribosomal protein L11, C-terminal	6.20E-19
4	Locus5895v1rpkm51.57_2	5.9	10	-95	9	18	10	Ribosomal_L11	75	144	Ribosomal protein L11, C-terminal	1.80E-13
4	Locus17035v1rpkm14.30_2	3.9	1	-7.1	10	21	1	Ribosomal_L11_N	80	137	Ribosomal protein L11, N-terminal	8.80E-29
4	Locus5895v1rpkm51.57_2	5.9	10	-95	9	18	10	Ribosomal_L11_N	13	70	Ribosomal protein L11, N-terminal	4.50E-17
4	Locus11773v1rpkm23.56_3	3.5	1	-5.1	6	20	1	Ribosomal_L12	118	185	Ribosomal protein L7/L12, C-terminal	3.00E-24
4	Locus1439v1rpkm184.62_3	5.5	10	-64	11	24	13	Ribosomal_L13	12	126	Ribosomal protein L13	9.60E-28
4	Locus1585v1rpkm172.19_2	4.6	4	-27	11	18	4	Ribosomal_L13	3	70	Ribosomal protein L13	1.20E-11
4	Locus2521v1rpkm117.04_3	5.6	10	-63	11	22	10	Ribosomal_L13	2	107	Ribosomal protein L13	1.10E-23
4	Locus6298v1rpkm48.35_3	4	1	-3.4	10	28	1	Ribosomal_L13	113	239	Ribosomal protein L13	2.60E-51
4	Locus9423v1rpkm31.19_5	4.8	5	-21	10	22	5	Ribosomal_L13	2	106	Ribosomal protein L13	1.30E-22
4	Locus1052v1rpkm233.27_2	5.9	13	-101	11	24	18	Ribosomal_L13e	6	184	Ribosomal protein L13e	2.20E-91
4	Locus1374v1rpkm190.93_1	4.3	1	-3.7	10	16	1	Ribosomal_L13e	1	118	Ribosomal protein L13e	1.60E-53
4	Locus22652v1rpkm8.84_2	5.6	6	-49	10	16	6	Ribosomal_L13e	1	118	Ribosomal protein L13e	1.10E-51
4	Locus4655v1rpkm65.08_3	5.8	9	-67	11	24	12	Ribosomal_L13e	6	184	Ribosomal protein L13e	6.50E-90
4	Locus1779v1rpkm155.89_3	5.1	4	-31	10	13	4	Ribosomal_L14	6	125	Ribosomal protein L14b/L23e	1.10E-34
4	Locus11889v1rpkm23.30_3	5.9	10	-77	10	15	10	Ribosomal_L14e	45	118	Ribosomal protein L14	1.40E-29
4	Locus6018v1rpkm50.46_3	5.5	4	-31	12	24	6	Ribosomal_L15e	2	193	Ribosomal protein L15e	2.10E-98
4	Locus6417v1rpkm47.33_3	3.2	1	-1.2	12	13	1	Ribosomal_L15e	2	106	Ribosomal protein L15e	1.50E-49
4	Locus1241v1rpkm205.49_7	5.5	7	-67	11	25	14	Ribosomal_L16	5	166	Ribosomal protein L10e/L16	4.40E-33
4	Locus1527v1rpkm177.32_5	5.2	5	-54	11	25	9	Ribosomal_L16	5	166	Ribosomal protein L10e/L16	2.60E-33
4	Locus19413v1rpkm11.64_5	4.8	1	-8.8	12	24	1	Ribosomal_L17	114	210	Ribosomal protein L17	3.80E-36
4	Locus3887v1rpkm77.32_5	3.8	1	-1.9	11	21	1	Ribosomal_L18ae	7	128	Ribosomal protein L18a/LX	1.80E-56
4	Locus6967v1rpkm43.33_4	5.9	19	-110	11	21	19	Ribosomal_L18ae	7	128	Ribosomal protein L18a/LX	6.50E-58
4	Locus16717v1rpkm14.69_3	6.2	16	-119	11	21	16	Ribosomal_L18e	2	123	Ribosomal protein L18e/L15P	1.10E-22
4	Locus2456v1rpkm119.98_2	5.7	7	-40	11	16	7	Ribosomal_L18e	20	144	Ribosomal protein L18e/L15P	1.60E-34
4	Locus3028v1rpkm98.17_3	6.2	20	-155	11	21	21	Ribosomal_L18e	2	122	Ribosomal protein L18e/L15P	1.90E-24
4	Locus3346v1rpkm89.49_1	5.4	6	-35	11	16	6	Ribosomal_L18e	20	137	Ribosomal protein L18e/L15P	5.80E-31
4	Locus4669v1rpkm64.86_3	6.1	15	-103	11	21	15	Ribosomal_L18e	2	122	Ribosomal protein L18e/L15P	1.20E-23
4	Locus2025v1rpkm140.35_6	5.9	19	-184	5	22	20	Ribosomal_L18p	32	61	Ribosomal protein L18/L5	8.40E-07
4	Locus2187v1rpkm132.12_5	6.1	18	-144	9	35	22	Ribosomal_L18p	26	172	Ribosomal protein L18/L5	9.20E-43
4	Locus6186v1rpkm49.15_6	6	17	-150	5	26	21	Ribosomal_L18p	1	100	Ribosomal protein L18/L5	1.00E-21
4	Locus1411v1rpkm187.21_6	4.2	1	-15	12	15	2	Ribosomal_L19e	3	128	Ribosomal protein L19/L19e	5.30E-56
4	Locus2141v1rpkm134.29_2	5.1	2	-8.2	12	21	3	Ribosomal_L19e	2	116	Ribosomal protein L19/L19e	1.00E-46
4	Locus4362v1rpkm69.39_3	4.6	1	-1.1	12	13	1	Ribosomal_L19e	17	69	Ribosomal protein L19/L19e	8.60E-19
4	Locus4362v1rpkm69.39_3	4.6	1	-1.1	12	13	1	Ribosomal_L19e	66	101	Ribosomal protein L19/L19e	3.10E-10
4	Locus667v1rpkm323.08_2	5.7	13	-85	11	28	27	Ribosomal_L2	13	90	Ribosomal Proteins L2, RNA binding domain	4.30E-17
4	Locus1297v1rpkm201.16_2	3.5	1	-3.6	11	19	1	Ribosomal_L2_C	17	149	Ribosomal protein L2, C-terminal	2.40E-41

4	Locus667v1rpkm323.08_2	5.7	13	-85	11	28	27	Ribosomal_L2_C	98	230	Ribosomal protein L2, C-terminal	7.00E-41
4	Locus2754v1rpkm107.80_2	5.9	8	-49	10	19	8	Ribosomal_L21e	4	99	Ribosomal protein L21e	1.80E-39
4	Locus3673v1rpkm81.39_2	4.2	1	-6.3	11	12	2	Ribosomal_L22	17	104	Ribosomal protein L22/L17	2.70E-12
4	Locus5046v1rpkm60.23_3	5	4	-33	11	21	5	Ribosomal_L22	17	153	Ribosomal protein L22/L17	4.40E-33
4	Locus5778v1rpkm52.70_3	4.8	2	-19	11	21	2	Ribosomal_L22	17	153	Ribosomal protein L22/L17	2.40E-33
4	Locus6890v1rpkm43.90_3	5.6	6	-40	10	20	6	Ribosomal_L22	17	153	Ribosomal protein L22/L17	2.40E-34
4	Locus5203v1rpkm58.39_1	5.7	9	-55	10	14	11	Ribosomal_L22e	12	125	Ribosomal protein L22e	2.70E-46
4	Locus4476v1rpkm67.59_1	5.8	11	-74	10	17	11	Ribosomal_L23	73	151	Ribosomal protein L25/L23	2.00E-18
4	Locus4476v1rpkm67.59_1	5.8	11	-74	10	17	11	Ribosomal_L23eN	13	67	Ribosomal protein L23/L25, N-terminal	2.00E-19
4	Locus3612v1rpkm83.12_2	5.2	4	-24	11	18	7	Ribosomal_L24e	3	73	Ribosomal protein L24e-related	1.60E-36
4	Locus8656v1rpkm34.22_3	3.8	1	-8.1	10	19	1	Ribosomal_L27	56	136	Ribosomal protein L27	2.10E-38
4	Locus4634v1rpkm65.37_1	5.6	5	-29	11	16	5	Ribosomal_L27e	52	135	Ribosomal protein L27e	2.10E-32
4	Locus1127v1rpkm220.82_2	6	14	-97	11	16	14	Ribosomal_L28e	6	135	Ribosomal protein L28e	2.60E-46
4	Locus5833v1rpkm52.11_2	5.6	12	-90	11	16	12	Ribosomal_L28e	6	135	Ribosomal protein L28e	1.40E-44
4	Locus87v1rpkm1018.40_1	5.4	6	-46	11	16	6	Ribosomal_L28e	5	135	Ribosomal protein L28e	7.80E-44
4	Locus16006v1rpkm15.69_1	5.6	9	-56	11	14	9	Ribosomal_L29	7	64	Ribosomal protein L29	3.40E-17
4	Locus14654v1rpkm17.79_6	4.2	2	-10	11	20	2	Ribosomal_L29	67	123	Ribosomal protein L29	7.30E-17
4	Locus3995v1rpkm74.99_1	5.7	10	-67	11	14	10	Ribosomal_L29	7	64	Ribosomal protein L29	1.10E-17
4	Locus19378v1rpkm11.67_5	4.5	7	-59	10	30	7	Ribosomal_L3	1	261	Ribosomal protein L3	3.60E-105
4	Locus11721v1rpkm23.70_3	4.8	3	-32	11	31	3	Ribosomal_L3	80	274	Ribosomal protein L3	4.00E-41
4	Locus12923v1rpkm20.99_6	5.8	15	-91	10	27	15	Ribosomal_L3	1	191	Ribosomal protein L3	1.30E-79
4	Locus5389v1rpkm56.58_6	6	21	-129	10	27	41	Ribosomal_L3	1	191	Ribosomal protein L3	1.30E-82
4	Locus671v1rpkm321.37_9	6.2	32	-200	10	45	58	Ribosomal_L3	50	343	Ribosomal protein L3	1.10E-124
4	Locus7780v1rpkm38.64_6	5.6	7	-51	10	28	9	Ribosomal_L3	50	240	Ribosomal protein L3	7.40E-78
4	Locus1259v1rpkm204.28_10	5.7	10	-76	10	29	11	Ribosomal_L30	85	136	Ribosomal protein L30, ferredoxin-like fold domain	1.90E-20
4	Locus2265v1rpkm128.70_6	5.7	10	-73	10	29	10	Ribosomal_L30	85	136	Ribosomal protein L30, ferredoxin-like fold domain	1.90E-20
4	Locus5562v1rpkm54.83_3	6.2	14	-83	10	17	20	Ribosomal_L30	1	40	Ribosomal protein L30, ferredoxin-like fold domain	1.50E-14
4	Locus1259v1rpkm204.28_10	5.7	10	-76	10	29	11	Ribosomal_L30_N	13	83	Ribosomal protein L30, N-terminal	1.20E-23
4	Locus2265v1rpkm128.70_6	5.7	10	-73	10	29	10	Ribosomal_L30_N	13	83	Ribosomal protein L30, N-terminal	1.10E-24
4	Locus558v1rpkm362.01_2	4.9	2	-18	11	13	3	Ribosomal_L34e	1	95	Ribosomal protein L34Ae	1.50E-34
4	Locus1118v1rpkm222.09_1	5.8	7	-55	12	12	7	Ribosomal_L36e	6	102	Ribosomal protein L36e	6.20E-43
4	Locus18052v1rpkm13.09_3	5.1	2	-21	10	10	2	Ribosomal_L37ae	2	91	Ribosomal protein L37ae	1.40E-38
4	Locus1634v1rpkm167.93_5	6.4	35	-218	11	34	71	Ribosomal_L4	1	169	Ribosomal protein L4/L1e	6.90E-32
4	Locus12742v1rpkm21.38_2	4.4	4	-33	6	31	4	Ribosomal_L4	76	257	Ribosomal protein L4/L1e	7.10E-56
4	Locus559v1rpkm361.33_2	6.6	39	-225	11	33	68	Ribosomal_L4	26	267	Ribosomal protein L4/L1e	1.30E-44
4	Locus1010v1rpkm239.39_3	5.8	7	-44	10	21	7	Ribosomal_L5	9	62	Ribosomal protein L5	3.60E-21
4	Locus1010v1rpkm239.39_3	5.8	7	-44	10	21	7	Ribosomal_L5_C	66	165	Ribosomal protein L5	4.30E-21
4	Locus1607v1rpkm170.04_3	5.8	12	-112	10	20	13	Ribosomal_L6	3	77	Ribosomal protein L6, alpha-beta domain	5.10E-15
4	Locus1607v1rpkm170.04_3	5.8	12	-112	10	20	13	Ribosomal_L6	89	168	Ribosomal protein L6, alpha-beta domain	5.20E-14
4	Locus2691v1rpkm110.38_3	5.9	17	-118	10	21	17	Ribosomal_L6	12	86	Ribosomal protein L6, alpha-beta domain	8.60E-15
4	Locus2691v1rpkm110.38_3	5.9	17	-118	10	21	17	Ribosomal_L6	98	177	Ribosomal protein L6, alpha-beta domain	5.00E-14
4	Locus4542v1rpkm66.53_3	5.8	14	-132	10	21	14	Ribosomal_L6	98	177	Ribosomal protein L6, alpha-beta domain	2.40E-14
4	Locus4542v1rpkm66.53_3	5.8	14	-132	10	21	14	Ribosomal_L6	12	86	Ribosomal protein L6, alpha-beta domain	1.00E-14
4	Locus2121v1rpkm135.23_3	5.5	11	-98	10	26	21	Ribosomal_L6e	128	235	Ribosomal protein L6E	3.50E-41
4	Locus432v1rpkm429.23_1	4.4	4	-30	7	12	6	Ribosomal_L6e	24	51	Ribosomal protein L6E	3.00E-10
4	Locus2121v1rpkm135.23_3	5.5	11	-98	10	26	21	Ribosomal_L6e_N	7	59	Ribosomal protein L6, N-terminal	1.50E-22
4	Locus14766v1rpkm17.58_1	5.5	11	-75	10	12	11	Ribosomal_L7Ae	13	105	Ribosomal protein L7Ae/L30e/S12e/Gadd45	4.40E-24
4	Locus2656v1rpkm111.68_3	5.6	9	-90	5	15	13	Ribosomal_L7Ae	23	116	Ribosomal protein L7Ae/L30e/S12e/Gadd45	3.10E-30
4	Locus2887v1rpkm102.82_4	5.4	5	-34	5	15	5	Ribosomal_L7Ae	23	116	Ribosomal protein L7Ae/L30e/S12e/Gadd45	7.90E-30
4	Locus5614v1rpkm54.34_3	6.1	19	-108	10	29	31	Ribosomal_L7Ae	114	203	Ribosomal protein L7Ae/L30e/S12e/Gadd45	5.80E-25
4	Locus8270v1rpkm36.24_4	6	13	-64	10	29	20	Ribosomal_L7Ae	115	204	Ribosomal protein L7Ae/L30e/S12e/Gadd45	2.60E-25
4	Locus16725v1rpkm14.68_1	5.3	8	-39	10	13	10	Ribosomal_S10	22	118	Ribosomal protein S10	1.20E-31
4	Locus2708v1rpkm109.72_4	5.4	4	-38	10	15	4	Ribosomal_S11	28	137	Ribosomal protein S11	3.30E-37
4	Locus3024v1rpkm98.47_7	5.6	7	-51	11	16	7	Ribosomal_S11	29	147	Ribosomal protein S11	4.70E-42
4	Locus7681v1rpkm39.05_4	5.6	7	-51	11	16	7	Ribosomal_S11	29	147	Ribosomal protein S11	4.70E-42
4	Locus11349v1rpkm24.59_2	5.6	5	-53	10	16	5	Ribosomal_S12	10	141	Ribosomal protein S12/S23	3.50E-41
4	Locus2734v1rpkm108.62_3	6	14	-89	11	18	16	Ribosomal_S13	14	142	Ribosomal protein S13	3.00E-42
4	Locus4778v1rpkm63.51_3	4.4	1	-1.4	10	19	1	Ribosomal_S13	52	156	Ribosomal protein S13	5.00E-37
4	Locus9692v1rpkm30.00_3	6.1	15	-88	11	18	15	Ribosomal_S13	14	142	Ribosomal protein S13	3.00E-42
4	Locus4648v1rpkm65.22_3	5.7	9	-79	11	17	9	Ribosomal_S13_N	1	60	Ribosomal protein S13/S15, N-terminal	3.10E-31
4	Locus4648v1rpkm65.22_3	5.7	9	-79	11	17	9	Ribosomal_S15	66	148	Ribosomal protein S15	3.40E-24
4	Locus1319v1rpkm198.16_3	5.4	6	-37	11	18	6	Ribosomal_S17	74	143	Ribosomal protein S17	4.10E-30
4	Locus13591v1rpkm19.65_3	5.5	6	-37	11	18	6	Ribosomal_S17	74	143	Ribosomal protein S17	4.10E-30
4	Locus1137v1rpkm219.91_1	5.6	13	-76	10	16	17	Ribosomal_S17e	1	120	Ribosomal protein S17e	3.10E-60
4	Locus5308v1rpkm57.33_4	5.2	12	-52	11	17	19	Ribosomal_S19	53	134	Ribosomal protein S19/S15	6.30E-35

