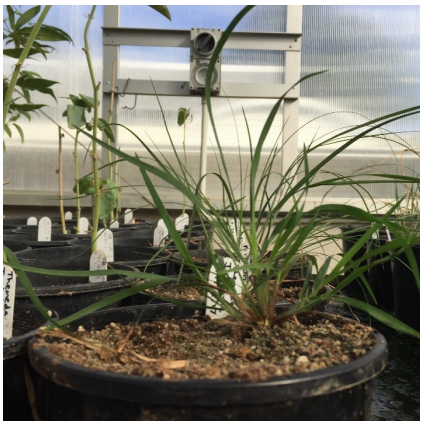
Surveying the diversity of Rubisco activase genes in Australian wild crop relatives

Aaron L. Phillips

BSc (Macquarie University)


Supervisors:

Professor Brian Atwell

Dr. Rachael Gallagher

Department of Biological Sciences,

Faculty of Science and Engineering

Macquarie University, Sydney NSW Australia

This thesis was submitted in partial fulfilment of the requirements for the degree of Masters of Research

Submission date: 30 October 2017

Declaration

I certify that the material of this thesis has not been previously submitted as part of the requirements for a higher degree to any other university or institution.

This thesis contains no material previously published or written by any other person. I certify that all information sources and literature used are indicated in the thesis.

I wish to acknowledge the following assistance with the research detailed in this report:

- Belinda Fabian, Timothy Maher, Katie Berthon, and Anthony Manea assisted in the collection of physiological data;
- Dr. Steve Van Sluyter provided all chemicals for protein extraction, tutored me in the protein extraction protocol detailed in this thesis, ran initial samples on a mass spectrometer at the University of Sydney, and assisted in the curation of mass spectrometry data and its analysis;
- Dr Renee Marchin Prokopavicius determined the concentration of protein in my protein extracts; and
- Mass spectrometry was carried out by the Australian Proteome Analysis Facility supported under the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS)
- Professor Brian Atwell and Dr. Rachael Gallagher, my supervisors, assisted in method development and edited my drafts.

All other research described in this report is my own original work.

Aaron Phillips

Note to examiners

This thesis is written in the form of a journal article from the Journal of Experimental Botany. The majority of the author guidelines have been followed, except for minor deviations detailed here and where the guidelines clash with Macquarie University thesis formatting requirements. All figures and tables have been presented at the appropriate places in the text to enhance readability. A 'Highlight' section was not included due to space limitations. A running title was not included to reduce page clutter.

Acknowledgments

First and foremost, I would like to acknowledge my mentors of the past year.

Professor Brian Atwell, my official supervisor, was a constant source of knowledge and inspiration throughout my Masters journey. He supported me when support was needed, but wasn't afraid to let me struggle and learn and grow – hindsight teaches me that this is one of the most valuable things Brian could have done for me.

Dr. Rachael Gallagher, my co-supervisor, was many times the voice of reason. Her advice, support and words of kindness instilled me with the confidence that helped me cross the finish line and I will carry that with me from here *ad infinitum*.

Julian May taught me that it's ok not to know what to do in the lab environment but also provided me with some of the tools I needed to grow comfortable as a researcher.

Steve Van Sluyter was incredibly generous in his material contributions, the time he allocated to my tutelage, and in the emotional support he offered during times of acute stress.

Thanks go to the following people for their help: **Ian Chivers**, **Richard Jobson**, and **Dr. Louise Gilfedder** provided *Themeda* seed stock; **Michelle Demers** provided help with primer design; **Paul Worden**, **Drew Allen** and **Fonti Kar** offered help with R coding and data analysis; **Mohammed Masood** provided help in glasshouse maintenance; and **Anthony Manea**, **Katie Berthon**, and **Tim Maher** helped with physiology measurements. Funding from **Macquarie University**, **The Linnean Society of NSW**, and the **Australian Plants Society (North Shore Group)** made this work possible.

To the Sheehan household, particularly to my wonderful **Joey** and the stern **Angus Sheehan**, I owe massive thanks! My thanks also extend to my good friends at the **gaming shed**, and to the members of **The Sandstorms** futsal team, particularly to **Jarrod Kelly**. You all provided a place of refuge both during times of stress and during times of intense focus. Likewise, the huge support given to me by my family, particularly from **Mum** and **Dad**, made everything I did this year achievable.

Lastly, but certainly not least, I extend the sincerest gratitude to my scientific rock, lab buddy, and moral guide, **Belinda Fabian**. Without her friendship, I would not have reached the place I am at today. I am truly humbled to have shared my experience with her.

Table of Contents

General Abstract

Rubisco activase (RCA) is the regulatory partner of Rubisco – the enzyme that fixes CO_2 into sugars. RCA removes inhibitory compounds from the active sites of Rubisco. Thus, RCA can reactivate Rubisco, increasing the efficiency of biomass accumulation in plants. Under heat stress the association between RCA and Rubisco breaks down and plant productivity declines. However, wild crop relatives that occur in hot environments have divergent RCA isoforms that can function during heat stress. Because average daily temperatures are increasing, RCA has garnered interest for transgenic studies.

This study aimed to characterise the heat response of photosynthesis and RCA abundance, and the sequence diversity of RCA in domestic and wild species of cotton (*Gossypium*) and tobacco (*Nicotiana*) and three populations of kangaroo grass (*Themeda triandra*). I found that wild species' photosynthetic rates demonstrate a range of responses to increased temperature. I identified four isoforms of RCA in both *Gossypium* and *Nicotiana* species, and one isoform in *Themeda* as well as three RCA mutant isoforms in *Gossypium* and one in *Nicotiana*. This study showed that while the total RCA content of each of these species did not change in response to heat, thermotolerant species of both genera demonstrated unique patterns of RCA accumulation, indicating the importance of RCA stoichiometry in thermotolerance. Finally, I report the occurrence of polymorphisms in RCA from *Gossypium*, *Nicotiana*, and *T. triandra* which may contribute to changes in RCA rigidity but which do not fit with current ideas about RCA thermostability.

1. Introduction

1.1 Context for Study

Climate models predict that average global temperatures will rise inexorably throughout the 21st century, along with longer, more intense and more frequent heat waves (IPCC, 2014a). Estimates of the magnitude of these changes varies widely, depending on the model being considered (Phillips *et al.*, 2014; Cowtan *et al.*, 2015). The common consensus suggests an increase in daily temperatures of at least 3^{o}C by 2100 (IPCC, 2014b). As average temperatures increase beyond the narrow tolerance range of plants, photosynthetic efficiency will fall, likely resulting in reduced yield in food and fibre crops (Hatfield *et al.*, 2011; Bita and Gerats, 2013; Kaushal *et al.*, 2016). This will jeopardise the resources required for nine billion people by 2050. It has therefore been recommended that the yield of food and fibre crops should increase to match corresponding demand increases of 100-110% in 2050 (Tilman *et al.*, 2011). Accordingly, plant biologists are looking increasingly towards the use of biotechnology to improve yield potential (Bortesi and Fischer, 2014; Hüner *et al.*, 2014; Lozano-Juste and Cutler, 2014; Palmgren *et al.*, 2015), with an emphasis on engineering photosynthesis; biotechnology promises to provide the next Green Revolution (Evans, 2013; Maurino and Weber, 2013; Nölke *et al.*, 2014; Ort *et al.*, 2015; Hüner *et al.*, 2016; Kromdijk and Long, 2016).

A key aspect of this work is the identification of genes that code for positive phenotypes (e.g. thermotolerant vegetation and increased yield). The search for novel genetic material largely centres around the investigation of crop landraces – populations of the same species that have evolved in specific environmental conditions (Mercer and Perales, 2010; Lopes *et al.*, 2015). This approach has led to the discovery of genes that contribute to agriculturally desirable traits, demonstrating the usefulness of diverse germplasm (Morrell *et al.*, 2011; Wei *et al.*, 2013). However, little attention has been given to photosynthesis-related genes in wild crop relatives. There is now growing interest in the diversity of genes and gene products from wild relatives within the same genus (Fitzgerald and Shapter, 2011; Vincent *et al.*, 2013; Atwell *et al.*, 2014; Warschefsky *et al.*, 2014).