4	Locus16253v1rpkm15.32_5	5.7	10	-69	10	16	10	Ribosomal_S19e	6	140	Ribosomal protein S19e	2.20E-58
4	Locus4019v1rpkm74.57_3	5.7	10	-65	10	16	10	Ribosomal_S19e	7	141	Ribosomal protein S19e	6.30E-58
4	Locus1999v1rpkm142.35_4	6.1	12	-128	5	33	22	Ribosomal_S2	16	183	Ribosomal protein S2	9.10E-39
4	Locus3204v1rpkm92.98_6	6.1	11	-114	5	33	11	Ribosomal_S2	16	183	Ribosomal protein S2	5.50E-39
4	Locus34864v1rpkm3.08_1	5.7	6	-45	10	10	6	Ribosomal_S24e	25	91	Ribosomal protein S24e	5.60E-28
4	Locus3504v1rpkm85.43_2	5.7	6	-47	11	16	6	Ribosomal_S24e	25	107	Ribosomal protein S24e	1.90E-35
4	Locus33668v1rpkm3.42_2	5.6	7	-26	11	12	7	Ribosomal_S25	1	106	Ribosomal protein S25	7.10E-46
4	Locus23393v1rpkm8.35_3	4.4	1	-2.6	11	14	1	Ribosomal_S26e	1	110	Ribosomal protein S26e	2.70E-53
4	Locus3530v1rpkm84.83_2	5.1	4	-33	11	15	4	Ribosomal_S26e	1	110	Ribosomal protein S26e	7.70E-54
4	Locus3031v1rpkm98.14_2	5.6	4	-31	10	18	4	Ribosomal_S27	101	147	Ribosomal protein S27a	2.70E-26
4	Locus1939v1rpkm145.89_1	3.1	1	-2.2	9	9.6	2	Ribosomal_S27e	30	84	Ribosomal protein S27e	1.10E-26
4	Locus8400v1rpkm35.60_4	6	15	-123	10	26	15	Ribosomal_S3_C	105	188	Ribosomal protein S3, C-terminal	1.60E-18
4	Locus99v1rpkm945.26_6	6	18	-151	10	30	25	Ribosomal_S3_C	139	222	Ribosomal protein S3, C-terminal	2.20E-18
4	Locus1331v1rpkm196.96_3	5.9	13	-138	10	30	20	Ribosomal_S3Ae	11	221	Ribosomal protein S3Ae	1.30E-83
4	Locus3009v1rpkm98.84_4	5.8	13	-134	10	30	18	Ribosomal_S3Ae	12	222	Ribosomal protein S3Ae	1.80E-83
4	Locus656v1rpkm328.29_5	6.1	12	-88	10	23	17	Ribosomal_S4	7	108	Ribosomal protein S4/S9, N-terminal	7.40E-26
4	Locus2338v1rpkm125.15_4	6.3	27	-172	10	28	36	Ribosomal_S4e	76	152	Ribosomal protein S4e, central	6.10E-31
4	Locus2707v1rpkm109.74_7	6.2	19	-121	10	23	19	Ribosomal_S4e	29	105	Ribosomal protein S4e, central	3.90E-32
4	Locus1224v1rpkm207.31_2	5.8	12	-80	10	20	15	Ribosomal_S5	1	61	Ribosomal protein S5, N-terminal	1.90E-26
4	Locus600v1rpkm341.65_3	6	14	-124	10	30	31	Ribosomal_S5	89	154	Ribosomal protein S5, N-terminal	5.30E-29
4	Locus1224v1rpkm207.31_2	5.8	12	-80	10	20	15	Ribosomal_S5_C	78	144	Ribosomal protein S5, C-terminal	1.40E-23
4	Locus600v1rpkm341.65_3	6	14	-124	10	30	31	Ribosomal_S5_C	171	237	Ribosomal protein S5, C-terminal	3.10E-23
4	Locus30623v1rpkm4.49_4	5.8	6	-65	10	13	6	Ribosomal_S6e	1	118	Ribosomal protein S6e	7.00E-52
4	Locus952v1rpkm251.00_4	6.1	15	-120	11	28	20	Ribosomal_S6e	1	128	Ribosomal protein S6e	1.40E-55
4	Locus3517v1rpkm85.05_4	5.6	9	-63	10	22	9	Ribosomal_S7	46	200	Ribosomal protein S7 domain	4.80E-40
4	Locus3701v1rpkm80.78_4	5.6	11	-98	10	22	11	Ribosomal_S7	47	201	Ribosomal protein S7 domain	2.30E-40
4	Locus6221v1rpkm48.96_4	5.5	11	-98	10	23	11	Ribosomal_S7	50	204	Ribosomal protein S7 domain	2.50E-40
4	Locus1262v1rpkm204.06_4	5.4	7	-61	10	22	7	Ribosomal_S7e	6	190	Ribosomal protein S7e	5.60E-83
4	Locus5393v1rpkm56.55_3	5.4	6	-46	10	22	7	Ribosomal_S7e	6	190	Ribosomal protein S7e	4.10E-83
4	Locus6158v1rpkm49.34_3	5.6	8	-62	10	22	8	Ribosomal_S7e	6	190	Ribosomal protein S7e	1.30E-79
4	Locus6586v1rpkm46.10_3	5.8	9	-55	10	15	9	Ribosomal_S8	6	129	Ribosomal protein S8	1.70E-25
4	Locus15905v1rpkm15.85_4	4.5	1	-7.1	11	23	1	Ribosomal_S9	95	216	Ribosomal protein S9	7.90E-44
4	Locus5196v1rpkm58.47_1	6	12	-69	10	17	12	Ribosomal_S9	15	147	Ribosomal protein S9	2.10E-37
4	Locus13894v1rpkm19.10_11	4.4	1	-6.3	7	63	1	Rieske	110	193	Rieske [2Fe-2S] iron-sulphur domain	8.10E-19
4	Locus14320v1rpkm18.34_7	4.4	3	-22	9	30	3	Rieske	191	261	Rieske [2Fe-2S] iron-sulphur domain	1.70E-16
4	Locus4450v1rpkm68.08_5	4.9	2	-7.9	9	24	2	Rieske	137	203	Rieske [2Fe-2S] iron-sulphur domain	3.60E-17
4	Locus5815v1rpkm52.29_8	4.5	3	-26	9	30	3	Rieske	195	262	Rieske [2Fe-2S] iron-sulphur domain	8.30E-16
4	Locus16355v1rpkm15.20_3	4.2	1	-7	7	35	1	RIP	42	236	Ribosome-inactivating protein	2.00E-38
4	Locus380v1rpkm467.37_6	5.4	9	-78	8	34	10	RIP	42	238	Ribosome-inactivating protein	2.80E-36
4	Locus44195v1rpkm1.41_5	4.5	2	-14	8	31	2	RIP	16	210	Ribosome-inactivating protein	7.40E-29
4	Locus698v1rpkm314.64_3	5.9	12	-109	6	34	15	RIP	40	234	Ribosome-inactivating protein	6.70E-38
4	Locus7256v1rpkm41.54_3	6.9	58	-366	6	32	80	RIP	37	232	Ribosome-inactivating protein	1.40E-31
4	Locus8331v1rpkm35.99_3	4.4	3	-25	9	33	3	RIP	21	226	Ribosome-inactivating protein	4.80E-21
4	Locus35126v1rpkm3.00_9	5.5	1	-1.3	9	54	1	RMMBL	435	466	RNA-metabolising metallo-beta-lactamase	1.10E-08
4	Locus3450v1rpkm86.80_14	5.9	29	-244	6	51	46	RPE65	4	443	Carotenoid oxygenase	4.30E-131
4	Locus38497v1rpkm2.23_4	3.4	1	-3.5	9	31	1	RRM_1	184	253	RNA recognition motif domain	8.10E-16
4	Locus38497v1rpkm2.23_4	3.4	1	-3.5	9	31	1	RRM_1	88	157	RNA recognition motif domain	2.40E-16
4	Locus4001v1rpkm74.88_6	3.5	1	-4.6	4	35	2	RRM_1	140	209	RNA recognition motif domain	7.90E-19
4	Locus4001v1rpkm74.88_6	3.5	1	-4.6	4	35	2	RRM_1	234	304	RNA recognition motif domain	4.10E-21
4	Locus2338v1rpkm125.15_4	6.3	27	-172	10	28	36	RS4NT	1	22	Ribosomal protein S4e, N-terminal	3.00E-07
4	Locus1612v1rpkm169.43_1	6.6	31	-135	9	20	31	RuBisCO_small	69	177	Ribulose bisphosphate carboxylase small chain, domain	5.10E-38
4	Locus201v1rpkm695.83_1	6.6	38	-164	9	20	41	RuBisCO_small	69	177	Ribulose bisphosphate carboxylase small chain, domain	1.00E-37
4	Locus32140v1rpkm3.94_26	4.3	2	-7.5	6	89	2	S_locus_glycop	216	323	S-locus glycoprotein	1.10E-20
4	Locus33011v1rpkm3.65_14	2.9	1	-1.2	6	90	1	S_locus_glycop	207	316	S-locus glycoprotein	9.00E-32
4	Locus7955v1rpkm37.72_5	3.2	1	-3.2	5	39	1	S1	20	93	Ribosomal protein S1, RNA-binding domain	4.60E-13
4	Locus1797v1rpkm154.57_2	5.6	8	-53	10	13	11	S10_plectin	1	51	Plectin/S10, N-terminal	1.30E-23
4	Locus4760v1rpkm63.67_2	5.7	9	-60	10	20	15	S10_plectin	3	96	Plectin/S10, N-terminal	1.30E-43
4	Locus6069v1rpkm50.05_4	5.3	4	-40	10	20	4	S10_plectin	3	95	Plectin/S10, N-terminal	2.70E-42
4	Locus2338v1rpkm125.15_4	6.3	27	-172	10	28	36	S4	29	72	RNA-binding S4	6.20E-06
4	Locus656v1rpkm328.29_5	6.1	12	-88	10	23	17	S4	109	152	RNA-binding S4	3.00E-12
4	Locus3247v1rpkm91.99_8	4.4	1	-3.7	6	28	2	S6PP	6	199	Sucrose-phosphate synthase	1.10E-17
4	Locus8288v1rpkm36.18_7	5.2	4	-30	8	51	4	Sad1_UNC	311	441	Sad1/UNC-like, C-terminal	3.60E-39
4	Locus22526v1rpkm8.92_2	4.7	2	-11	10	27	4	SAM_1	203	252	Sterile alpha motif, type 1	4.50E-07
4	Locus2922v1rpkm101.38_11	5.6	7	-76	6	44	7	SapB_1	379	399	Saposin-like type B, 1	1.80E-04
4	Locus5869v1rpkm51.81_10	4.5	1	-4.3	5	33	1	SapB_1	170	207	Saposin-like type B, 1	5.10E-13
4	Locus711v1rpkm308.68_7	4.8	3	-16	5	36	3	SapB_1	300	337	Saposin-like type B, 1	1.40E-14
4	Locus2922v1rpkm101.38_11	5.6	7	-76	6	44	7	SapB_2	317	349	Saposin-like type B, 2	1.20E-11
4	Locus5869v1rpkm51.81_10	4.5	1	-4.3	5	33	1	SapB_2	106	140	Saposin-like type B, 2	2.50E-12
4	Locus711v1rpkm308.68_7	4.8	3	-16	5	36	3	SapB_2	237	271	Saposin-like type B, 2	1.20E-13
4	Locus3233v1rpkm92.34_5	3.6	1	-6.1	9	37	2	SecY	9	323	SecY protein	2.20E-74
4	Locus4554v1rpkm66.42_5	5.1	4	-32	9	38	4	SecY	77	321	SecY protein	3.20E-61
4	Locus8522v1rpkm34.93_4	4.1	1	-2.6	9	26	1	SNARE	141	202	Target SNARE coiled-coil domain	1.50E-13
4	Locus32686v1rpkm3.75_27	5.9	2	-2.3	5	146	2	SNF2_N	356	640	SNF2-related	3.80E-92

4	Locus45226v1rpkm1.30_4	2.9	1	-1.3	5	33	1	SOR_SNZ	18	229	Vitamin B6 biosynthesis protein	1.80E-107
4	Locus1166v1rpkm214.72_6	3.8	2	-19	5	28	3	SOUL	62	243	SOUL haem-binding protein	1.70E-43
4	Locus19366v1rpkm11.69_5	3.5	1	-5.1	9	21	1	SPC25	19	179	Signal peptidase complex subunit 2	3.10E-43
4	Locus15900v1rpkm15.86_4	4.6	2	-14	9	29	2	SRPRB	51	230	Signal recognition particle receptor, beta subunit	1.40E-37
4	Locus17842v1rpkm13.33_3	3.9	1	-8.3	8	19	2	SRPRB	18	147	Signal recognition particle receptor, beta subunit	2.30E-28
4	Locus23946v1rpkm7.97_3	4.6	2	-17	6	11	2	SRPRB	51	89	Signal recognition particle receptor, beta subunit	1.10E-07
4	Locus10072v1rpkm28.69_4	5.2	7	-53	6	41	9	Str_synth	153	240	Strictosidine synthase, conserved region	1.30E-31
4	Locus22164v1rpkm9.19_3	4.5	4	-30	8	42	8	Str_synth	160	246	Strictosidine synthase, conserved region	1.40E-29
4	Locus2501v1rpkm117.62_4	4.7	3	-27	6	35	6	Str_synth	150	233	Strictosidine synthase, conserved region	3.30E-27
4	Locus36546v1rpkm2.65_7	4.9	7	-55	6	41	7	Str_synth	154	240	Strictosidine synthase, conserved region	1.10E-33
4	Locus3301v1rpkm90.74_11	4.1	2	-13	6	69	2	Succ_DH_flav_C	495	627	Fumarate reductase/succinate dehydrogenase flavoprotein, C-terminal	1.60E-46
4	Locus36448v1rpkm2.67_10	3.9	2	-13	7	71	3	Succ_DH_flav_C	515	650	Fumarate reductase/succinate dehydrogenase flavoprotein, C-terminal	1.30E-46
4	Locus7932v1rpkm37.88_8	4.6	5	-42	6	41	8	Surf_Ag_VNR	188	253	Surface antigen variable number	1.20E-04
4	Locus15609v1rpkm16.29_10	6.7	42	-243	6	59	143	Synthase_beta	1	43	ATP synthase, F1 beta subunit	1.10E-08
4	Locus2786v1rpkm106.74_2	4.3	2	-14	10	14	4	Synthase_beta	1	46	ATP synthase, F1 beta subunit	9.50E-12
4	Locus18083v1rpkm13.06_4	4.2	2	-11	9	33	3	Tetraspannin	7	267	Tetraspanin	1.70E-28
4	Locus1484v1rpkm181.14_2	4.3	1	-4.5	9	31	2	Tetraspannin	8	260	Tetraspanin	2.20E-39
4	Locus13524v1rpkm19.77_7	5.2	5	-43	5	47	5	Thioredoxin	30	130	Thioredoxin domain	5.60E-29
4	Locus13524v1rpkm19.77_7	5.2	5	-43	5	47	5	Thioredoxin	160	258	Thioredoxin domain	3.90E-31
4	Locus1581v1rpkm172.47_13	6.4	53	-435	5	56	91	Thioredoxin	384	486	Thioredoxin domain	7.40E-30
4	Locus1581v1rpkm172.47_13	6.4	53	-435	5	56	91	Thioredoxin	40	147	Thioredoxin domain	1.00E-32
4	Locus21099v1rpkm10.00_9	3.8	2	-16	5	60	2	Thioredoxin	78	178	Thioredoxin domain	3.90E-15
4	Locus21099v1rpkm10.00_9	3.8	2	-16	5	60	2	Thioredoxin	416	519	Thioredoxin domain	2.10E-16
4	Locus31897v1rpkm4.03_8	3.7	2	-14	5	60	2	Thioredoxin	418	520	Thioredoxin domain	8.90E-16
4	Locus31897v1rpkm4.03_8	3.7	2	-14	5	60	2	Thioredoxin	79	180	Thioredoxin domain	5.90E-15
4	Locus3231v1rpkm92.36_6	4.9	5	-58	6	26	14	Thioredoxin	148	233	Thioredoxin domain	2.20E-31
4	Locus3231v1rpkm92.36_6	4.9	5	-58	6	26	14	Thioredoxin	30	133	Thioredoxin domain	9.20E-34
4	Locus36626v1rpkm2.63_9	5.2	5	-43	5	47	7	Thioredoxin	30	130	Thioredoxin domain	2.90E-29
4	Locus36626v1rpkm2.63_9	5.2	5	-43	5	47	7	Thioredoxin	160	258	Thioredoxin domain	7.80E-31
4	Locus4712v1rpkm64.38_5	5.4	8	-56	8	17	8	Thioredoxin	65	158	Thioredoxin domain	5.90E-29
4	Locus6776v1rpkm44.75_9	5.3	13	-143	5	64	25	Thioredoxin	434	538	Thioredoxin domain	2.30E-21
4	Locus6776v1rpkm44.75_9	5.3	13	-143	5	64	25	Thioredoxin	96	196	Thioredoxin domain	7.40E-32
4	Locus6806v1rpkm44.51_9	5.7	18	-180	5	26	23	Thioredoxin	106	208	Thioredoxin domain	1.80E-30
4	Locus15043v1rpkm17.17_2	3.7	1	-2.6	10	26	1	Tic22	30	237	Tic22-like	1.80E-64
4	Locus1571v1rpkm173.12_8	5.2	6	-60	6	25	6	TIM	6	232	Triosephosphate isomerase	1.20E-84
4	Locus15113v1rpkm17.05_3	2.8	1	-6.2	6	22	1	Tim17	54	154	Mitochondrial inner membrane translocase subunit Tim17/Tim22/Tim23/peroxisomal protein PMP24	3.60E-20
4	Locus22526v1rpkm8.92_2	4.7	2	-11	10	27	4	Tim17	53	175	Mitochondrial inner membrane translocase subunit Tim17/Tim22/Tim23/peroxisomal protein PMP24	7.40E-12
4	Locus28120v1rpkm5.58_5	4.6	3	-13	9	15	3	Tim17	28	143	Mitochondrial inner membrane translocase subunit Tim17/Tim22/Tim23/peroxisomal protein PMP24	4.10E-13
4	Locus9119v1rpkm32.36_20	3.9	1	-2.1	6	135	1	TIP120	1040	1198	TATA-binding protein interacting (TIP20)	2.50E-53
4	Locus10639v1rpkm26.79_15	4.1	2	-12	10	81	2	TLC	179	658	ADP/ATP carrier protein	7.20E-205
4	Locus11946v1rpkm23.15_9	4.3	2	-11	8	32	2	TLC	1	220	ADP/ATP carrier protein	4.60E-98
4	Locus5561v1rpkm54.84_4	3.4	1	-2.7	5	20	2	TPP_enzyme_C	33	158	Thiamine pyrophosphate enzyme, C-terminal TPP-binding	2.90E-16
4	Locus14674v1rpkm17.75_17	5	6	-37	6	109	10	TPPII	482	676	Peptidase S8A, tripeptidyl peptidase II	1.80E-63
4	Locus4838v1rpkm62.70_7	4	1	-3.2	6	55	2	TPR_1	401	434	Tetratricopeptide TPR-1	5.80E-05
4	Locus4838v1rpkm62.70_7	4	1	-3.2	6	55	2	TPR_1	367	399	Tetratricopeptide TPR-1	2.20E-06
4	Locus4838v1rpkm62.70_7	4	1	-3.2	6	55	2	TPR_2	224	256	Tetratricopeptide TPR2	1.90E-05
4	Locus1201v1rpkm210.03_5	5.7	5	-48	9	33	5	TPT	138	282	Domain of unknown function DUF250	3.40E-36
4	Locus6124v1rpkm49.61_5	3.5	1	-1.4	10	44	1	TPT	244	388	Domain of unknown function DUF250	3.20E-37
4	Locus8441v1rpkm35.36_3	2.7	1	-4.5	10	30	1	TPT	115	261	Domain of unknown function DUF250	5.30E-44
4	Locus1266v1rpkm203.87_5	4.1	2	-23	5	45	3	Transferase	3	383	Transferase	8.20E-70
4	Locus2337v1rpkm125.16_5	3.9	1	-1.6	8	43	1	Transferase	6	382	Transferase	5.80E-90
4	Locus4086v1rpkm73.26_8	5.7	7	-36	6	41	17	Transferase	5	326	Transferase	3.60E-32
4	Locus769v1rpkm294.92_15	4	2	-9.2	5	68	4	Transket_pyr	325	495	Transketolase-like, pyrimidine-binding domain	5.30E-42
4	Locus769v1rpkm294.92_15	4	2	-9.2	5	68	4	Transketolase_C	520	608	Transketolase, C-terminal	7.10E-08
4	Locus769v1rpkm294.92_15	4	2	-9.2	5	68	4	Transketolase_N	1	307	Transketolase, N-terminal	1.60E-135
4	Locus5419v1rpkm56.28_4	4	1	-3	5	27	2	TRAP_alpha	23	238	Translocon-associated protein (TRAP), alpha subunit	8.80E-23