1.2 Effects of Heat on Crop Plants

1.2.1 Yield Reduction

As ambient temperature increases beyond the optimal range for primary metabolism, a number of whole-plant and cell-level physiological and biochemical changes occur that impair plant performance (Hasanuzzaman *et al.*, 2013). The magnitude of heat-stress effects on plants is species dependent, though within a genotype the responses vary based on developmental stage and tissue type (Luo, 2011; Bita and Gerats, 2013; Jing *et al.*, 2016; Wang *et al.*, 2016). For example, plants in the reproductive growth phase (i.e. during flower development or pollen production; Bishop *et al.*, 2016) are generally more susceptible to fluctuations in temperature than plants exhibiting vegetative growth (Luo, 2011).

It has been suggested that just a 1^{o}C increase beyond the optimal temperature will impose heat stress and impair plant performance (Hasanuzzaman *et al.*, 2013). For example, increases in global temperature have been attributed to declining yields of wheat, with a predicted 6% decline in yield for every 1^{o}C increase beyond the global mean temperature (Asseng *et al.*, 2014). Additionally, grain weight in wheat has been predicted to fall by 2.5% for each 1^{o}C increase in average temperature, and by 17% following heat waves (Luo, 2011). In cotton, boll size has been shown to be reduced by up to 99.4% under high-temperature stress (40^{o}C) relative to optimal temperature (Reddy *et al.*, 1992).

Patterns of vegetative growth (those that drive the development of structural and functional plant tissues) are altered by exposure to heat stress (Hasanuzzaman *et al.*, 2013; Jing *et al.*, 2016). Species demonstrating sensitivity to high temperatures experience changes in their gross morphology, have stunted growth, and often exhibit changes in development (Hasanuzzaman *et al.*, 2013; Jing *et al.*, 2016). This is consistent with the observations of studies in domestic and wild species of rice, which report increased allocation of carbon resources to photosynthetic tissue under heat stress in tolerant wild species, while above-ground growth of sensitive domestic species was reduced (Scafaro *et al.*, 2010; Scafaro *et al.*, 2016). Similarly, Reddy *et al.* (1993) report significant change to the expansion rates and final areas of cotton leaves exposed to high temperature stress.

1.2.2 Molecular Aspects of Yield Reduction

Despite the variability of heat stress effects, there is a core set of processes that are particularly prone to thermally-induced damage across all species. Direct effects of heat on plants include increased membrane fluidity and protein aggregation and denaturation, while indirect effects include changes in inactivation of enzymes, increased production of reactive oxygen species (ROS), inhibition of synthesis pathways and acceleration of senescence (Kaushal *et al.*, 2016). These injurious effects result from fundamental impacts on metabolism that are common to all plants (Bita and Gerats, 2013; Kaushal *et al.*, 2016). In particular, heat-induced changes in membrane and protein dynamics are major determinants of plant thermotolerance because of their intimate links with transport and energy generation (Cavanagh and Kubien, 2014; Kaushal *et al.*, 2016). In the case of photosynthesis, damage by heat stress is believed to relate to impaired electron transport and compromised activity of the Calvin-Benson cycle enzymes (Sharkey, 2005; Carmo-Silva *et al.*, 2015).

1.3 Is Photosynthesis Especially Vulnerable to Heat Stress?

Losses in yield, such as those discussed in Section 1.2.1, have been linked to decreasing photosynthetic efficiency at high temperature (Bita and Gerats, 2013). Because photosynthesis is a two-stage system, a great deal of attention has been paid to relative impact on the light reactions versus the Calvin-Benson cycle. There is still no consensus over which process limits plant productivity under heat stress. Historically, the light reactions were thought to be the main limitation, though a relatively recent focus on carbon-fixation research has now challenged this view (Salvucci and Crafts-Brandner, 2004b; Scafaro *et al.*, 2016). Because the light and dark reactions are co-dependent, it has been challenging to resolve the principal lesion under heat. Progress in these questions under heat stress has been made through the use of A-C_i curves and C_{trans} values (Farquhar *et al.*, 1980; von Caemmerer and Farquhar, 1981; Long and Bernacchi, 2003; Scafaro *et al.*, 2012). While the rate-limiting reaction of photosynthesis at high temperature seems to be species-dependent (Scafaro *et al.*, 2012), it has been shown that Rubisco is deactivated at temperatures well below the thermal tolerance ranges of PSII complexes, highlighting the sensitivity of the Calvin-Benson Cycle to heat stress (Sharkey and Zhang, 2010; Ashraf and Harris, 2013).