4	Locus6252v1rpkm48.75_7	3.6	1	-2.5	5	28	1	TRAP_alpha	22	237	Translocon-associated protein (TRAP), alpha subunit	1.50E-24
4	Locus2162v1rpkm133.26_3	5.4	6	-48	10	21	6	TRAP_beta	14	191	Translocon-associated beta	3.40E-52
4	Locus4973v1rpkm61.05_6	4.6	2	-12	5	50	3	Tubulin	3	222	Tubulin/FtsZ, GTPase domain	1.40E-71
4	Locus1276v1rpkm202.86_4	4	1	-3.4	5	20	1	Tubulin_C	1	126	Tubulin/FtsZ, 2-layer sandwich domain	1.10E-47
4	Locus4973v1rpkm61.05_6	4.6	2	-12	5	50	3	Tubulin_C	261	382	Tubulin/FtsZ, 2-layer sandwich domain	2.10E-48
4	Locus19057v1rpkm12.00_4	4.3	1	-6.2	4	23	1	UBA	174	208	Ubiquitin-associated/translation elongation factor EF1B, N-terminal	6.50E-05
4	Locus3031v1rpkm98.14_2	5.6	4	-31	10	18	4	ubiquitin	6	74	Ubiquitin	1.90E-34
4	Locus4401v1rpkm68.83_3	4.7	4	-28	7	17	18	ubiquitin	82	150	Ubiquitin	1.90E-34
4	Locus4401v1rpkm68.83_3	4.7	4	-28	7	17	18	ubiquitin	6	74	Ubiquitin	1.90E-34
4	Locus8025v1rpkm37.44_6	3	1	-1.3	6	21	1	U-box	2	56	U box domain	8.00E-06
4	Locus1107v1rpkm224.69_1	4.9	1	-6.2	10	15	1	UCR_14kD	15	112	Cytochrome d ubiquinol oxidase, 14kDa subunit	1.60E-32
4	Locus14320v1rpkm18.34_7	4.4	3	-22	9	30	3	UCR_TM	95	146	Ubiquinol cytochrome reductase, transmembrane domain	5.70E-13
4	Locus5815v1rpkm52.29_8	4.5	3	-26	9	30	3	UCR_TM	96	147	Ubiquinol cytochrome reductase, transmembrane domain	4.00E-13
4	Locus10627v1rpkm26.82_29	4	2	-10	6	184	4	UDP-g_GGTase	992	1196	UDP-glucose:Glycoprotein Glucosyltransferase	2.40E-64
4	Locus1679v1rpkm164.47_6	4.3	5	-40	7	26	5	UDPGP	2	232	UTP--glucose-1-phosphate uridylyltransferase	4.30E-105
4	Locus1102v1rpkm225.42_5	5.4	8	-55	7	25	17	UDPGP	1	193	UTP--glucose-1-phosphate uridylyltransferase	9.60E-83
4	Locus7033v1rpkm42.89_6	4.7	3	-20	8	31	3	UDPGT	83	231	UDP-glucuronosyl/UDP-glucosyltransferase	7.90E-08
4	Locus26030v1rpkm6.73_4	5	2	-14	5	23	2	UPF0172	5	204	Uncharacterised protein family UPF0172	5.90E-61
4	Locus2703v1rpkm109.97_12	3.8	1	-4.7	5	92	1	UVR	419	454	UvrB/UvrC protein	1.50E-08
4	Locus706v1rpkm311.27_14	5.8	40	-366	6	93	153	V_ATPase_I	42	815	ATPase, V0/A0 complex, 116kDa subunit	2.80E-242
4	Locus15040v1rpkm17.17_2	5.7	8	-50	5	14	18	V-ATPase_C	5	128	ATPase, V1 complex, subunit C	2.00E-29
4	Locus3216v1rpkm92.68_6	6.4	32	-267	6	41	102	V-ATPase_C	5	367	ATPase, V1 complex, subunit C	5.70E-125
4	Locus8278v1rpkm36.22_7	6.3	22	-180	6	42	69	V-ATPase_C	5	367	ATPase, V1 complex, subunit C	1.90E-123
4	Locus18798v1rpkm12.27_12	6	16	-134	7	51	35	V-ATPase_H_C	329	442	ATPase, V1 complex, subunit H, C-terminal	1.10E-41
4	Locus10055v1rpkm28.77_6	6	13	-105	7	33	42	V-ATPase_H_C	165	279	ATPase, V1 complex, subunit H, C-terminal	1.60E-42
4	Locus4457v1rpkm67.94_2	5.6	10	-41	6	12	26	V-ATPase_H_C	1	93	ATPase, V1 complex, subunit H, C-terminal	3.10E-32
4	Locus18798v1rpkm12.27_12	6	16	-134	7	51	35	V-ATPase_H_N	5	323	ATPase, V1 complex, subunit H	2.70E-88
4	Locus10055v1rpkm28.77_6	6	13	-105	7	33	42	V-ATPase_H_N	2	159	ATPase, V1 complex, subunit H	1.50E-48
4	Locus2992v1rpkm99.31_8	6.5	34	-242	7	39	88	V-ATPase_H_N	7	325	ATPase, V1 complex, subunit H	2.60E-90
4	Locus1763v1rpkm156.53_6	6	18	-136	5	31	69	vATP-synt_AC39	5	262	ATPase, V0/A0 complex, subunit C/D	1.70E-71
4	Locus22759v1rpkm8.77_4	6.1	25	-181	5	41	64	vATP-synt_AC39	15	346	ATPase, V0/A0 complex, subunit C/D	1.00E-95
4	Locus24474v1rpkm7.63_5	6.8	55	-182	9	26	57	vATP-synt_E	16	225	ATPase, V1/A1 complex, subunit E	3.40E-78
4	Locus2452v1rpkm120.09_6	6.9	64	-218	9	27	89	vATP-synt_E	16	225	ATPase, V1/A1 complex, subunit E	7.70E-78
4	Locus3671v1rpkm81.40_4	6.2	26	-134	7	27	39	vATP-synt_E	16	225	ATPase, V1/A1 complex, subunit E	2.20E-81
4	Locus36984v1rpkm2.54_7	6	4	-32	9	27	4	vATP-synt_E	84	215	ATPase, V1/A1 complex, subunit E	4.10E-46
4	Locus6645v1rpkm45.74_5	6.2	22	-125	7	27	22	vATP-synt_E	16	225	ATPase, V1/A1 complex, subunit E	1.50E-78
4	Locus8842v1rpkm33.50_5	6.7	50	-141	8	30	76	vATP-synt_E	48	257	ATPase, V1/A1 complex, subunit E	4.20E-80
4	Locus9378v1rpkm31.37_5	6.3	23	-116	9	26	23	vATP-synt_E	16	225	ATPase, V1/A1 complex, subunit E	1.90E-78
4	Locus15878v1rpkm15.90_12	3.3	1	-9.7	5	88	1	Vps35	12	749	Vacuolar protein sorting-associated protein 35	0.00E+00
4	Locus16489v1rpkm14.99_13	4.8	1	-1.1	7	68	1	VWA	175	355	von Willebrand factor, type A	1.00E-21
4	Locus13210v1rpkm20.39_10	4.2	1	-2.2	9	98	1	VWA	396	535	von Willebrand factor, type A	3.40E-10
4	Locus44047v1rpkm1.42_6	5.2	1	-1.1	8	40	1	Wax2_C	189	352	Uncharacterised domain Wax2, C-terminal	4.10E-67
4	Locus1501v1rpkm179.93_4	5.5	15	-145	8	36	22	WD40	300	325	WD40 repeat, subgroup	3.40E-06
4	Locus1501v1rpkm179.93_4	5.5	15	-145	8	36	22	WD40	155	188	WD40 repeat, subgroup	2.20E-12
4	Locus1501v1rpkm179.93_4	5.5	15	-145	8	36	22	WD40	195	230	WD40 repeat, subgroup	8.20E-09
4	Locus1501v1rpkm179.93_4	5.5	15	-145	8	36	22	WD40	236	270	WD40 repeat, subgroup	5.10E-01
4	Locus1501v1rpkm179.93_4	5.5	15	-145	8	36	22	WD40	107	142	WD40 repeat, subgroup	6.30E-11
4	Locus1501v1rpkm179.93_4	5.5	15	-145	8	36	22	WD40	64	100	WD40 repeat, subgroup	2.10E-08
4	Locus1501v1rpkm179.93_4	5.5	15	-145	8	36	22	WD40	7	43	WD40 repeat, subgroup	3.20E-05
4	Locus1547v1rpkm175.02_5	4.9	8	-70	8	17	8	WD40	1	15	WD40 repeat, subgroup	9.70E-04
4	Locus1547v1rpkm175.02_5	4.9	8	-70	8	17	8	WD40	62	97	WD40 repeat, subgroup	8.80E-02
4	Locus1547v1rpkm175.02_5	4.9	8	-70	8	17	8	WD40	22	57	WD40 repeat, subgroup	2.60E-09
4	Locus1547v1rpkm175.02_5	4.9	8	-70	8	17	8	WD40	127	152	WD40 repeat, subgroup	3.90E-07

4	Locus22837v1rpkm8.73_7	3	1	-2.9	6	45	1	WD40	208	238	WD40 repeat, subgroup	4.70E-04
4	Locus22837v1rpkm8.73_7	3	1	-2.9	6	45	1	WD40	173	196	WD40 repeat, subgroup	1.30E-01
4	Locus22837v1rpkm8.73_7	3	1	-2.9	6	45	1	WD40	337	370	WD40 repeat, subgroup	2.60E-04
4	Locus10440v1rpkm27.48_8	4.9	3	-12	5	47	3	X8	314	398	X8	4.70E-22
4	Locus12436v1rpkm22.01_10	4.9	4	-40	5	53	6	X8	372	456	X8	2.40E-23
4	Locus46503v1rpkm1.19_9	3.1	1	-1.9	5	45	1	X8	290	374	X8	2.90E-27
4	Locus16489v1rpkm14.99_13	4.8	1	-1.1	7	68	1	zf-C3HC4	11	37	Zinc finger, C3HC4 RING-type	7.60E-05
4	Locus13210v1rpkm20.39_10	4.2	1	-2.2	9	98	1	zf-C3HC4	159	201	Zinc finger, C3HC4 RING-type	4.60E-05
4	Locus24470v1rpkm7.64_15	3.7	1	-1.3	6	101	1	zf-C3HC4	831	869	Zinc finger, C3HC4 RING-type	2.60E-07
4	Locus24616v1rpkm7.54_3	5.1	1	-1.4	6	29	1	zf-C3HC4	48	89	Zinc finger, C3HC4 RING-type	1.60E-09
4	Locus35848v1rpkm2.81_4	3.6	1	-2.7	9	36	1	zf-LYAR	30	57	Zinc finger, C2H2, LYAR-type	5.00E-15
4	Locus35848v1rpkm2.81_4	3.6	1	-2.7	9	36	1	zf-met	94	118	NULL	3.80E-06

Appendix F: multiple sequence alignment

V-ATPase clustal omega multiple sequence alignment

```
CLUSTAL O(1.2.1) multiple sequence alignment


2       ----------------------MASRYWMVSLPV--------------------QSSASS
3       ----------------------MASRYWVVSLPV--------------------QGSASS
4       ----------------------MASRYWVVSLPV--------------------QSSASS
1       MGDYANLSRGGGCCPTMDLFRSEAMQLVQIIIPMESAHVTLSYLGELGLLQFKDLNADKS
8       --------------------------------MAMDRAELST-----------EQVLKRDI
5       --------------------------------MDRAELST-----------EQVLKRDI
6       ------------------------------------------------------------
7       ------------------------------------------------------------


2       LWSRLQESVSKKAFDTPLYRFNAPDLRVGTLDSLLALSDDLLKSN--AFIEGVSH-----
3       LWSRLQESVSKKAFDTSLYRFNTPDLRVGTLDSLLALSDDLLKSN--AFIEGVSH-----
4       PWSRLQESVSKQAFDTPLYRFSTPDLRIGTLDSLLALSDDLLKSN--AFIEGVSH-----
1       PFQRT---------------YATQIKRCGEMARKLRLFKEQMT--KAGI----SPAAM--
8       PWET----------------YMTTKLITGTCLQLLRRYDHKSESQRAALLDDEGPAYVRV
5       PWEA----------------YITTKLISGTCLQLLRRYDHKSESQRAALLEDEGPAYVRV
6       ------------------------------------------------------------
7       ------------------------------------------------------------


2       --KIRRQI---EEMERAAG-VDGGALTVD-----------------GVPVDSYLTRFVWD
3       --KIRRQI---EELERAAG-VDGGALTVD-----------------GVPVDSYLTRFVWD
4       --KIRRQI---EELERAAG-VDGGALTVD-----------------GVPVDSYLTRFVWD
1       --PTARNHIHLDDLEIRLGELEAELIEVNANSE--------KLQRSYNELLEYMLVLRKA
8       FVSILRDISKEDTIEYVLALI-DEMLTANPKRARLFHDSSLSSTDTYEPFLRWL----WR
5       FVSILRDISKEETVEYVIALI-DEMLTANPKRAMLFHDSSLSSTDIYEPFLRWL----WN
6       ------------------------------------------------------------
7       ------------------------------------------------------------


2       EAKYPT------------------------------MSPLREIVDGIHVQVMVYYHLIPE
3       EAKYPT------------------------------VSPLREIVDGIHVQVAKIEDDM--
4       EAKYPT------------------------------MSPLREIVDGIHVQVAKIEDDM--
1       GEFFHSAQSNA--------------------TTEQREIEARQAGDGLDSPLLLEQEMLT-
8       GNWFIQEKSCKILSLIVSVRPKRLEGTVSNGEATHSKSTFTSINDVLDSLVEWLCSQMKN
5       GNWFIQEKSCKILSLIMSVRPKPHECIVSNGEATHSKSTFTSINDVLNSLVEWLCSQMRN
6       ----------------------------------------------MEWLCSQMKN
7       ------------------------------------------------------------


2       YQTEESMYNAVRRFGKVKYDTLRLPSTVVAREADGSVKFGQGEGSAYLFDPDIYT-----
3       -KVRSAEYNNVR-----------SQLNAINRKQTGSLAVRDL--SNLVKPEDIIT-----
4       -KVRSSEYNNVR-----------SQLNAINRKQTGSLAVRDL--SNLVKPEDIIT-----
1       ---DPS------KQVKLG--FVSGLVPKVKSMAFERILFRATRGNIFLKQAVIDDPVTDP
8       PS-HPS------RSVPIAVNCLSTLLRE----STVRASFVQADGVKL------LIPLITP
5       PS-HSS------RSVPIAINCLSTLLRE----STVRASFVQADGVKL------LIPLISP
```

```
6       PS-HPS------RSVPIAINCLSTLLRE----STVRASFVQADGVKL------LIPLISP
7       ------------------------------------------------------------


2       -DGRISVQGIDSVLFPPEDELLAWVKGFLGCCVFLG-----------------------
3       -SEHL----VTLLAVVPKYSQKDWLSSYETLTTYVVPRSSKKLHEDNEYALY-TVTLFGR
4       -SEHL----VTLLAVVPKYSQKDWLASYETLTTYVVPRSSTKLHEDNEYALY-TVTLFGR
1       VSGEKVVKNVFVIFYSGER---AKSKILKICEAF-G---------ANRYPFTEDVSKQMQ
8       ASTQQSTQL-----------------LYETCLCV-WLLSY--YDAAVDYLATTRV--LPR
5       ASTQQSIQL-----------------LYETCLCV-WLLSY--YDAAVDYLATARV--LPR
6       ASTQQSIQL-----------------LYETCLCV-WLLSY--YDAAVDYLATTRV--LPR
7       ------------------------------------------------------------


2       ------------------------------------------------------------
3       VADNFKTSAREKGFQIREFEYSPEAQEGRKQELEKLMQD---QDTLRSSLLQWCYA-SYG
4       VADNFKTSAREKGFQIREFEYSPEAQEGRKQELEKLMQD---QDTMRSSLLQWCYA-SYG
1       MIDEVSGKISELKTTID--------------I-GLIHRGNLLKNISYQFEQWNNLVRKE
8       LVEVVKGSTKEKVV-----------------------------RVVILTFRNLLAKG
5       LVEVVKGSTKEKVV-----------------------------RVVILTFRNLLSKG
6       LVEVVKGSTKEKVV-----------------------------RVVILTFRNLLSKG
7       ------------------------------------------------------------


2       ------------------------------------------------------------
3       EVFSSWIHFCAVRVFVESILRYGLPPSFLAAVLAPPTKSE----------------KKVR
4       EVFSSWMHFCAVRVFVESILRYGLPPSFLATVLAPPTKSE----------------KKVR
1       KSVYHTLNMLSLDVTKKCLVAEGWSPVFAT---NQIQDALQ----R-ATFDSKS---QVG
8       -TFGVQMVDLGLPQIVQSLKAQAWSDEDLLDALNQLEEGLKDNIRRLSSFDKYKQE----
5       -AFGAQMVDLGLPQIVQSLKAQAWSDEDLLDALNQLEEGLKDNIRRLSSFDKYRQEVLLG
6       TAFGAQMVDLGLPQIVQSLKAQAWSDEDLLDALNQLEEGLKDNIRRLSSFDKYKQEVLLG
7       ------------------------------------------------------------


2       ------------------------------------------------------------
3       SILEQL-CGNVNSTYWKAE--------------------------EDVSIAGLGGEVEA
4       SILERL-CGNVNSTYWKAE--------------------------EDVSIAGLGGEMDA
1       SIFQVLHTTELPPTYFQTNK-------YT-------------TAFQEIVDAYGIAKYQEA
8       ------------------------------------------------------------
5       HL--DWSPMHKDPGFWRENITKFEENDFQILRVLITITDTSNDPTALAVACYDLSQFMQC
6       HL--DWSTMHKDPGFWRENITNFEENDFQILRVLITIMDTSNDPTALAVACYDLSQFMQY
7       --------MHKDPGFWRENITNFEENDFQILRVLITIMDTSNDPTALAVACYDLSQFMQY


2       ------------------------------------------------------------
3       YPYV--------------------------------------------------------
4       HPYVSFTINIT-------------------------------------------------
1       NPGVYTIVTFPFLFAVMFGDWGHGLCLLAATLYFLFREKKLSSQKLGDIMEMTFGGRYVI
8       ------------------------------------------------------------
5       HPGGRIVVA----------------------DLKAKARVMKLMNHENSK---VTKSA----
6       HPGGRIVVA----------------------NLKAKERVMKLMNHENSE---VTKNA----
7       HPGGRIVAA----------------------DLKAKERVMKLMTHENAE---VTKNA----


2       ------------------------------------------------------------
3       ------------------------------------------------------------
4       ------------------------------------------------------------
1       LMMAVFSIYTGFIYNEFFSVPFEIFGHSAYACRDASCSDATTSGLIKVRPAYAFGVDPKW
8       ------------------------------------------------------------
5       -LLCIQRLFLSAKYASFLQA------------------------------------------
6       -LLCIQRLFLSSKYASFLQA------------------------------------------
7       -LLCIQRLFLSAKYASFLQA------------------------------------------


2       ------------------------------------------------------------
3       ------------------------------------------------------------
4       ------------------------------------------------------------
1       HGSRSELPFLNSLKMKMSILIGVAQMNLGIMLSYFNAKFFRNSVNVWFQFIPQLIFLNSL
8       ------------------------------------------------------------
5       ------------------------------------------------------------
```

```
6       ------------------------------------------------------------
7       ------------------------------------------------------------


2       ------------------------------------------------------------
3       ------------------------------------------------------------
4       ------------------------------------------------------------
1       FGYLSLLVIVKWCTGSQADLYHVMIYMFLSPTDDLGENQLFPGQRLLQLVLLALALIAVP
8       ------------------------------------------------------------
5       ------------------------------------------------------------
6       ------------------------------------------------------------
7       ------------------------------------------------------------


2       ------------------------------------------------------------
3       ------------------------------------------------------------
4       ------------------------------------------------------------
1       WMLFPKPFLLKKQHEERHQGQSYAILQSTDTDMLEEQDHGSHDHEEFDFSEVFVHQLIHT
8       ------------------------------------------------------------
5       ------------------------------------------------------------
6       ------------------------------------------------------------
7       ------------------------------------------------------------


2       ------------------------------------------------------------
3       ------------------------------------------------------------
4       ------------------------------------------------------------
1       IEFVLGAVSNTASYLRLWALSLAHSELSTVFYEKVLLLAWGYNNIFILLIGGIVFIFATV
8       ------------------------------------------------------------
5       ------------------------------------------------------------
6       ------------------------------------------------------------
7       ------------------------------------------------------------


2       -------------------------------------------
3       -------------------------------------------
4       -------------------------------------------
1       GVLLVMETLSAFLHALRLHWVEFQNKFYEGNGYKFSPYSFALL
8       -------------------------------------------
5       -------------------------------------------
6       -------------------------------------------
7       -------------------------------------------
```

Ppase cluustal omega multiple sequence alignment

CLUSTAL O(1.2.1) multiple sequence alignment

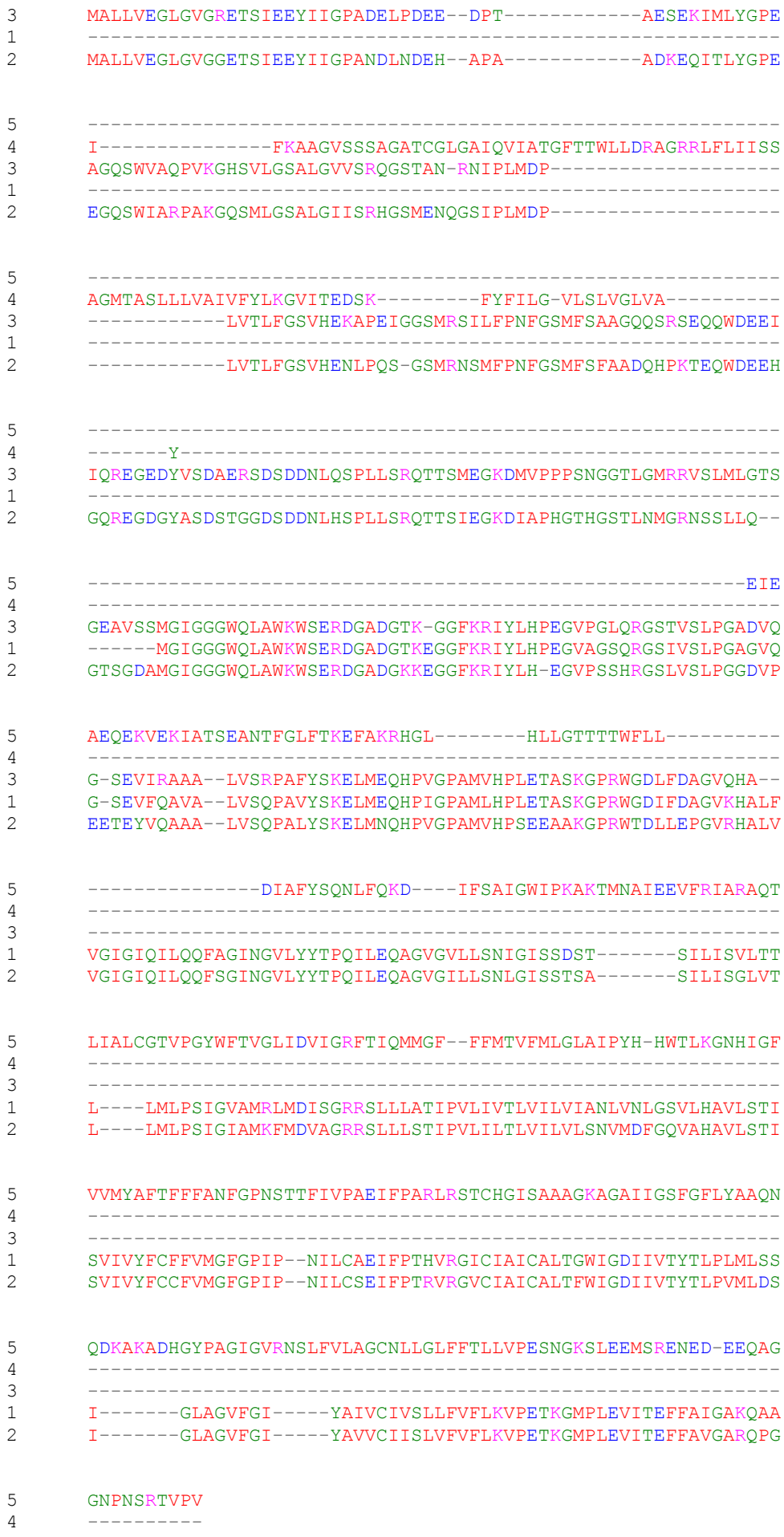
```
5       MGAAILSDLVTEILIPIAAVIGIAFSLVQWLLVAKVKLSPEAQTPGAHGGKKNGYSDYLI
2       MGAPVLSEFVTEIVIPVAAVIGIAFSLVQWLLVSKVKVSSDSHGAS-NKKKNGGYGDYLL
1       MGAPVLSDVITEILIPVAAVIGIAFSLVQWVLVSKVKLSPDSHGAN--SKKNGGYRDYLL
6       ------------------------------------------------------------
7       MGAPILSDVITEIVIPVAAVIGIAFSLFQWMLVSKVKLSPDSHGAN--SKKNGGHGDYLL
3       ------------------------------------------------------------
4       ------------------------------------------------------------


5       EEEEGLNDHNVVVKCAEIQSAISEGATSFLFTEYQYVGIFMAVFAVLIFVFLGSVEGFST
2       EEEEGISDHSVVSKCAEIQLAISEGATSFLFTEYQYVGVFMVIFAVLIFLFLGSVEGFST
1       EEEEGISDHSVVSKCAEIQSAISEGATSFLFTEYQYVGVFMVAFAALIFLFLGSVEGFST
6       ------------------------------------------------------------
7       EEEEGISDHSVVSKCAEIQSAISEGATSFLFTEYQYVGVFMVAFAALIFLFLGSVEGFST
3       ------------------------------------------------------------
4       ------------------------------------------------------------


5       ESRPCTYDKFKTCKPALSNAIFSTVSFLLGAITSVVSGFLGMKIATYANARTTLEARKGV
2       KGQPCTYSKGKTCKPALFNAIFSTVAFLLGAITSVVSGFLGMKIATYANARTTLEARKGV
1       K-----------------------------------------------------------
6       ------------------------------------------------------------
7       KGQPCTYSKDKTCKPALFNAIFSTVAFLLGAVTSVVSGFLGMKIATFANARTTLEARKGV
```

```
3       ------------------------------------------------------------
4       ------------------------------------------------------------

5       GKAFITAFRSGAVMGFLLAANGLLVLYISINLFKLYYGEDWEGLFEAITGYGLGGSSMAL
2       GKAFITAFRSGAVMGFLLAANGLLVLYIAINLFKLYYGDDWEGLFEAITGYGLGGSSMAL
1       ------------------------------------------------------------
6       ------------------------------------------------------------
7       GKAFITAFRSGAVMGFLLAANGLFVLYVSINLFKLYYGDDWEGLFEAITGYGLGGSSMAL
3       ------------------------------------------------------------
4       ------------------------------------------------------------

5       FGRVGGGIYTKAADVGADLVGKVERNIPEDDPRNPAVIADNVGDNVGDIAGMGSDLFGSY
2       FGRVGGGIYTKAADVGADLVGKVERNIPEDDPRNPAVIADNVGDNVGDIAGMGSDLFGSY
1       ------------------------------------------------------------
6       ----------------------------------------------------MGSDLFGSY
7       FGRVGGGIYTKAADVGADLVGKVERNIPEDDPRNPAVIADNVGDNVGDIAGMGSDLFGSY
3       ------------------------------------------------------------
4       ------------------------------------------------------------

5       AESSCAALVVASISSFGINHELTAMMYPLLISSMGIIVCLITTLFATDFFEIKDVKEIEP
2       AESSCAALVVASISSFGINHDFTGMCFPLLVSSMGIIVCLITTLFATDFFEIKAVKEIEP
1       ------------------------------------------------------------
6       AESSCAALVVASISSFGINHELTAMMYPLLVSSMGIIVCLITTLFATDFFEIKAVKEIEP
7       AESSCAALVVASISSFGINHDLTGMCYPLLVSSMGIIVCLITTLFATDFFEIKGVTEIEP
3       ------------------------------------------------------------
4       ------------------------------------------------------------

5       ALKKQLIISTALMTVGIAVVSWIALPASFTIFNFGVQKEVKNWELFFCVAIGLWAGLVIG
2       ALKKQLIISTALMTLGIALVSWLALPPSFTIFNFGAQKEVKNWELFFCVAIGLWAG----
1       ------------------------------------------------------------
6       ALKMQLIISTALMTVGIAVVSWISLPASFTIFNFGVQKEVKNWELFFCVAIGLWAGLVIG
7       ALKKQLIISTALMTVGIAVVSWLALPSSFTIFNFGAQKEVKNWELFFCVAIGLWAGLVIG
3       ------------------------------------------------------------
4       ------------MTLGIALVSWLALPSSFTIFNFGAQKEVKNWELFFCVAIGLWAGLVIG

5       FVTEYYTSNAYSPVQDVADSCRTGAATNVIFGLALGYKSVIIPIFAIAISIFVSFSFAVM
2       ------------------------------------------------------------
1       ------------------------------------------------------------
6       FVTEYYTSNAYSPVQDVADSCRTGAATNVIFGLALGYKSVIIPIFAIAISIFVSFSFAAM
7       FVTEYYTSNAYSPVQDVADSCRTGAATNVIFGLALGYKSVIIPIFAIAVSIFVSFSFAAM
3       ------------------------------------------------------------
4       FVTEYYTSNAYSPVQDVADSCRTGAATNVIFGLALGYKSVIIPIFAIAISIFVSFSLAAM

5       YGIAVAALGMLSTLATGLAIDAYGPICDNAGGIAEMAGMSHRIRERTDALDA--------
2       ------------------------------------------------------------
1       ------------------------------------------------------------
6       YGIAVAALGMLSTIATGLAIDAYGPISDNAGGIAEMAGMSHRIRERTDALDAAGNTTAAI
7       YGIAVAALGMLSTIATGLAIDAYGPISDNAGGIAEMAGMSHRVRERTDALDAAGNTTAAI
3       ------------------------------------------------------------
4       YGIAVAALGMLSTIATGLAIDAYGPISDNAGGIAEMAGMSHKIRERTDALDAAGNTTAAI

5       ------------------------------------------------------------
2       ------------------------------------------------------------
1       ------------------------------------------------------------
6       GKGFAIGSAALVSLALFGAFVSRAGISTVDVLTPKVFIGLLVGAMLPYWFSAMTMKSVGS
7       GKGFAIGSAALVSLALFGAFVSRAAISTVDVLTPKVFIGLIVGAMLPYWFSAMTMKSVGS
3       ------------------------------------------------------------
4       GKGFAIGSAALVSLALFGAFVSRAAISTVDVLTPKVFIGLIVGAMLPYWFSAMTMKSVGS

5       ------------------------------------------------------------
2       ------------------------------------------------------------
1       ------------------------------------------------------------
```

```
6       AALKMVEEVRRQFNTIPGLMEGTAKPDYATCVKISTDASIKEMIPPGALVMLTPLIVGTL
7       AALKMVEEVRRQFNTIPGLMEDTAKPDYATCVKISTDASIKEMIPPGALVMLTPLIVGTL
3       -------------------MEGTGKPDYATCVKISTDASIKEMIPPGALVMLTPLIVGIL
4       AALKMVEEVRRQFNTIPGLMEGTAKPDYATCVKISTDASIKEMIPPGALVMLTPLIVGTL


5       ------------------------------------------------------------
2       ------------------------------------------------------------
1       ------------------------------------------------------------
6       FGVETLSGVLAGSLVSGVQIAISASNTGGAWDNAKKYIEAGASEHA--------------
7       FGVETLSGVLAGSLVSGVQIAISASNTGGAWDNAKKYIEAGASEHARTLGPKGSDPHKAA
3       FGVETLSGVLAGSLVSGVQIAISASNTGGAWDNAKKYIEAGASEHARTLGPKGSDPHKAA
4       FGVETLSGVLAGSLVSGVQIAISASNTGGAWDNAKKYIEAGASDHARTLGPKGSDPHKAA


5       ---------------------------------------------
2       ---------------------------------------------
1       ---------------------------------------------
6       ---------------------------------------------
7       VIGDTIGDPLKDTSGPSLNILVKLMAVESLVFAPFFATHGGLLFKIF
3       VIGDTIGDPLKDTSGPSLNILIKLMAVESLVFAPFFATHGGLLFKIF
4       VIGDTIGDPLKDTSGPSLNILIKLMAVESLVFAPFFATHGGLLFKIF
```

Sugar transporters clustal omega multiple sequence alignment

```
CLUSTAL O(1.2.1) multiple sequence alignment


5       ----------------------------MGFFTDAYDLF-----------------CISLVT
4       MSFRGDESGGEDGGLRKPFLHTGSWYRMGMGSRQSSLMDKSSSGSVIRDSSVSVVLCTLI
3       ------------------------------------------------------MGAVLIAIA
1       ------------------------------------------------------------
2       ------------------------------------------------------------


5       KLLGRIYYHVDGSETPGV-----------LPPNVSAAVNGVAFCGTLLGQLFFGWLGDKM
4       VALGPIQFGFTGGYSSPTQDAIIKDLGLSIS--EFSIFGSLSNVGAMVGAIASGQIAEYI
3       AAIGNLLQGWDNATIAGSVLYIKKEFNLESEPAIEGLIVAMSLIGATVITTFSGAISDAF
1       ------------------------------------------------------------
2       ------------------------------------------------------------


5       GRKRVYGMTLMLMVICSVASGLSFGHKAKGVMATLCFFRFWLGFGIGGDYPLSATIMSEY
4       GRKGSLMIASIPNII----GWLAISF--AKDSSFLYMGRLLEGFGVGVISYTVPVYIAEI
3       GRRPMLIVSSLLYFL----SGIVMFC--SPNIYVLLLARLIDGLGIGLSVTLVPMYISET
1       ------------------------------------------------------------
2       ------------------------------------------------------------


5       ANKKTRGAFIAAVFAMQGFGILTGGAVALIVSAAFKNEFKAPTYEQNAVASTVPEADYVW
4       APQNMRGGLGSVNQLSVTIGIM--------LAYI-------------F--GMF----LPW
3       APSDIRGLLNTLPQFTGSCGMF--------LSYC-------------MVFGMSLRVKPDW
1       ------------------------------------------------------------
2       --------------------MF--------LSYC-------------MVFSMSLLPQPNW


5       RIILMFGALPAAMTYYW-RMKMPETARYTALVAKNAKQAAADMSKVLQV-----------
4       RLLAVMGVLP-CTVLIPGLFFIPESPRWLAKMGMM-----EDFEASLQVLRGFDTDISVE
3       RLMLGVLSIPSLLYFALTIFYLPESPRWLVSKGRM-----IEAKHVLQRLRGREDVS-GE
1       ------------------------------------------------------------
2       RLMLGVLSIPSLLYFALTIFYLPESPRWLVSKGRM-----TEAKKVLQRLRGREDVA-GE


5       ------------------------------------------------------------
4       VNEIKRSVASGTRRTTIRF-----SDLKQRRYKLPLMIGIGLLVLQQLSGINGILFYANN
```