1.3.1 Effects of Heat on Carbon Fixation

The Calvin-Benson cycle limits carbon gain under heat stress in species such as rice, cotton, tobacco and maize (Law and Crafts-Brandner, 1999; Sharkey *et al.*, 2001; Crafts-Brandner and Salvucci, 2002; Sharkey, 2005; Carmo-Silva *et al.*, 2012; Scafaro *et al.*, 2012; Scafaro *et al.*, 2016). Rubisco, the enzyme that uses the energy generated in the light reactions to fix CO_2 into biologically available molecules (i.e. sugar), is inefficient even under benign conditions (Parry *et al.*, 2006; Carmo-Silva *et al.*, 2015; Ort *et al.*, 2015). Photorespiration is exacerbated under heat (Sicher, 2015), further decreasing the efficiency of Rubisco and lowering plant productivity. In addition, Rubisco loses efficiency at high temperatures because the active sites of the enzyme become blocked and carbon fixation declines (Salvucci and Crafts-Brandner, 2004a; Salvucci and Crafts-Brandner, 2004b; Salvucci and Crafts-Brandner, 2004; Carmo-Silva *et al.*, 2012).

1.4 Improving the Heat Tolerance of Photosynthesis

1.4.1 Rubisco is Encoded by a Highly Conserved Gene

The Rubisco holoenzyme is a hexadecameric complex made up of eight large and eight small subunits (Tabita *et al.*, 2008). The large, catalytically active subunits of Rubisco are highly conserved proteins with little sequence variation which suggests some kind of evolutionary equilibrium (Parry *et al.*, 2003; Studer *et al.*, 2014). This is in spite of the fact that the molecular design of Rubisco still allows for substantial oxygenase activity and therefore futile fixation of O_2 (Maurino and Peterhansel, 2010). Nonetheless, there is great interest in understanding how the kinetics of Rubisco can be improved (Carmo-Silva *et al.*, 2015). Studies of the variation and catalytic efficiency of Rubisco between domestic and wild species of wheat show little sequence variation, but significant changes in function (Orr *et al.*, 2016; Prins *et al.*, 2016). The role of the small subunits in moderating such variation in Rubisco catalysis has garnered interest in recent years (Ishikawa *et al.*, 2011; Morita *et al.*, 2014). Sites for modification in Rubisco subunits have thus been identified as potential targets for improvement. However, with limited sequence variation it seems unlikely that dramatic improvements in Rubisco structure and function will be realised, even through the use of genetic modification. It is likely that any change in Rubisco structure is a deviation from an evolutionarily stable protein and will result in loss of function (Parry *et al.*, 2006; Tcherkez and Farquhar, 2006; Studer *et al.*, 2014). For example, the Rubisco-engineering literature commonly refers to a phenomenon in which increased Rubisco function often comes with reduced substrate specificity, and therefore higher rates of photorespiration

(Tcherkez and Farquhar, 2006; Morita *et al.*, 2014). Additionally, there are reports of faulty Rubisco assembly when utilising manipulated subunits (Gutteridge and Gatenby, 1995), though these challenges have been overcome in tobacco (Whitney *et al.*, 2009).

However, accessory proteins of the carbon-fixing reactions have been identified that show great promise for improving Rubisco indirectly (Parry *et al.*, 2013; Hauser *et al.*, 2015). For example, Rubisco activase (RCA) is recognised as a route towards the generation of a more thermostable photosynthesis (Parry *et al.*, 2013; Mueller-Cajar *et al.*, 2014; Carmo-Silva *et al.*, 2015).

1.4.2 Rubisco Activase

Present in all plants, RCA is a member of the AAA+ protein family – it hydrolyses ATP to remove inhibitory sugar-phosphate groups from the active sites of Rubisco, increasing its efficiency (Fig. 1; Portis, 2003; Wachter and Henderson, 2015). Because RCA and Rubisco are closely associated, RCA possesses conserved sequence regions that are related to its interaction with Rubisco and its catalytic function. However, the overall sequence diversity of RCA is greater than that of Rubisco (Mueller-Cajar *et al.*, 2014). It may be that because RCA is only an accessory to the Calvin-Benson cycle, it is more likely to undergo evolutionary change than Rubisco itself (Alvarez-Ponce *et al.*, 2016; Hsu *et al.*, 2016). Overall, the sequence diversity in RCA, coupled with its presence in all plants, and its ability to activate Rubisco means that RCA is a good candidate for the improvement of photosynthesis (Parry *et al.*, 2013). The presence of conserved residues indicates that RCA genes might be transferrable within a genus with no loss of function, while their diversity allows for differences in efficiency to arise in populations from different habitats.