3 MALLVEGLGVGRETSIEEYIIGPADELPDEE--DPT----------AESEKIMLYGPE
1 ----------
2 MALLVEGLGVGGETSIEEYIIGPANDLNDEH--APA----------ADKEQITLYGPE
5 ----------
4 I----------FKAAGVSSSAGATCGLGAIQVIATGFTTWLLDRAGRRLFLIISS
3 AGQSWVAQPVKGHSVLGSALGVVSRQGSTAN-RNIPLMDP----------
1 ----------
2 EGQSWIARPAKGQSMLGSALGIISRHGSMENQGSIPLMDP----------
5 ----------
4 AGMTASLLLVAIVFYLKGVITEDSK---------FYFILG-VLSLVGLVA----------
3 ----------LVTLFGSVHEKAPEIGGSMRSILFPNFGSMFSAAGQQSRSEQQWDEEI
1 ----------
2 ----------LVTLFGSVHENLPQS-GSMRNSMFPNFGSMFSFAADQHPKTEQWDEEH
5 ----------
4 -------Y----------
3 IQREGEDYVSDAERSDSDDNLQSPLLSRQTTSMEGKDMVPPPSNGGTLGMRRVSLMLGTS
1 ----------
2 GQREGDGYASDSTGGDSDDNLHSPLLSRQTTSIEGKDIAPHGTHGSTLNMGRNSSLLQ--
5 ----------EIE
4 ----------
3 GEAVSSMGIGGGWQLAWKWSERDGADGTK-GGFKRIYLHPEGVPGLQRGSTVSLPGADVQ
1 ------MGIGGGWQLAWKWSERDGADGTKEGGFKRIYLHPEGVAGSQRGSIVSLPGAGVQ
2 GTSGDAMGIGGGWQLAWKWSERDGADGKKEGGFKRIYLH-EGVPSSHRGSLVSLPGGDVP
5 AEQEKVEKIATSEANTFGLFTKEFAKRHGL--------HLLGTTTTWFLL----------
4 ----------
3 G-SEVIRAAA--LVSRPAFYSKELMEQHPVGPAMVHPLETASKGPRWGDLFDAGVQHA--
1 G-SEVFQAVA--LVSQPAVYSKELMEQHPIGPAMLHPLETASKGPRWGDIFDAGVKHALF
2 EETEYVQAAA--LVSQPALYSKELMNQHPVGPAMVHPSEEAAKGPRWTDLLEPGVRHALV
5 ----------DIAFYSQNLFQKD----IFSAIGWIPKAKTMNAIEEVFRIARAQT
4 ----------
3 ----------
1 VGIGIQILQQFAGINGVLYYTPQILEQAGVGVLLSNIGISSDST-------SILISVLTT
2 VGIGIQILQQFSGINGVLYYTPQILEQAGVGILLSNLGISSTSA-------SILISGLVT
5 LIALCGTVPGYWFTVGLIDVIGRFTIQMMGF--FFMTVFMLGLAIPYH-HWTLKGNHIGF
4 ----------
3 ----------
1 L----LMLPSIGVAMRLMDISGRRSLLLATIPVLIVTLVILVIANLVNLGSVLHAVLSTI
2 L----LMLPSIGIAMKFMDVAGRRSLLLSTIPVLILTLVILVLSNVMDFGQVAHAVLSTI
5 VVMYAFTFFFANFGPNSTTFIVPAEIFPARLRSTCHGISAAAGKAGAIIGSFGFLYAAQN
4 ----------
3 ----------
1 SVIVYFCFFVMGFGPIP--NILCAEIFPTHVRGICIAICALTGWIGDIIVTYTLPLMLSS
2 SVIVYFCCFVMGFGPIP--NILCSEIFPTRVRGVCIAICALTFWIGDIIVTYTLPVMLDS
5 QDKAKADHGYPAGIGVRNSLFVLAGCNLLGLFFTLLVPESNGKSLEEMSRENED-EEQAG
4 ----------
3 ----------
1 I-------GLAGVFGI-----YAIVCIVSLLFVFLKVPETKGMPLEVITEFFAIGAKQAA
2 I-------GLAGVFGI-----YAVVCIISLVFVFLKVPETKGMPLEVITEFFAVGARQPG
5 GNPNSRTVPV
4 ----------

```
3       ----------
1       GN--------
2       RT--------
```

Appendix G: Stomatal distribution and Densities

The stomata of three *Agave* cultivars varying in succulence were compared under contrasting water regimes.

Stomata in *Agave* occur on both surfaces of the leaves (amphistomatous). Impressions of the upper and lower surface of the leaf were made to measure stomatal characteristics, using clear nail varnish and tape. Once the nail varnish dried, clear tape was pressed gently over the area and peeled off, and placed on a microscope slide. Pictures were taken under the light microscope (Leica DM RB).At least 25 stomata were measured per leaf per surface (upper and lower). Stomatal dimensions in average 32.25 areas of $1mm^2$ were used to estimate stomatal density, under 40x magnification of light microscope. The stomata of three *Agave* cultivars varying in succulence were compared under contrasting water regimes.

References

Abd El-Hafez, S. (1990) 'Major problems and constraints of soil and water management in Kuwait', *Soil and Water Division, Public Authority for Agricultural A airs and Fisheries Resources, Kuwait*.

Acevedo, E., Badilla, I. and Nobel, P.S. (1983) 'Water relations, diurnal acidity changes, and productivity of a cultivated cactus, Opuntia ficus-indica', *Plant Physiology*, 72(3), pp. 775-780.

Adams, P., Thomas, J.C., Vernon, D.M., Bohnert, H.J. and Jensen, R.G. (1992) 'Distinct cellular and organismic responses to salt stress', *Plant and Cell Physiology*, 33(8), pp. 1215-1223.

Ahumada-Santos, Y.P., Montes-Avila, J., de Jesús Uribe-Beltrán, M., Díaz-Camacho, S.P., López-Angulo, G., Vega-Aviña, R., López-Valenzuela, J.Á., Heredia, J.B. and Delgado-Vargas, F. (2013) 'Chemical characterization, antioxidant and antibacterial activities of six Agave species from Sinaloa, Mexico', *Industrial Crops and Products*, 49, pp. 143-149.

Al-Ragom, F.A. (2004) 'Achieving energy efficiency in buildings that utilize subsidized electrical energy', *Energy engineering*, 101(2), pp. 16-38.

Alejandra, M., Paulo Tamaso, M., Adriana Yepes, M., Luciano, F. and Helenice, M. (2013) *CAM Photosynthesis in Bromeliads and Agaves: What Can We Learn from These Plants?*

Alexandersson, E., Saalbach, G., Larsson, C. and Kjellbom, P. (2004) 'Arabidopsis plasma membrane proteomics identifies components of transport, signal transduction and membrane trafficking', *Plant and Cell Physiology*, 45(11), pp. 1543-1556.

Antony, E. and Borland, A.M. (2009) 'The role and regulation of sugar transporters in plants with crassulacean acid metabolism', in *Progress in Botany*. Springer, pp. 127-143.

Antony, E., Taybi, T., Courbot, M., Mugford, S.T., Smith, J.A.C. and Borland, A.M. (2008) 'Cloning, localization and expression analysis of vacuolar sugar transporters in the CAM plant Ananas comosus (pineapple)', *Journal of Experimental Botany*, 59(7), pp. 1895-1908.

Arizaga, S. and Ezcurra, E. (2002) 'Propagation mechanisms in Agave macroacantha (Agavaceae), a tropical arid-land succulent rosette', *American Journal of Botany*, 89(4), pp. 632-641.

Arrizon, J., Morel, S., Gschaedler, A. and Monsan, P. (2010) 'Comparison of the water-soluble carbohydrate composition and fructan structures of Agave tequilana plants of different ages', *Food Chemistry*, 122(1), pp. 123-130.

Bartholomew, D.M., Rees, D.J.G., Rambaut, A. and Smith, J.A.C. (1996) 'Isolation and sequence analysis of a cDNA encoding the c subunit of a vacuolar-type H+-ATPase from the CAM plant Kalanchoe daigremontiana', *Plant Molecular Biology*, 31(2), pp. 435-442.

Bartholomew, D.P. and Kadzimin, S.B. (1977) 'Pineapple. In'Ecophysiology of Tropical Crops'.(Eds P. de T. Alvim and TT Kozlowski.) pp. 113-56'. Academic Press: New York.

Bettey, M. and Smith, J.A.C. (1993) 'Dicarboxylate transport at the vacuolar membrane of the CAM plant *Kalanchoë daigremontiana* sensitivity to protein-modifying and sulphydryl reagents', *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1152(2), pp. 270-279.

Black, C., Carnal, N.W. and Kenyon, W.H. (1982) 'Compartmentation and the regulation of CAM', *American Society of Plant Physiology*, Proceedings 5th Annual Symposium in Botany. eds I. P. Ting and M. Gibbs. Waverly Press, Baltimore, Maryland. , pp. 51-68.

Black, C.C., Chen, J.Q., Doong, R.L., Angelov, M.N. and Sung, S.J.S. (1996) 'Alternative Carbohydrate Reserves Used in the Daily Cycle of Crassulacean Acid Metabolism', in Winter, K. and Smith, J.A. (eds.) *Crassulacean Acid Metabolism*. Springer Berlin Heidelberg, pp. 31-45.

Blunden, G., Yi, Y. and Jewers, K. (1973) 'The comparative leaf anatomy of Agave, Beschorneria, Doryanthes and Furcraea species (Agavaceae: Agaveae)', *Botanical Journal of the Linnean Society*, 66(2), pp. 157-179.

Boller, T. and Wiemken, A. (1986) 'Dynamics of vacuolar compartmentation', *Annual Review of Plant Physiology*, 37(1), pp. 137-164.

Borland, A., Maxwell, K. and Griffiths, H. (2000) 'Ecophysiology of Plants with Crassulacean Acid Metabolism', in Leegood, R., Sharkey, T. and von Caemmerer, S. (eds.) *Photosynthesis*. Springer Netherlands, pp. 583-605.

Borland, A.M., Barrera Zambrano, V.A., Ceusters, J. and Shorrock, K. (2011) 'The photosynthetic plasticity of crassulacean acid metabolism: an evolutionary innovation for sustainable productivity in a changing world', *New Phytologist*, 191(3), pp. 619-633.

Borland, A.M. and Dodd, A.N. (2002) 'Carbohydrate partitioning in crassulacean acid metabolism plants: reconciling potential conflicts of interest', *Functional Plant Biology*, 29(6), pp. 707-716.

Borland, A.M. and Griffiths, H. (1989) 'The regulation of citric acid accumulation and carbon recycling during CAM in Ananas comosus', *Journal of Experimental Botany*, 40(1), pp. 53-60.

Borland, A.M. and Griffiths, H. (1997) 'A comparative study on the regulation of C3 and C4 carboxylation processes in the constitutive crassulacean acid metabolism (CAM) plant Kalanchoë daigremontiana and the C3-CAM intermediate Clusia minor', *Planta*, 201(3), pp. 368-378.

Borland, A.M., Griffiths, H., Broadmeadow, M.S.J., Fordham, M.C. and Maxwell, C. (1993) 'Short-term changes in carbon-isotope discrimination in the C3-CAM intermediate Clusia minor L. growing in Trinidad', *Oecologia*, 95(3), pp. 444-453.

Borland, A.M., Griffiths, H., Broadmeadow, M.S.J., Fordham, M.C. and Maxwell, C. (1994) 'Carbon-isotope composition of biochemical fractions and the regulation of carbon balance in leaves of the C3-crassulacean acid metabolism intermediate Clusia minor L. growing in Trinidad', *Plant Physiology*, 106(2), pp. 493-501.

Borland, A.M., Griffiths, H., Hartwell, J. and Smith, J.A.C. (2009) 'Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands', *Journal of Experimental Botany*, 60(10), pp. 2879-2896.

Borland, A.M., Griffiths, H., Maxwell, C., Fordham, M.C. and Broadmeadow, M.S.J. (1996) 'CAM induction in Clusia minor L. during the transition from wet to dry season in Trinidad: the role of organic acid speciation and decarboxylation', *Plant, Cell & Environment*, 19(6), pp. 655-664.

Borland, A.M., Hartwell, J., Jenkins, G.I., Wilkins, M.B. and Nimmo, H.G. (1999) 'Metabolite Control Overrides Circadian Regulation of Phosphoenolpyruvate Carboxylase Kinase and CO_2 Fixation in Crassulacean Acid Metabolism', *Plant Physiology*, 121(3), pp. 889-896.

Borland, A.M. and Taybi, T. (2004) 'Synchronization of metabolic processes in plants with Crassulacean acid metabolism', *Journal of Experimental Botany*, 55(400), pp. 1255-1265.

Borland, A.M., Tesci LI, Leegood RC and Walker RP. (1998) 'Inducibility of Crassulacean acid metabolism in Clusia species: physiological/biochemical characterization and intercellular localization of carboxylation and decarboxylation processes in three species which exhibit different degrees of CAM. ', *Planta* 205, pp. 342-351.

Bradford, M.M. (1976) 'A Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding', *Analytical Biochemistry*, 72, pp. 248-254.

Bremberger, C. and Lüttge, U. (1992) 'Dynamics of tonoplast proton pumps and other tonoplast proteins of Mesembryanthemum crystallinum L. during the induction of Crassulacean acid metabolism', *Planta*, 188(4), pp. 575-580.

Büttner, M. (2007) 'The monosaccharide transporter (-like) gene family in Arabidopsis', *FEBS Letters*, 581(12), pp. 2318-2324.
Carere, C.R., Sparling, R., Cicek, N. and Levin, D.B. (2008) 'Third generation biofuels via direct cellulose fermentation', *International Journal of Molecular Sciences*, 9(7), pp. 1342-1360.

Carnal, N.W. and Black, C.C. (1989) 'Soluble Sugars as the Carbohydrate Reserve for CAM in Pineapple Leaves Implications for the Role of Pyrophosphate: 6-Phosphofructokinase in Glycolysis', *Plant Physiology*, 90(1), pp. 91-100.

Carpita, N.C. and McCann, M.C. (2008) 'Maize and sorghum: genetic resources for bioenergy grasses', *Trends in Plant Science*, 13(8), pp. 415-420.
Carroll, A. and Somerville, C. (2009) 'Cellulosic Biofuels', *Annual Review of Plant Biology*, 60(1), pp. 165-182.

Carter, P.J., Fewson, C.A., Nimmo, G.A., Nimmo, H.G. and Wilkins, M.B. (1996) 'Roles of circadian rhythms, light and temperature in the regulation of phosphoenolpyruvate carboxylase in crassulacean acid metabolism', in *Crassulacean acid metabolism*. Springer, pp. 46-52.

Carter, P.J., Nimmo, H.G., Fewson, C.A. and Wilkins, M.B. (1990) 'Bryophyllum fedtschenkoi protein phosphatase type 2A can dephosphorylate phosphoenolpyruvate carboxylase', *FEBS letters*, 263(2), pp. 233-236.

Carter, P.J., Nimmo, H.G., Fewson, C.A. and Wilkins, M.B. (1991) 'Circadian rhythms in the activity of a plant protein kinase', *The EMBO journal*, 10(8), p. 2063.

Cedeño, M.C. (1995) 'Tequila production', *Critical Reviews in Biotechnology*, 15(1), pp. 1-11.

Ceusters, J. and Borland, A.M. (2011) 'Impacts of elevated CO2 on the growth and physiology of plants with crassulacean acid metabolism', in *Progress in Botany 72*. Springer, pp. 163-181.

Ceusters, J., Borland, A.M., Londers, E., Verdoodt, V., Godts, C. and De Proft, M.P. (2009) 'Differential usage of storage carbohydrates in the CAM bromeliad Aechmea ‘Maya’during acclimation to drought and recovery from dehydration', *Physiologia plantarum*, 135(2), pp. 174-184.

Chávez-Guerrero, L. (2013) 'Agave Salmiana Plant as an Alternative Source of Energy: Use of Abandoned and Arid Land'.

Chávez-Guerrero, L. and Hinojosa, M. (2010) 'Bagasse from the mezcal industry as an alternative renewable energy produced in arid lands', *Fuel*, 89(12), pp. 4049-4052.

Chen, L.-S. and Nose, A. (2004) 'Day–night changes of energy-rich compounds in crassulacean acid metabolism (CAM) species utilizing hexose and starch', *Annals of botany*, 94(3), pp. 449-455.

Christopher, J.T. and Holtum, J.A.M. (1996) 'Patterns of carbon partitioning in leaves of crassulacean acid metabolism species during deacidification', *Plant Physiology*, 112(1), pp. 393-399.

Christopher, J.T. and Holtum, J.A.M. (1998) 'Carbohydrate partitioning in the leaves of Bromeliaceae performing C3 photosynthesis or Crassulacean acid metabolism', *Functional Plant Biology*, 25(3), pp. 371-376.

Cleveland, C.J. (2007) 'Encyclopedia of earth', *Washington, DC: Environmental*.

Cockburn W, T.I., Sternberg LO (1979) 'Relationships between Stomatal Behavior and Internal carbon dioxide concentration in Crassulacean Acid Metabolism plants', *Plant Physiology*, 63, pp. 1029-1032.

Colunga-García Marín, P., Larqué Saavedra, A., Eguiarte, L.E. and Zizumbo-Villareal, D. (2007) 'En lo ancestral hay futuro: del tequila, los mezcales y otros agaves', *Centro de Investigación Cient fica de Yucatán, AC, Mérida, Yucatán, México*.

Coninck, B.d., Ende, W., Roy, K.l., Norio, S., Noureddine, B. and Shuichi, O. (2007) 'Fructan ExoHydrolases (FEHs) in plants: properties, occurrence and 3-D structure', *Recent advances in fructooligosaccharides research*, pp. 157-179.

Crafts-Brandner, S.J. and Salvucci, M.E. (2000) 'Rubisco activase constrains the photosynthetic potential of leaves at high temperature and CO_2', *Proceedings of the National Academy of Sciences*, 97(24), pp. 13430-13435.

Croome, D.J. (1991) 'The determinants of architectural form in modern buildings within the Arab world', *Building and Environment*, 26(4), pp. 349-362.

Cunningham, S.A., Summerhayes, B. and Westoby, M. (1999) 'Evolutionary divergences in leaf structure and chemistry, comparing rainfall and soil nutrient gradients', *Ecological Monographs*, 69(4), pp. 569-588.

Cushman, J.C. (2001) 'Crassulacean acid metabolism. A plastic photosynthetic adaptation to arid environments', *Plant Physiology*, 127(4), pp. 1439-1448.

Cushman, J.C., Agarie, S., Albion, R.L., Elliot, S.M., Taybi, T. and Borland, A.M. (2008) 'Isolation and characterization of mutants of common ice plant deficient in crassulacean acid metabolism', *Plant Physiology*, 147(1), pp. 228-238.

da Silva, B.P., de Sousa, A.C., Silva, G.M., Mendes, T.P. and Parente, J.P. (2002) 'A new bioactive steroidal saponin from Agave attenuata', *Zeitschrift fur Naturforschung C*, 57(5/6), pp. 423-428.

Davis, S.C., Dohleman, F.G. and Long, S.P. (2011a) 'The global potential for Agave as a biofuel feedstock', *GCB Bioenergy*, 3(1), pp. 68-78.

Davis, S.C., Griffiths, H., Holtum, J., Saavedra, A.L. and Long, S.P. (2011b) 'The Evaluation of Feedstocks in GCBB Continues with a Special Issue on Agave for Bioenergy', *GCB Bioenergy*, 3(1), pp. 1-3.

Davis, S.C. and Long, S.P. (2015) 'Sisal/Agave', in *Industrial Crops*. Springer, pp. 335-349.

de Santo, A.V., Alfani, A., Russo, G. and Fioretto, A. (1983) 'Relationship between CAM and succulence in some species of Vitaceae and Piperaceae', *Botanical gazette*, pp. 342-346.

Dittrich, P. (1976) 'Nicotinamide adenine dinucleotide-specific “malic” enzyme in Kalanchoe daigremontiana and other plants exhibiting Crassulacean acid metabolism', *Plant physiology*, 57(2), pp. 310-314.

Dittrich, P., Campbell, W.H. and Black, C.C. (1973) 'Phosphoenolpyruvate carboxykinase in plants exhibiting crassulacean acid metabolism', *Plant Physiology*, 52(4), pp. 357-361.

Dodd, A.N., Borland, A.M., Haslam, R.P., Griffiths, H. and Maxwell, K. (2002) 'Crassulacean acid metabolism: plastic, fantastic', *Journal of Experimental Botany*, 53(369), pp. 569-580.

Dodd, A.N., Griffiths, H., Taybi, T., Cushman, J.C. and Borland, A.M. (2003) 'Integrating diel starch metabolism with the circadian and environmental regulation of Crassulacean acid metabolism in Mesembryanthemum crystallinum', *Planta*, 216(5), pp. 789-797.

Dubois, M., K.A., G., K, H.J., P.A., R. and Fred, S. (1956) 'Colorimetric Method for Determination of Sugars and Related Substances', *Analytical Chemistry*, 28(3), pp. 350-356.

Duncan, A.C., Jäger, A.K. and van Staden, J. (1999) 'Screening of Zulu medicinal plants for angiotensin converting enzyme (ACE) inhibitors', *Journal of Ethnopharmacology*, 68(1), pp. 63-70.

Ehrler, W.L. (1969) 'Daytime stomatal closure in Agave americana as related to enhanced water-use efficiency'.

Eickmeier, W.G. and Adams, M.S. (1978) 'Gas Exchange in Agave lecheguilla Torr. (Agavaceae) and Its Ecological Implications', *The Southwestern Naturalist*, 23(3), pp. 473-485.

El-Katiri, L. and Husain, M. (2014) 'Prospects for Renewable Energy in GCC States…', *art. cité*, p. 14.

Elenga, R.G., Dirras, G.F., Maniongui, J.G., Djemia, P. and Biget, M.P. (2009) 'On the microstructure and physical properties of untreated raffia textilis fiber', *Composites Part A: Applied Science and Manufacturing*, 40(4), pp. 418-422.

Emmerlich, V., Linka, N., Reinhold, T., Hurth, M.A., Traub, M., Martinoia, E. and Neuhaus, H.E. (2003) 'The plant homolog to the human sodium/dicarboxylic cotransporter is the vacuolar malate carrier', *Proceedings of the National Academy of Sciences*, 100(19), pp. 11122-11126.

Endler, A., Meyer, S., Schelbert, S., Schneider, T., Weschke, W., Peters, S.W., Keller, F., Baginsky, S., Martinoia, E. and Schmidt, U.G. (2006) 'Identification of a vacuolar sucrose transporter in barley and Arabidopsis mesophyll cells by a tonoplast proteomic approach', *Plant Physiology*, 141(1), pp. 196-207.

Escamilla-Treviño, L.L. (2012) 'Potential of plants from the genus Agave as bioenergy crops', *Bioenergy Research*, 5(1), pp. 1-9.

FAO (2010) 'Food and Agricultural Organization (FAO) of the United Nations, FAO Statistics Division. Vol. accessed June 29, 2010.

FAO (2012) 'Food and Agricultural Organization (FAO) of the United Nations, FAO Statistics Division. Vol. accessed November 29, 2014.'.

Flores-Sahagun, T.H.S., Dos Santos, L.P., Dos Santos, J., Mazzaro, I. and Mikowski, A. (2013) 'Characterization of blue agave bagasse fibers of Mexico', *Composites Part A: Applied Science and Manufacturing*, 45, pp. 153-161.

French, A.D. (1989) 'Chemical and physical properties of fructans', *Journal of Plant Physiology*, 134(2), pp. 125-136.

Frenken, K. (2009) 'Irrigation in the Middle East region in figures AQUASTAT Survey-2008', *Water Reports*, (34).

Freschi, L., Takahashi, C.A., Cambui, C.A., Semprebom, T.R., Cruz, A.B., Mioto, P.T., de Melo Versieux, L., Calvente, A., Latansio-Aidar, S.R., Aidar, M.P. and Mercier, H. (2010) 'Specific leaf areas of the tank bromeliad Guzmania monostachia perform distinct functions in response to water shortage', *J Plant Physiol*, 167(7), pp. 526-33.

Gallagher, S.R. and Leonard, R.T. (1982) 'Effect of vanadate, molybdate, and azide on membrane-associated ATPase and soluble phosphatase activities of corn roots', *Plant Physiology*, 70(5), pp. 1335-1340.

Garcia-Moya, E., Romero-Manzanares, A. and Nobel, P.S. (2011) 'Highlights for Agave productivity', *Gcb Bioenergy*, 3(1), pp. 4-14.

Garcia-Reyes, R. and Rangel-Mendez, J. (2009) 'Contribution of agro-waste material main components (hemicelluloses, cellulose, and lignin) to the removal of chromium (III) from aqueous solution', *Journal of chemical technology and biotechnology*, 84(10), pp. 1533–1538.

García Mendoza, A. (2002) 'Distribution of agave (Agavaceae) in Mexico', *Cact. Succ. J.(USA)*, 74(4), pp. 177-187.

Garnier, E. (1992) 'Growth analysis of congeneric annual and perennial grass species', *Journal of Ecology*, pp. 665-675.

Garnier, E. and Laurent, G. (1994) 'Leaf anatomy, specific mass and water content in congeneric annual and perennial grass species', *New Phytologist*, 128(4), pp. 725-736.

Gentry, H.S. (2004) *Agaves of Continental North America*. University of Arizona Press.

Gibeaut, D.M. and Thomson, W.W. (1989) 'Leaf ultrastructure of Peperomia obtusifolia, P. camptotricha, and P. scandens', *Botanical Gazette*, pp. 108-114.

Gibson, A.C. (1982) *Crassulacean acid metabolism.[Proc. V ann. Symp. Bot., Univ. Calif., Riverside.] Amer. Soc. Plant Physiologists: rockville, USA*.

Gil, F. (1986) 'Origin of CAM as an alternative photosynthetic carbon fixation pathway', *Photosynthetica*, 20(4), pp. 494-507.

Gómez-Pompa, A. (1963) 'El género Agave'.

González, H.H., Del Real, L.J.I. and Solís, A.J.F. (2007) 'Manejo de plagas del agave tequilero', *Colegio de Postgraduados y Tequila Sauza, SA de CV México. 123p*.

Good-Avila, S.V., Souza, V., Gaut, B.S. and Eguiarte, L.E. (2006) 'Timing and rate of speciation in Agave (Agavaceae)', *Proceedings of the National Academy of Sciences*, 103(24), pp. 9124-9129.

Graham, E.A. and Nobel, P.S. (1996) 'Long-term effects of a doubled atmospheric CO_2 concentration on the CAM species Agave deserti', *Journal of Experimental Botany*, 47(1), pp. 61-69.

Grams, T.E.E., Borland, A.M., Roberts, A., Griffiths, H., Beck, F. and Luttge, U. (1997) 'On the mechanism of reinitiation of endogenous crassulacean acid metabolism rhythm by temperature changes', *Plant Physiology*, 113(4), pp. 1309-1317.

Griffiths, H. (1988) *Crassulacean acid metabolism: a re-appraisal of physiological plasticity in form and function*. Academic Press.

Griffiths, H. (1989) 'Crassulacean Acid Metabolism: a Re-appraisal of Physiological Plasticity in Form and Function', in Callow, J.A. (ed.) *Advances in Botanical Research*. Academic Press, pp. 43-92.

Griffiths, H. (1992) 'Carbon isotope discrimination and the integration of carbon assimilation pathways in terrestrial CAM plants', *Plant, Cell & Environment*, 15(9), pp. 1051-1062.

Griffiths, H., Broadmeadow, M.S.J., Borland, A.M. and Hetherington, C.S. (1990) 'Short-term changes in carbon-isotope discrimination identify transitions between C3 and C4 carboxylation during Crassulacean acid metabolism', *Planta*, 181(4), pp. 604-610.

Griffiths, H., Robe, W.E., Girnus, J. and Maxwell, K. (2008) 'Leaf succulence determines the interplay between carboxylase systems and light use during Crassulacean acid metabolism in Kalanchoë species', *Journal of Experimental Botany*, 59(7), pp. 1851-1861.

Güçlü-Üstündağ, Ö. and Mazza, G. (2007) 'Saponins: properties, applications and processing', *Critical Reviews in Food Science and Nutrition*, 47(3), pp. 231-258.

Gupta, N., Bandeira, N., Keich, U. and Pevzner, P.A. (2011) 'Target-decoy approach and false discovery rate: when things may go wrong', *Journal of the American Society for Mass Spectrometry*, 22(7), pp. 1111-1120.

Hafke, J.B., Hafke, Y., Smith, J.A.C., Lüttge, U. and Thiel, G. (2003) 'Vacuolar malate uptake is mediated by an anion-selective inward rectifier', *The Plant Journal*, 35(1), pp. 116-128.

Hajiah, A.E.H. (2006) *Proceedings of the Fifteenth Symposium on Improving Building Systems in Hot and Humid Climates, Orlando, FL*.

Hartsock, T.L. and Nobel, P.S. (1976) 'Watering converts a CAM plant to daytime CO2 uptake', *Nature*, 262(5569), pp. 574-576.

Hartwell, J., Gill, A., Nimmo, G.A., Wilkins, M.B., Jenkins, G.I. and Nimmo, H.G. (1999) 'Phosphoenolpyruvate carboxylase kinase is a novel protein kinase regulated at the level of expression', *The Plant Journal*, 20(3), pp. 333-342.

Hartwell, J., Smith, L.H., Wilkins, M.B., Jenkins, G.I. and Nimmo, H.G. (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(6), pp. 1071-1078.

Hedrich, R., Kurkdjian, A., Guern, J. and Flügge, U.I. (1989) 'Comparative studies on the electrical properties of the H+ translocating ATPase and pyrophosphatase of the vacuolar-lysosomal compartment', *The EMBO journal*, 8(10), p. 2835.

Heldt, H.-W. and Piechulla, B. (2004) *Plant biochemistry*. Academic Press.

Holtum, J.A.M., Smith, J.A.C. and Neuhaus, H.E. (2005) 'Intracellular transport and pathways of carbon flow in plants with crassulacean acid metabolism', *Functional Plant Biology*, 32(5), pp. 429-449.

Holum, J.R. (1994) 'Elements of General, Organic and Biological Chemistry, 1995'. John Wiley & Sons Publishing. Fundamentals of General, Organic and Biological Chemistry.

Honda, H., Okamoto, T. and Shimada, H. (1996) 'Isolation of a cDNA for a phosphoenolpyruvate carboxylase from a monocot CAM-plant, Aloe arborescens: structure and its gene expression', *Plant and cell physiology*, 37(6), pp. 881-888.

Idso, S.B., Kimball, B.A., Anderson, M.G. and Szarek, S.R. (1986) 'Growth response of a succulent plant, Agave vilmoriniana, to elevated CO2', *Plant physiology*, 80(3), pp. 796-797.

Iñiguez-Covarrubias, G., Díaz-Teres, R., Sanjuan-Dueñas, R., Anzaldo-Hernández, J. and Rowell, R. (2001b) 'Utilization of by-products from the tequila industry. Part 2: Potential value of *Agave tequilana* Weber azul leaves', *Bioresource Technology*, 77(2), pp. 101-108.

Iñiguez-Covarrubias, G., Dıaz-Teres, R., Sanjuan-Dueñas, R., Anzaldo-Hernández, J. and Rowell, R.M. (2001) 'Utilization of by-products from the tequila industry. Part 2: potential value of Agave tequilana Weber azul leaves', *Bioresource Technology*, 77(2), pp. 101-108.

Jaquinod, M., Villiers, F., Kieffer-Jaquinod, S., Hugouvieux, V., Bruley, C., Garin, J. and Bourguignon, J. (2007) 'A proteomics dissection of Arabidopsis thaliana

vacuoles isolated from cell culture', *Molecular & Cellular Proteomics*, 6(3), pp. 394-412.

Johnson, R. (2006) 'How to sequence tryptic peptides using low energy CID data'.

Johnson, R. and Ryan, C.A. (1990) 'Wound-inducible potato inhibitor II genes: enhancement of expression by sucrose', *Plant molecular biology*, 14(4), pp. 527-536.

Kaiser, G. and Heber, U. (1984) 'Sucrose transport into vacuoles isolated from barley mesophyll protoplasts', *Planta*, 161(6), pp. 562-568.

Kaul, R.B. (1977) 'The role of the multiple epidermis in foliar succulence of Peperomia (Piperaceae)', *Botanical Gazette*, pp. 213-218.

Keeley, J.E. and Rundel, P.W. (2003) 'Evolution of CAM and C4 Carbon-Concentrating Mechanisms', *International Journal of Plant Sciences*, 164(S3), pp. S55-S77.

Kenyon, W.H., Severson, R.F. and Black, C.C. (1985) 'Maintenance Carbon Cycle in Crassulacean Acid Metabolism Plant Leaves Source and Compartmentation of Carbon for Nocturnal Malate Synthesis', *Plant Physiology*, 77(1), pp. 183-189.

Kessner, D., Chambers, M., Burke, R., Agus, D. and Mallick, P. (2008) 'ProteoWizard: open source software for rapid proteomics tools development', *Bioinformatics*, 24(21), pp. 2534-2536.

Kirby, R.H. (1963) 'Vegetable fibres, botany, cultivation and utilization', *Vegetable fibres, botany, cultivation and utilization.*

Kluge, C., Lahr, J., Hanitzsch, M., Bolte, S., Golldack, D. and Dietz, K.-J. (2003) 'New insight into the structure and regulation of the plant vacuolar H+-ATPase', *Journal of Bioenergetics and Biomembranes*, 35(4), pp. 377-388.

Kluge, M. and Brulfert, J. (1996) 'Crassulacean acid metabolism in the genus Kalanchoë: ecological, physiological and biochemical aspects', in *Crassulacean acid metabolism*. Springer, pp. 324-335.

Kluge, M., Brulfert, J., Lipp, J., Ravelomanana, D. and Ziegler, H. (1993) 'A Comparative Study by δ13C-Analysis of Crassulacean Acid Metabolism (CAM) in Kalanchoë (Crassulaceae) Species of Africa and Madagascar', *Botanica Acta*, 106(4), pp. 320-324.

Kluge, M., Razanoelisoa, B. and Brulfert, J. (2001) 'Implications of Genotypic Diversity and Phenotypic Plasticity in the Ecophysiological Success of CAM Plants, Examined by Studies on the Vegetation of Madagascar1', *Plant Biology*, 3(3), pp. 214-222.

Kristen, U. (1969) 'Untersuchungen uber den Zusammenhang zwischen dem CO_2-Gaswechsel und der Luftwegigkeit an den CAM-Sukkulenten Bryophyllum daigremontianum Berg. und Agave americana L', *Flora Abt A Physiol Biochem*.
Laemmli (1970) 'CLEAVAGE OF STRUCTURAL PROTEINS DURING ASSEMBLY OF HEAD OF BACTERIOPHAGE-T4.', *Nature*, 227, pp. 680-685.

Lalonde, S., Wipf, D. and Frommer, W.B. (2004) 'Transport mechanisms for organic forms of carbon and nitrogen between source and sink', *Annu. Rev. Plant Biol.*, 55, pp. 341-372.

Lambers, H. and Poorter, H. (1992) 'Inherent variation in growth rate between higher plants: a search for physiological causes and ecological consequences'.

Lawlor, D.W. and Cornic, G. (2002) 'Photosynthetic carbon assimilation and associated metabolism in relation to water deficits in higher plants', *Plant, Cell & Environment*, 25(2), pp. 275-294.

Le Houerou, H.N. (1984) 'Rain use efficiency: a unifying concept in arid-land ecology', *Journal of Arid Environments*, 7(3), pp. 213-247.