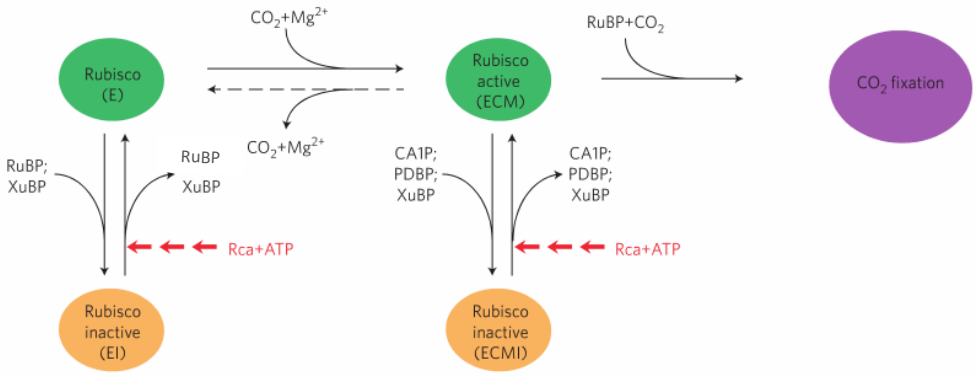


Figure 1. Mechanism of RCA activity on Rubisco. RCA hydrolyses ATP to return inactive Rubisco (orange) to a functional state (green), facilitating CO_2 fixation. Adapted from Hauser *et al.* (2015).

At high temperature, the RCA isoforms of many plants lose the ability to maintain Rubisco in an active state and photosynthesis is compromised (Crafts-Brandner and Salvucci, 2000). Therefore, it has been speculated that improved thermal stability of RCA would allow it to maintain its interaction with Rubisco and will lead to improved thermotolerance in C3 plants at high temperatures (Parry *et al.*, 2013). This has been demonstrated in *Arabidopsis thaliana* plants with chimeric *Arabidopsis*-tobacco RCA, where the fusion product increased photosynthetic rate under heat stress (Kumar *et al.*, 2009), indicating some level of cross-species compatibility, though the extent of this cross compatibility has not been established. For example, alternative forms of RCA are expected to have less impact on photosynthetic efficiency in C4 plants, with other factors playing a larger role in thermotolerance (Hendrickson *et al.*, 2007).

1.5 Sequence Polymorphisms as a Path to Thermotolerance

1.5.1 Local Adaptation to Climates

Terrestrial plant species are limited in their distribution by their ability to respond to environmental conditions (Criddle *et al.*, 1994). Plants are characterised by plasticity in their response to stress and by increased levels of genetic variation driven by whole or tandem genome duplication events. Whole genome duplications (WGDs) have garnered much support as drivers of plant evolution, especially in the angiosperm lineage (Zhu *et al.*, 2016). Seminal theories regarding the function of WGDs in organismal evolution posit that the resulting duplicated chromosomes provide immense sources of raw genetic material that can be acted upon by selective forces (Ohno, 1970; McGrath and Lynch, 2012). The functional redundancy of duplicated genes means their function can drift away from their ancestral state with no fitness penalties for the organism (Ohno, 1970). Depending on the function of the gene in the context of the whole genome, the duplicated gene can be maintained as it is, be subject to purifying selection and be lost, or experience positive selection and undergo functional divergence. While this is especially true for regulatory genes like heat shock factors (Scharf *et al.*, 2012; McGrath *et al.*, 2014), functional redundancy also provides scope for variation to accumulate in genes that contribute more directly to stress tolerance, even within a species where populations are geographically isolated. In such a situation, local adaptation to environment can be said to have occurred.