Leegood, R.C. and Osmond, C.B. (1990) *The flux of metabolites in C4 and CAM plants*. Longman Sci. Tech. Essex, UK.

Lemoine, R. (2000) 'Sucrose transporters in plants: update on function and structure', *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1465(1), pp. 246-262.

Lepiniec, L., Vidal, J., Chollet, R., Gadal, P. and Crétin, C. (1994) 'Phosphoenolpyruvate carboxylase: structure, regulation and evolution', *Plant Science*, 99(2), pp. 111-124.

Lewis, D.H. (1984) *Storage Carbohydrates in Vascular Plants: Distribution, Physiology, and Metabolism*. CUP Archive.

Li, H., Foston, M., Kumar, R., Samuel, R., Gao, X., Hu, F., Ragauskas, A. and Wyman, C. (2012a) 'Chemical composition and characterization of cellulose for Agave as a fast-growing, drought-tolerant biofuels feedstock', *RSC Advances*, 2(11), pp. 4951–4958.

Li, H., Foston, M.B., Kumar, R., Samuel, R., Gao, X., Hu, F., Ragauskas, A.J. and Wyman, C.E. (2012b) 'Chemical composition and characterization of cellulose for Agave as a fast-growing, drought-tolerant biofuels feedstock', *RSC Advances*, 2(11), pp. 4951-4958.

Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S.L., Kebrom, T.H., Provart, N., Patel, R. and Myers, C.R. (2010) 'The developmental dynamics of the maize leaf transcriptome', *Nature genetics*, 42(12), pp. 1060-1067.

Linton, M.J. and Nobel, P.S. (1999) 'Loss of water transport capacity due to xylem cavitation in roots of two CAM succulents', *American journal of botany*, 86(11), pp. 1538-1543.

Liu, Z., Mouradov, A., Smith, K.F. and Spangenberg, G. (2011) 'An improved method for quantitative analysis of total fructans in plant tissues', *Analytical Biochemistry*, 418(2), pp. 253-259.

Livingston Iii, D.P., Hincha, D.K. and Heyer, A.G. (2009) 'Fructan and its relationship to abiotic stress tolerance in plants', *Cellular and Molecular Life Sciences*, 66(13), pp. 2007-2023.

López-Alvarez, A., Díaz-Pérez, A.L., Sosa-Aguirre, C., Macías-Rodríguez, L. and Campos-García, J. (2012) 'Ethanol yield and volatile compound content in fermentation of agave must by Kluyveromyces marxianus UMPe-1 comparing with Saccharomyces cerevisiae baker's yeast used in tequila production', *Journal of bioscience and bioengineering*, 113(5), pp. 614-618.

López, M.G. and Mancilla-Margalli, N.A. (2007) 'The nature of fructooligosaccharides in Agave plants', *Recent Advances in Fructoologosaccharides Research*, 2, pp. 47-67.

Lopez, M.G., Mancilla-Margalli, N.A. and Mendoza-Diaz, G. (2003) 'Molecular structures of fructans from Agave tequilana Weber var. azul', *Journal of Agricultural and Food Chemistry*, 51(27), pp. 7835-7840.

Lüttge, U. (1986) 'Nocturnal water storage in plants having Crassulacean acid metabolism', *Planta*, 168(2), pp. 287-289.

Lüttge, U. (1987) 'Carbon dioxide and water demand: Crassulacean acid metabolism (CAM), a versatile ecological adaptation exemplifying the need for integration in ecophysiological work', *New Phytologist*, 106(4), pp. 593-629.

Lüttge, U. (2006) 'Photosynthetic flexibility and ecophysiological plasticity: questions and lessons from *Clusia*, the only CAM tree, in the neotropics', *New Phytologist*, 171(1), pp. 7-25.

Lüttge, U. and Nobel, P.S. (1984) 'Day-night variations in malate concentration, osmotic pressure, and hydrostatic pressure in *Cereus validus*', *Plant physiology*, 75(3), pp. 804-807.

Madore, M.A. (1992) 'Nocturnal stachyose metabolism in leaf tissues of Xerosicyos danguyi H. Humb', *Planta*, 187(4), pp. 537-541.

Maeshima, M. (1992) 'Characterization of the major integral protein of vacuolar membrane', *Plant Physiology*, 98(4), pp. 1248-1254.

Maeshima, M. (2000) 'Vacuolar H< sup>+</sup>-pyrophosphatase', *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1465(1), pp. 37-51.

Maeshima, M. (2001) 'Tonoplast transporters: organization and function', *Annual review of Plant Biology*, 52(1), pp. 469-497.

Maeshima, M., Mimura, T. and Sato, T. (1994) 'Distribution of vacuolar H+-pyrophosphatase and a membrane integral protein in a variety of green plants', *Plant and cell physiology*, 35(2), pp. 323-328.

Mallick, P. and Kuster, B. (2010) 'Proteomics: a pragmatic perspective', *Nat Biotech*, 28(7), pp. 695-709.

Mancilla-Margalli, N.A. and López, M.G. (2006) 'Water-soluble carbohydrates and fructan structure patterns from Agave and Dasylirion species', *Journal of Agricultural and Food Chemistry*, 54(20), pp. 7832-7839.

Marquardt, G. and Lüttge, U. (1987) 'Proton transporting enzymes at the tonoplast of leaf cells of the CAM plant Kalanchoë daigremontiana. II. The pyrophosphatase', *Journal of Plant Physiology*, 129(3), pp. 269-286.

Martin, A., Martins, M., Mattoso, L. and Silva, O. (2009) 'Chemical and structural characterization of sisal fibers from *Agave sisalana* variety', *Polímeros*, 19(1), pp. 40–46.

MartÍNez-Torres, J., Barahona-PÉRez, F., Lappe-Oliveras, P., Colunga GarcÍA-MarÍN, P., Magdub-MÉNdez, A., Vergara-Yoisura, S. and LarquÉ-Saavedra, A. (2011) 'Ethanol production from two varieties of henequen (Agave fourcroydes Lem)', *GCB Bioenergy*, 3(1), pp. 37-42.

Martínez Salvador, M., Arias, H.R. and Rubio, A.O. (2005) 'Population structure of maguey (Agave Salmiana ssp. Crassispina) in southeast Zacatecas, México', *Arid Land Research and Management*, 19(2), pp. 101-109.

Martinoia, E. (1992) 'Transport processes in vacuoles of higher plants', *Botanica acta*, 105(4), pp. 232-245.

Martinoia, E., Kaiser, G., Schramm, M.J. and Heber, U. (1987) 'Sugar transport across the plasmalemma and the tonoplast of barley mesophyll protoplasts. Evidence for different transport systems', *Journal of plant physiology*, 131(5), pp. 467-478.

Martinoia, E., Klein, M., Geisler, M., Bovet, L., Forestier, C., Kolukisaoglu, Ü., Müller-Röber, B. and Schulz, B. (2002) 'Multifunctionality of plant ABC transporters–more than just detoxifiers', *Planta*, 214(3), pp. 345-355.

Martinoia, E., Maeshima, M. and Neuhaus, H.E. (2007) 'Vacuolar transporters and their essential role in plant metabolism', *Journal of Experimental Botany*, 58(1), pp. 83-102.

Martinoia, E., Massonneau, A. and Frangne, N. (2000) 'Transport processes of solutes across the vacuolar membrane of higher plants', *Plant and Cell Physiology*, 41(11), pp. 1175-1186.

Marys, E. and Izaguirre-Mayoral, M.L. (1995) 'Isolation and characterisation of a new Venezuelan strain of cassava common mosaic virus', *Annals of Applied Biology*, 127(1), pp. 105-112.

Matsuura-Endo, C., Maeshima, M. and Yoshida, S. (1992) 'Mechanism of the decline in vacuolar H+-ATPase activity in mung bean hypocotyls during chilling', *Plant physiology*, 100(2), pp. 718-722.

Maxwell , VonCaemmerer S and JR, E. (1997) 'Is a low internal conductance to CO_2 diffusion a consequence of succulence in plants with crassulacean acid metabolism?', *Australian Journal of Plant Physiology*, 24, pp. 777-786.

Maxwell, K., Badger, M.R. and Osmond, C.B. (1998) 'A comparison of CO_2 and O_2 exchange patterns and the relationship with chlorophyll fluorescence during photosynthesis in C3 and CAM plants', *Functional Plant Biology*, 25(1), pp. 45-52.

Maxwell, K., Griffiths, H., Helliker, B., Roberts, A., Haslam, R.P., Girnus, J., Robe, W.E. and Borland, A.M. (2002) 'Regulation of Rubisco activity in crassulacean acid metabolism plants: better late than never', *Functional Plant Biology*, 29(6), pp. 689-696.

Maxwell, K., Marrison, J.L., Leech, R.M., Griffiths, H. and Horton, P. (1999) 'Chloroplast acclimation in leaves of Guzmania monostachia in response to high light', *Plant Physiology*, 121(1), pp. 89-96.

McRae, S.R., Christopher, J.T., Smith, J.A.C. and Holtum, J.A.M. (2002) 'Sucrose transport across the vacuolar membrane of Ananas comosus', *Functional Plant Biology*, 29(6), pp. 717-724.

Mellado-Mojica, E. and López, M.G. (2012) 'Fructan Metabolism in A. tequilana Weber Blue Variety along Its Developmental Cycle in the Field', *Journal of Agricultural and Food Chemistry*, 60(47), pp. 11704-11713.

MEW (1999) 'Electrical Energy Statistical Year Book (1999), KPP, Kuwait City'.

MEW (2003) 'Statistical year book: Electrical energy. Ministry of Electricity and Water, Kuwait'.

MOP (1998) 'Annual Statistical Abstract, Statistics and Information Sector,Ministry of Planning, Kuwait', Edition 35.

Muñoz-Gutiérrez, I., Rodríguez-Alegría, M.E. and Munguía, A.L. (2009) 'Kinetic behaviour and specificity of β-fructosidases in the hydrolysis of plant and microbial fructans', *Process Biochemistry*, 44(8), pp. 891-898.

Mylsamy, K. and Rajendran, I. (2010) 'Investigation on physio-chemical and mechanical properties of raw and alkali-treated *Agave americana* fiber', *Journal of Reinforced Plastics and Composites*, 29(19), pp. 2925–2935.

Narváez-Zapata, J.A. and Sánchez-Teyer, L.F. (2010) 'Agaves as a raw material: recent technologies and applications', *Recent Patents on Biotechnology*, 3(3), pp. 185-191.

Nasrallah, H.A., Balling, R.C., Selover, N.J. and Vose, R.S. (2001) 'Development of a seasonal forecast model for Kuwait winter precipitation', *Journal of arid environments*, 48(2), pp. 233-242.

Nava-Cruz, N.Y., Medina-Morales, M.A., Martinez, J.L., Rodriguez, R. and Aguilar, C.N. (2014) 'Agave biotechnology: an overview', *Critical Reviews in Biotechnology*, (0), pp. 1-14.

Neales, T.F., Patterson, A.A. and Hartney, V.J. (1968) 'Physiological adaptation to drought in the carbon assimilation and water loss of xerophytes'.

Nelson, E.A. and Sage, R.F. (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(7), pp. 1841-1850.

Nelson, E.A., Sage, T.L. and Sage, R.F. (2005) 'Functional leaf anatomy of plants with crassulacean acid metabolism', *Functional Plant Biology*, 32(5), p. 409.

Neuhaus, H.E. (2007) 'Transport of primary metabolites across the plant vacuolar membrane', *FEBS letters*, 581(12), pp. 2223-2226.

Nimmo, G.A., Nimmo, H.G., Fewson, C.A. and Wilkins, M.B. (1984) 'Diurnal changes in the properties of phosphoenolpyruvate carboxylase in Bryophyllum leaves: a possible co valent modification', *FEBS Letters*, 178(2), pp. 199-203.

Nimmo, G.A., Nimmo, H.G., Hamilton, I.D., Fewson, C.A. and Wilkins, M.B. (1986) 'Purification of the phosphorylated night form and dephosphorylated day form of phosphoenolpyruvate carboxylase from *Bryophyllum fedtschenkoi*', *Biochem. J*, 239, pp. 213-220.

Nobel, P.S. (1985) 'Par, Water, and Temperature Limitations on the Productivity of Cultivated Agave fourcroydes (Henequen)', *Journal of Applied Ecology*, 22(1), pp. 157-173.

Nobel, P.S. (1991) 'Achievable productivities of certain CAM plants: basis for high values compared with C3 and C4 plants', *New Phytologist*, 119(2), pp. 183-205.

Nobel, P.S. (1994) *Remarkable agaves and cacti*. Oxford University Press.

Nobel, P.S. (1996) 'High productivity of certain agronomic CAM species', in *Crassulacean acid metabolism*. Springer, pp. 255-265.

Nobel, P.S. (2003) *Environmental biology of agaves and cacti*. Cambridge University Press.

Nobel, P.S. (2010) *Desert wisdom-agaves and cacti: CO2 CO [tief] 2, water, climate change*. IUniverse.

Nobel, P.S., Castañeda, M., North, G., Pimienta-Barrios, E. and Ruiz, A. (1998) 'Temperature influences on leaf CO_2 exchange, cell viability and cultivation range forAgave tequilana', *Journal of Arid Environments*, 39(1), pp. 1-9.

Nobel, P.S., GarcÍA-Moya, E. and Quero, E. (1992) 'High annual productivity of certain agaves and cacti under cultivation', *Plant, Cell & Environment*, 15(3), pp. 329-335.

Nobel, P.S. and Hartsock, T.L. (1978) 'Resistance Analysis of Nocturnal Carbon Dioxide Uptake by a Crassulacean Acid Metabolism Succulent, Agave deserti', *Plant Physiology*, 61(4), pp. 510-514.

Nobel, P.S. and Hartsock, T.L. (1986) 'Short-term and long-term responses of crassulacean acid metabolism plants to elevated CO_2', *Plant physiology*, 82(2), pp. 604-606.

Nunez, H.M., Rodriguez, L.F. and Khanna, M. (2011) 'Agave for tequila and biofuels: an economic assessment and potential opportunities', *GCB Bioenergy*, 3(1), pp. 43-57.

Olivares, E. and Medina, E. (1990) 'Carbon Dioxide Exchange, Soluble Carbohydrates and Acid Accumulation in a Fructan Accumulating Plant: Fourcroya humboldtiana Treal', *Journal of Experimental Botany*, 41(5), pp. 579-585.

Oliver, R.J., Finch, J.W. and Taylor, G. (2009) 'Second generation bioenergy crops and climate change: a review of the effects of elevated atmospheric CO_2 and drought on water use and the implications for yield', *GCB Bioenergy*, 1(2), pp. 97-114.

Olsen, J.V., de Godoy, L.M.F., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S. and Mann, M. (2005) 'Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap', *Molecular & Cellular Proteomics*, 4(12), pp. 2010-2021.

Omar, S. (2001) 'Mapping the vegetation of Kuwait through reconnaissance soil survey', *Journal of Arid Environments*, 48(3), pp. 341-355.

Orestes Guerra, J., Meneses, A., Simonet, A.M., Macías, F.A., Nogueiras, C., Gómez, A. and Escario, J.A. (2008) 'Saponinas esteroidales de la planta Agave brittoniana (Agavaceae) con actividad contra el parásito Trichomona vaginalis', *Revista de Biología Tropical*, 56(4), pp. 1645-1652.

Osmond, B., Maxwell, K., Popp, M. and Robinson, S. (1999) 'On being thick: fathoming apparently futile pathways of photosynthesis and carbohydrate metabolism in succulent CAM plants', *Carbohydrate metabolism in plants. Oxford: BIOS Scientific Publishers*, pp. 183-200.

Osmond, C.B. (1978) 'Crassulacean Acid Metabolism: A Curiosity in Context', *Annual Review of Plant Physiology*, 29(1), pp. 379-414.

Owen, N.A. and Griffiths, H. (2013) 'A system dynamics model integrating physiology and biochemical regulation predicts extent of crassulacean acid metabolism (CAM) phases', *New Phytologist*, 200(4), pp. 1116-1131.

Owen, N.A. and 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(6), pp. 687-703.

Patil, V., Tran, K.-Q. and Giselrød, H.R. (2008) 'Towards sustainable production of biofuels from microalgae', *International Journal of Molecular Sciences*, 9(7), pp. 1188-1195.

Paul, M.J., Loos, K., Stitt, M. and Ziegler, P. (1993) 'Starch-degrading enzymes during the induction of CAM in Mesembryanthemum crystallinum', *Plant, Cell & Environment*, 16(5), pp. 531-538.

Perez-Pimienta, J., Lopez-Ortega, M., Varanasi, P., Stavila, V., Cheng, G., Singh, S. and Simmons, B. (2013) 'Comparison of the impact of ionic liquid pretreatment on recalcitrance of agave bagasse and switchgrass', *Bioresource Technology*, 127, pp. 18–24.

Petersen, J.-E. (2008) 'Energy production with agricultural biomass: environmental implications and analytical challenges†', *European Review of Agricultural Economics*, 35(3), pp. 385-408.

Pimienta-Barrios, E., Robles-Murguia, C. and Nobel, P.S. (2001) 'Net CO_2 Uptake for Agave tequilana in a Warm and a Temperate Environment', *Biotropica*, 33(2), pp. 312-318.

Pimienta-Barrios, E., Zañudo-Hernández, J. and García-Galindo, J. (2006) 'Fotosíntesis estacional en plantas jóvenes de Agave tequilana', *Agrociencia*, 40(6), pp. 699-709.

Pontis, H.G. (1989) 'Fructans and cold stress', *Journal of Plant Physiology*, 134(2), pp. 148-150.

Poorter, H., Niinemets, Ü., Poorter, L., Wright, I.J. and Villar, R. (2009) 'Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis', *New Phytologist*, 182(3), pp. 565-588.

Popp, M., Janett, H.P., Lüttge, U. and Medina, E. (2003) 'Metabolite gradients and carbohydrate translocation in rosette leaves of CAM and C3 bromeliads', *New Phytologist*, 157(3), pp. 649-656.

Popp, M., Kramer, D., Lee, H., Diaz, M., Ziegler, H. and Lüttge, U. (1987) 'Crassulacean acid metabolism in tropical dicotyledonous trees of the genus *Clusia*', *Trees*, 1(4), pp. 238-247.

Powles, S.B., Chapman, K.S.R. and Osmond, C.B. (1980) 'Photoinhibition of intact attached leaves of C4 plants: dependence on CO_2 and O_2 partial pressures', *Functional Plant Biology*, 7(6), pp. 737-747.