There is substantial evidence for local adaption to temperature and precipitation regimes for long-lived tree species that occur across broad latitudinal gradients. For example, Kaluthota *et al.*

(2015) found significant differences in the photosynthetic capacity of southern, central and northern populations of *Populus angustifolia* consistent with previous measures of genetic variation between the populations. Similarly, Benomar *et al.* (2016) demonstrated that rates of photosynthesis of individuals from six populations of white spruce (*Picea glauca*) exhibited signs of adaptive response to the temperature gradient they occurred along, and that these differences were likely driven by local genetic adaption. While no explicit link was made by either study to variation in the activity of RCA on Rubisco, it has previously been suggested that species distributions might be limited by the thermal tolerance range of their respective activases (Salvucci and Crafts-Brandner, 2004b). The activity of RCA is related to the environment in which a species occurs, which is driven either by the accumulation of beneficial mutations, changes in the abundance of isoforms, or by some kind of post-translational modification of the enzyme (Jiang *et al.*, 2013; Scafaro *et al.*, 2016).

1.5.2 Rubisco activase – a Case Study: Heat-adapted Species Have Polymorphic RCA

It has been demonstrated that heat-adapted plants possess RCA variants that are more resistant to high temperature stress than those found in cooler climates (Salvucci and Crafts-Brandner, 2004b; Carmo-Silva *et al.*, 2012; Scafaro *et al.*, 2012; Scafaro *et al.*, 2016). For example, a recent investigation into the *Oryza* genus demonstrated that Australian wild rice species have superior growth under heat stress when compared to domesticated rice, while at control temperatures, the domestic variety of rice outperformed the wild rice (Scafaro *et al.*, 2016). The differences between domestic and wild rice were attributed to 19 nonsynonymous polymorphisms in the primary structure of RCA. It has been hypothesised that the wild rice RCA has been selected towards a more thermostable structure to maintain function at the relatively high growth temperatures it experiences in nature, the penalty being a loss of activity at lower temperatures due to increased rigidity. Conversely, the domestic RCA may exhibit increased flexibility of its active sites to operate maximally at lower temperatures. Similar patterns of variation can be seen in the Arizonan creosote bush (*Larrea tridentata*) when compared with Antarctic hair grass (Deschampsia antarctica; Salvucci and Crafts-Brandner, 2004b; Weston *et al.*, 2007). These differences have been linked to changes in photosynthetic performance observed between the two species when exposed to high temperature stress, and thermal instability of spinach (Salvucci and Crafts-Brandner, 2004b). Stability-activity trade-offs such as these have been well described in studies of directed protein evolution in microbial species, where cold-adapted enzymes show increased active site flexibility, and increases in enzyme function are often accompanied by loses

of enzyme stability, a phenomenon that has also been explored for Rubisco itself (Studer *et al.*, 2014; Mehta *et al.*, 2015; Truongvan *et al.*, 2016).

1.6 Exploiting Diversity in Wild Crop Relatives for Abiotic Stress Tolerance

Wild crop relatives represent a vast reservoir of novel genetic material that can be used to improve the tolerance of domestic varieties to biotic and abiotic stressors (Dempewolf and Guarino, 2015; Kaur *et al.*, 2017). Many, if not all, crop species have suffered varying degrees of genetic erosion through the intensive breeding processes that have given rise to modern high-yielding cultivars. For example, when compared with wild species the diversity of the soybean (*Glycine max*) genome is reduced by ~50% and the presence of rare alleles reduced by ~60% (Hyten *et al.*, 2006). Such losses in variation render agriculturally important species susceptible to disease and environmental stresses. While different cultivars of the same species can show variation in traits associated with tolerance to extreme conditions, it is much more likely that the functional diversity needed to improve modern cultivars will be found in wild species adapted to marginal environments where local adaptation to extreme conditions has occurred (Atwell *et al.*, 2014; Dempewolf and Guarino, 2015). However, as seems to be the case for other plant accessory proteins (for example, see Scharf *et al.*, 2012), the interaction between RCA and Rubisco demonstrates a level of genus level specificity (Wachter *et al.*, 2013). Because of this, it may be that plants expressing RCA from another species might exhibit reduced Rubisco content or have lower rates of Rubisco activation (Fukayama *et al.*, 2012). The genus-level specificity of the RCA-Rubisco interaction represents a significant gap in the current understanding of carbon fixation. It is therefore vital to understand the genetic variation of RCA in wild species closely related to domesticated species of interest, especially if RCA is to be a viable tool for improving Rubisco catalysis via transgenic work. Likewise, it is critical to understand how different isoforms of RCA are regulated at high temperature in thermotolerant species.