Pucher, G.W., Vickery, H.B., Abrahams, M.D. and Leavenworth, C.S. (1949) 'Studies in the metabolism of crassulacean plants: diurnal variation of organic

acids and starch in excised leaves of Bryophyllum calycinum', *Plant physiology*, 24(4), p. 610.

Pyankov, V.I., Kondratchuk, A.V. and Shipley, B. (1999) 'Leaf structure and specific leaf mass: the alpine desert plants of the Eastern Pamirs, Tadjikistan', *New Phytologist*, 143(1), pp. 131-142.

Ragauskas, A.J., Williams, C.K., Davison, B.H., Britovsek, G., Cairney, J., Eckert, C.A., Frederick, W.J., Hallett, J.P., Leak, D.J. and Liotta, C.L. (2006) 'The path forward for biofuels and biomaterials', *science*, 311(5760), pp. 484-489.

Ramírez-Malagón, R., Borodanenko, A., Pérez-Moreno, L., Salas-Araiza, M.D., Nuñez-Palenius, H.G. and Ochoa-Alejo, N. (2008) 'In vitro propagation of three Agave species used for liquor distillation and three for landscape', *Plant Cell, Tissue and Organ Culture*, 94(2), pp. 201-207.

Raveh, E., Wang, N. and Nobel, P.S. (1998) 'Gas exchange and metabolite fluctuations in green and yellow bands of variegated leaves of the monocotyledonous CAM species Agave americana', *Physiologia Plantarum*, 103(1), pp. 99-106.

Ravi, S., Lobell, D.B. and Field, C.B. (2012) *AGU Fall Meeting Abstracts*.
Rea, P.A. and Sanders, D. (1987) 'Tonoplast energization: two H+ pumps, one membrane', *Physiologia Plantarum*, 71(1), pp. 131-141.

Rees, T.A. (1994) 'Plant physiology: virtue on both sides', *Current Biology*, 4(6), pp. 557-559.

Reich, P.B. (1993) 'Reconciling Apparent Discrepancies Among Studies Relating Life Span, Structure and Function of Leaves in Contrasting Plant Life Forms and Climates:The Blind Men and the Elephant Retold'', *Functional Ecology*, pp. 721-725.

Ritsema, T. and Smeekens, S. (2003) 'Fructans: beneficial for plants and humans', *Current opinion in plant biology*, 6(3), pp. 223-230.

Ritz, D., Kluge, M. and Veith, H.J. (1986) 'Mass-spectrometric evidence for the double-carboxylation pathway of malate synthesis by crassulacean acid metabolism plants in light', *Planta*, 167(2), pp. 284-291.

Robert, M.L., Herrera-Herrera, J.L., Castillo, E., Ojeda, G. and Herrera-Alamillo, M.A. (2006) 'An efficient method for the micropropagation of agave species', in *Plant Cell Culture Protocols*. Springer, pp. 165-178.

Roberts, A., Borland, A.M. and Griffiths, H. (1997) 'Discrimination processes and shifts in carboxylation during the phases of crassulacean acid metabolism', *Plant Physiology*, 113(4), pp. 1283-1292.

Roepstorff, P. and Fohlman, J. (1984) 'Proposal for a common nomenclature for sequence ions in mass spectra of peptides', *Biomed Mass Spectrom*, 11(11), p. 601.

Rosegrant, M.W. (2008) *Biofuels and grain prices: impacts and policy responses*. International Food Policy Research Institute Washington, DC.

Roy, W. and Grealish, G. (2004) 'Mapping arable soils using GIS-based soil information database in Kuwait', *Management of Environmental Quality: An International Journal*, 15(3), pp. 229-237.

Sage, R.F. (2002) 'Are crassulacean acid metabolism and C4 photosynthesis incompatible?', *Functional Plant Biology*, 29(6), pp. 775-785.

Sanchez, A.E.M. (2009) 'Improved Agave Cultivars (*Agave angustifolia Haw*) for Profitable and Sustainable Bioethanol Production in Mexico'.

Sarafian, V. and Poole, R.J. (1989) 'Purification of an H+-translocating inorganic pyrophosphatase from vacuole membranes of red beet', *Plant Physiology*, 91(1), pp. 34-38.

Sayed, O.H. (1998) 'Phenomorphology and ecophysiology of desert succulents in eastern Arabia', *Journal of Arid Environments*, 40(2), pp. 177-189.

Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) 'Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels', *Analytical Chemistry*, 68(5), pp. 850-858.

Shimaoka, T., Ohnishi, M., Sazuka, T., Mitsuhashi, N., Hara-Nishimura, I., Shimazaki, K.-I., Maeshima, M., Yokota, A., Tomizawa, K.-I. and Mimura, T. (2004) 'Isolation of intact vacuoles and proteomic analysis of tonoplast from suspension-cultured cells of Arabidopsis thaliana', *Plant and Cell Physiology*, 45(6), pp. 672-683.

Shipley, B. (1995) 'Structured interspecific determinants of specific leaf area in 34 species of herbaceous angiosperms', *Functional Ecology*, pp. 312-319.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M. and Söding, J. (2011) 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Molecular Systems Biology*, 7(1), p. 539.

Silvera, K., Neubig, K.M., Whitten, W.M., Williams, N.H., Winter, K. and Cushman, J.C. (2010) 'Evolution along the crassulacean acid metabolism continuum', *Functional Plant Biology*, 37(11), pp. 995-1010.

Simpson, J., Martinez Hernandez, A., JazmÍN Abraham JuÁRez, M., Delgado Sandoval, S., Sanchez Villarreal, A. and Cortes Romero, C. (2011a) 'Genomic resources and transcriptome mining in *Agave tequilana'*, *GCB Bioenergy*, 3(1), pp. 25-36.

Smith, A.M. (2008) 'Prospects for increasing starch and sucrose yields for bioethanol production', *The Plant Journal*, 54(4), pp. 546-558.
Smith, J.A.C. and Bryce, J.H. (1992) *Seminar series*.

Smith, J.A.C. and Heuer, S. (1981) 'Determination of the volume of intercellular spaces in leaves and some values for CAM plants', *Annals of Botany*, pp. 915-917.

Smith, J.A.C., Ingram, J., Tsiantis, M.S., Barkla, B.J., Bartholomew, D.M., Bettey, M., Pantoja, O. and Pennington, A.J. (1996) 'Transport across the vacuolar membrane in CAM plants', in *Crassulacean Acid Metabolism*. Springer, pp. 53-71.

Smith, J.A.C., Schulte, P.J. and Nobel, P.S. (1987) 'Water flow and water storage in Agave deserti: osmotic implications of crassulacean acid metabolism', *Plant, Cell & Environment*, 10(8), pp. 639-648.

Smith, J.A.C., Uribe, E.G., Ball, E., Heuer, S. and LÜTtge, U. (1984) 'Characterization of the vacuolar ATPase activity of the crassulacean-acid-metabolism plant Kalanchoë daigremontiana Receptor modulating', *European journal of biochemistry*, 141(2), pp. 415-420.

Somerville, C., Youngs, H., Taylor, C., Davis, S.C. and Long, S.P. (2010) 'Feedstocks for lignocellulosic biofuels', *Science(Washington)*, 329(5993), pp. 790-792.

Spalding MH, S.D., Ku MSB, Burris RH, Edwards GE (1979) 'Crassulacean Acid Metabolism and Diurnal CO2 concentrations in *Sedum praelaltum* DC. ', *Australian Journal of Plant Physiology*, 6, pp. 557-567.

Steen, H. and Mann, M. (2004) 'The ABC's (and XYZ's) of peptide sequencing', *Nature reviews Molecular cell biology*, 5(9), pp. 699-711.

Suleiman, M.K. and Abdal, M.S. (2002) 'Water availability for the greening of Kuwait', *Limnologica-Ecology and Management of Inland Waters*, 32(4), pp. 322-328.

Sutton, B.G. (1975) 'Glycolysis in CAM plants', *Functional Plant Biology*, 2(3), pp. 389-402.

Szarek, S.R. and Ting, I.P. (1975) 'Photosynthetic Efficiency of CAM Plants in Relation to C3 and C4 Plants', in Marcelle, R. (ed.) *Environmental and Biological Control of Photosynthesis*. Springer Netherlands, pp. 289-297.

Takasu, A., Nakanishi, Y., Yamauchi, T. and Maeshima, M. (1997) 'Analysis of the substrate binding site and carboxyl terminal region of vacuolar H+-pyrophosphatase of mung bean with peptide antibodies', *Journal of biochemistry*, 122(4), pp. 883-889.

Taybi, T., Patil, S., Chollet, R. and Cushman, J.C. (2000) 'A minimal serine/threonine protein kinase circadianly regulates phosphoenolpyruvate carboxylase activity in crassulacean acid metabolism-induced leaves of the common ice plant', *Plant Physiology*, 123(4), pp. 1471-1482.

Teeri, J.A., Tonsor, S.J. and Turner, M. (1981) 'Leaf thickness and carbon isotope composition in the Crassulaceae', *Oecologia*, 50(3), pp. 367-369.

Terrier, N., Deguilloux, C., Sauvage, F.X., Martinoia, E. and Romieu, C. (1997) 'V-ATPase, inorganic pyrophosphatase and and anion transport on the tonoplast grape berries (Vitis viniferaL.)', *Plant Physiol Biochem*, 36, pp. 79-193.
Valenzuela-Zapata, A.G. (1985) 'The tequila industry in Jalisco, Mexico', *Desert plants (USA)*.

Valenzuela, A. (2011) 'A new agenda for blue agave landraces: food, energy and tequila', *Gcb Bioenergy*, 3(1), pp. 15-24.

Valluru, R. and Van den Ende, W. (2008) 'Plant fructans in stress environments: emerging concepts and future prospects', *Journal of Experimental Botany*, 59(11), pp. 2905-2916.

Vendramini, F., Díaz, S., Gurvich, D.E., Wilson, P.J., Thompson, K. and Hodgson, J.G. (2002) 'Leaf traits as indicators of resource-use strategy in floras with succulent species', *New Phytologist*, 154(1), pp. 147-157.

Verástegui, Á., Verde, J., García, S., Heredia, N., Oranday, A. and Rivas, C. (2008) 'Species of Agave with antimicrobial activity against selected pathogenic bacteria and fungi', *World Journal of Microbiology and Biotechnology*, 24(7), pp. 1249-1252.

Vieira, M., Heinze, T., Antonio-Cruz, R. and Mendoza-Martinez, A. (2002) 'Cellulose derivatives from cellulosic material isolated from *Agave lechuguilla* and *fourcroydes*', *Cellulose*, 9(2), pp. 203–212.

Vijn, I. and Smeekens, S. (1999) 'Fructan: more than a reserve carbohydrate?', *Plant physiology*, 120(2), pp. 351-360.

Vile, D., Garnier, E., Shipley, B., Laurent, G., Navas, M.-L., Roumet, C., Lavorel, S., Diaz, S., Hodgson, J.G. and Lloret, F. (2005) 'Specific leaf area and dry matter content estimate thickness in laminar leaves', *Annals of botany*, 96(6), pp. 1129-1136.

Von Caemmerer, S.v. and Farquhar, G.D. (1981) 'Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves', *Planta*, 153(4), pp. 376-387.

Wang, D., Li, X.F., Zhou, Z.J., Feng, X.P., Yang, W.J. and Jiang, D.A. (2010) 'Two Rubisco activase isoforms may play different roles in photosynthetic heat acclimation in the rice plant', *Physiologia plantarum*, 139(1), pp. 55-67.

Wang, M., Wu, M. and Huo, H. (2007) 'Life-cycle energy and greenhouse gas emission impacts of different corn ethanol plant types', *Environmental Research Letters*, 2(2), p. 024001.

Wang, N. and Nobel, P.S. (1998) 'Phloem transport of fructans in the crassulacean acid metabolism species Agave deserti', *Plant physiology*, 116(2), pp. 709-714.

Wang, Y. and Sze, H. (1985) 'Similarities and differences between the tonoplast-type and the mitochondrial H+-ATPases of oat roots', *Journal of Biological Chemistry*, 260(19), pp. 10434-10443.

Watson, R.T., Zinyowera, M.C. and Moss, R.H. (1996) *Climate Change 1995 impacts, adaptations and mitigation of climate change: Scientific-technical analysis*. Cambridge University Press.

Westbrook, J.W., Kitajima, K., Burleigh, J.G., Kress, W.J., Erickson, D.L. and Wright, S.J. (2011) 'What makes a leaf tough? Patterns of correlated evolution between leaf toughness traits and demographic rates among 197 shade-tolerant woody species in a Neotropical forest', *The American Naturalist*, 177(6), pp. 800-811.

White, P.J. and Smith, J.A.C. (1989) 'Proton and anion transport at the tonoplast in crassulacean-acid-metabolism plants: specificity of the malate-influx system in Kalanchoë daigremontiana', *Planta*, 179(2), pp. 265-274.

Williams, L.E., Lemoine, R. and Sauer, N. (2000) 'Sugar transporters in higher plants–a diversity of roles and complex regulation', *Trends in plant science*, 5(7), pp. 283-290.

Wilson, P.J., Thompson, K.E.N. and Hodgson, J.G. (1999) 'Specific leaf area and leaf dry matter content as alternative predictors of plant strategies', *New phytologist*, 143(1), pp. 155-162.

Wingenter, K., Schulz, A., Wormit, A., Wic, S., Trentmann, O., Hoermiller, I.I., Heyer, A.G., Marten, I., Hedrich, R. and Neuhaus, H.E. (2010) 'Increased activity of the vacuolar monosaccharide transporter TMT1 alters cellular sugar partitioning, sugar signaling, and seed yield in Arabidopsis', *Plant physiology*, 154(2), pp. 665-677.

Winter, H., Robinson, D.G. and Heldt, H.W. (1993) 'Subcellular volumes and metabolite concentrations in barley leaves', *Planta*, 191(2), pp. 180-190.

Winter, K. (1982) 'Properties of phosphoenolpyruvate carboxylase in rapidly prepared, desalted leaf extracts of the Crassulacean acid metabolism plant Mesembryanthemum crystallinum L', *Planta*, 154(4), pp. 298-308.

Winter, K. (1985) ' Crassulacean acid metabolism. Photosynthetic mechanisms and the environment. J. Barber & N. R. Baker (eds.)

', *Elsevier, London*, pp. 329-387

Winter, K., Foster, J., Schmitt, M. and Edwards, G. (1982) 'Activity and quantity of ribulose bisphosphate carboxylase-and phosphoenolpyruvate carboxylase-protein in two Crassulacean acid metabolism plants in relation to leaf age, nitrogen nutrition, and point in time during a day/night cycle', *Planta*, 154(4), pp. 309-317.

Winter, K., Garcia, M. and Holtum, J.A.M. (2008) 'On the nature of facultative and constitutive CAM: environmental and developmental control of CAM expression during early growth of Clusia, Kalanchoë, and Opuntia', *Journal of Experimental Botany*, 59(7), pp. 1829-1840.

Winter, K., Medina, E., Garcia, V., Mayoral, M.L. and Muniz, R. (1985) 'Crassulacean Acid Metabolism in Roots of a Leafless Orchid, Campylocentrum tyrridion Garay & Dunsterv', *Journal of plant physiology*, 118(1), pp. 73-78.

Winter, K. and Smith, J.A.C. (1996) 'An Introduction to Crassulacean Acid Metabolism. Biochemical Principles and Ecological Diversity', in Winter, K. and Smith, J.A. (eds.) *Crassulacean Acid Metabolism*. Springer Berlin Heidelberg, pp. 1-13.

Winter, K., Wallace BJ, Stocker GC and Z, R. (1983) 'Crassulacean acid metabolism in Australian vascular epiphytes and some related species', *Oecologia* 57, pp. 129-141.

Witkowski, E.T.F. and Lamont, B.B. (1991) 'Leaf specific mass confounds leaf density and thickness', *Oecologia*, 88(4), pp. 486-493.

Woodhouse, R.M., Williams, J.G. and Nobel, P.S. (1980) 'Leaf orientation, radiation interception, and nocturnal acidity increases by the CAM plant Agave deserti (Agavaceae)', *American Journal of Botany*, pp. 1179-1185.

Wormit, A., Trentmann, O., Feifer, I., Lohr, C., Tjaden, J., Meyer, S., Schmidt, U., Martinoia, E. and Neuhaus, H.E. (2006) 'Molecular identification and physiological characterization of a novel monosaccharide transporter from Arabidopsis involved in vacuolar sugar transport', *The Plant Cell Online*, 18(12), pp. 3476-3490.

Yakir, D., Ting, I. and DeNiro, M. (1994) 'Natural Abundance< sup> 2</sup> H/< sup> 1</sup> H Ratios of Water Storage in Leaves of Peperomia Congesta HBK during Water Stress', *Journal of plant physiology*, 144(4), pp. 607-612.

Yan, X., Tan, D.K.Y., Inderwildi, O.R., Smith, J.A.C. and King, D.A. (2011) 'Life cycle energy and greenhouse gas analysis for agave-derived bioethanol', *Energy & Environmental Science*, 4(9), pp. 3110-3121.

Yang, L., Carl, S., Lu, M., Mayer, J., Cushman, J., Tian, E. and Lin, H. (2015a) 'Biomass characterization of *Agave* and *Opuntia* as potential biofuel feedstocks', *Biomass and Bioenergy*, in revision.

Yang, X., Cushman, J.C., Borland, A.M., Edwards, E.J., Wullschleger, S.D., Tuskan, G.A., Owen, N.A., Griffiths, H., Smith, J.A.C. and De Paoli, H.C. (2015b) 'A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in hotter, drier world', *New Phytologist*.

Zapata, A.G.V. and Nabhan, G.P. (2003) *Tequila: a natural and cultural history*. University of Arizona press.

Zhang, N. and Portis, A.R. (1999) 'Mechanism of light regulation of Rubisco: a specific role for the larger Rubisco activase isoform involving reductive activation by thioredoxin-f', *Proceedings of the National Academy of Sciences*, 96(16), pp. 9438-9443.

Zhu, J., Bartholomew, D.P. and Goldstein, G. (1997) 'Effect of elevated carbon dioxide on the growth and physiological responses of pineapple, a species with crassulacean acid metabolism', *Journal of the American Society for Horticultural Science*, 122(2), pp. 233-237.