2. Objectives of the Present Study

I believe that if the thermostability of RCA plays a critical role in the photosynthetic response of the focal species to high temperature then variation will exist either in the primary sequence of the RCA enzyme, or there will be some difference in the regulation of RCA isoforms between species. Therefore, the present study has three main aims:

a) to establish the sensitivity of photosynthesis of domesticated species and their wild relatives to above-optimal temperatures (here, thermotolerance is defined as maintenance of photosynthetic rates at supra-optimal temperature),
b) to quantify the abundance of the various RCA isoforms following heat treatment in a selection of these species via quantitative mass spectrometry, and
c) to characterise the RCA genes in wild species of cotton and tobacco that are adapted to arid zones of Australia, as well as RCA in three populations of *Themeda triandra*

The species utilised in the present study to address the above aims are: *Gossypium hirsutum* (domestic cotton), *G. robinsonii, G. bickii, G. australe, G. sturtianum, Nicotiana tabacum* (domestic tobacco), *N. megalosiphon, N. benthamiana, N. gossei, N. africana*, as well as three populations of *Themeda triandra* (Tasmania, Sydney, and the Northern Territory).

3. Materials and Methods

3.1 Plant Material

3.1.1 Collection of germplasm

Experimental plants were grown from seed or seedlings sourced from various locations and sources (see Table 1).

Table 1. Source of germplasm, ploidy and genomes of the three genera used in this study.

Species	Source	Ploidy and Genome
Gossypium hirsutum	MQ seed collection	4n (allotetraploid); AADD
G. australe	MQ seed collection	2n; GG
G. sturtianum	MQ seed collection	2n; CC
G. bickii	MQ seed collection	2n; GG
G. robinsonii	Nindethana Seed Company	2n; CC
Nicotiana tabacum	MQ seed collection	4n (allotetraploid); SSTT
N. africana	MQ seed collection	4n (?)
N. megalosiphon	AusPGRIS	Unknown
N. gossei	AusPGRIS	Unknown
N. benthamiana	AusPGRIS	4n (allotetraploid); SSNN
Themeda triandra **(Mount Wellington, TAS)**	Collected and supplied by Louise Gilfedder, University of Tasmania	Unknown
T. triandra **(Myrtleford, VIC)**	Supplied by Ian Chivers, Native Seeds Pty Ltd	Unknown
T. triandra **(Jindabyne, NSW)**	Supplied by Ian Chivers, Native Seeds Pty Ltd	Unknown
T. triandra **(Glenbrook, NSW)**	Seedlings supplied by Blake Hawke, Dragonfly Environmental	Unknown
T. triandra **(Forbes, NSW)**	Supplied by Ian Chivers, Native Seeds Pty Ltd	Unknown
T. triandra **(Dalby, QLD)**	Supplied by Ian Chivers, Native Seeds Pty Ltd	Unknown
T. triandra **(Rainbow Valley, NT)**	Collected and supplied by Richard Jobson, Sydney RBG	Unknown

Footnote: MQ = Macquarie University; AusPGRIS = Australian Plant Genetic Resource Information Service; RBG = Royal Botanic Gardens.

3.1.2 Biogeography – Species Distribution and Climate Data

The breadth of temperature conditions experienced by each species of *Gossypium* and *Nicotiana* across its Australian range was approximated by matching the occurrence records in the Australian Virtual Herbarium (AVH) to long-term climate records for Australia. For each species, occurrence records (latitude and longitude coordinates associated with vouchered herbarium specimens) were downloaded from the AVH in March, 2017. The AVH provides digitised records from vouchered specimens held within Australia's nine major herbaria and is the largest source of occurrence data for this continental flora (CHAH, 2009).

A preliminary dataset of 1,242 species occurrence records was cleaned to remove taxonomic and spatial errors. The final dataset contained 1,041 occurrence records (*Gossypium bickii* (n = 172); *G. robinsonii* (n = 223); *Nicotiana benthamiana* (n = 273); *N. gossei* (n = 184); *N. megalosiphon* subsp. *megalosiphon* (n = 189)).

Cleaned occurrence records were imported into ArcGIS 10.3 (ESRI, 2011) as point shapefiles. Long-term average climate conditions for Australia over the period 1970-2014 were accessed from ANUCLIM1.0 (Hutchinson *et al.*, 2014). Daily maximum temperature (Tmax) was the environmental variable of interest. Daily maximum temperatures in ANUCLIM are represented as gridded data at a 0.01° resolution (approximately 1km) across Australia. Tmax is modelled in ANUCLIM1.0 by expressing each daily value as a difference anomaly with respect to the gridded 1976-2005 monthly mean daily maximum temperatures. To calculate the breadth of Tmax experienced across the range of each species, occurrence records were overlayed on gridded data and values extracted at each location. The niche breadth for Tmax was then calculated as the difference between the highest and lowest values of Tmax encountered by the species across the range. Tmax breadth was then used as an explanatory variable in regression analyses exploring the relationship between polymorphism and climatic conditions.

For *Themeda triandra*, Tmax values were extracted for each of the nine provenances used in the study. This approach deliberately contrasts with the approach for *Gossypium* and *Nicotiana* because the intention here was to look for a polymorphism response in relation to climatic conditions *within* a single species which spans a wide breadth of conditions.

3.1.3 Germination and plant establishment

The conditions required to germinate seeds varied between genera. *Nicotiana* seeds were sprayed with a 10 mM gibberellic acid and 5 mM $CaCl_2$ mix to promote the breaking of dormancy. Seeds were germinated on moistened absorbent cotton wool in petri dishes or through direct sowing into pots in a climate controlled growth cabinet for approximately one week. Cabinets were set at 30/22^{o}C in an equal day/night regime, with a peak irradiance of 1000 µmol photons m^{-2} s^{-1} midway through the photoperiod. Seedlings established on petri dishes were transplanted onto the soil surface (soil as described below) when the radicle was approximately 5 mm long. Seeds were distributed evenly on the soil surface (40:30:30 sandy clay loam) and left uncovered for exposure to natural light at a 30/22^{o}C day/night temperature. Direct sown seed was thinned to one seedling per pot approximately two weeks after germination. Soil was fertilised with 2.5 g/L of a 70-day fertiliser (Yates, Victoria, Australia) and sprayed lightly each day to keep moist.

Gossypium was germinated by imbibing seed in water for 30 minutes, followed by surface sterilisation with a five-minute wash in 50% bleach (White King Premium Bleach) and five one-minute washes in water to remove traces of bleach. Upon emergence of the radicle, seedlings were planted ~1 cm into 0.25-L punnets containing soil (as described above) and a fine layer of vermiculite was applied to maintain soil moisture. Punnets were kept in trays of water in a glasshouse at 30/22^{o}C day/night temperature. Once seedlings were ~15 cm tall they were transplanted into 2.2 L pots, with one plant per pot. Pots were kept in trays of water for four months, after which time the trays were allowed to dry and subsequent watering was done every two days by spraying the soil surface.

Themeda triandra seed were surface sterilised as described for *Gossypium*. Seed was placed onto moist cotton-lined petri dishes and exposed to a 35/15^{o}C day/night temperature regime to break dormancy. Germination of seed from all populations was very poor, especially for genotypes from the northern populations. Poor germination reduced the number of accessions which could be studied (see Table 1). Upon the emergence of the radicle, seedlings were planted ~0.5 cm into punnets containing soil (as described above) and a fine layer of vermiculite was applied. Punnets were initially kept outdoors where seedlings survived better. Then, at the four-leaf stage, the seedlings were transplanted to 2.2 L pots of soil.

Established seedlings of all species were maintained in climate controlled glasshouses with 12-h light and 30/22^{o}C day/night temperature regimes. Natural light (midday photosynthetic photon flux density [PPFD] = 1106 ± 13.5 µmol m^{-2} s^{-1}) was supplemented with blue/red LED lighting when the light levels in the glasshouses fell below 700 µmol photons m^{-2} s^{-1}. Soil and plants were sprayed with a lime sulphur solution (20 ml/L; Ausgro, Victoria, Australia) to control fungal infections while insect and soil pests (e.g. mites, mealy bugs, nematodes etc.) were controlled with a garlic-based organic infusion. Established seedlings were watered every two weeks with a commercial 23:3.95:14 nitrogen: phosphorus: potassium water-soluble fertiliser (with added trace elements) at a concentration of 1 g/L (AQUASOL, Yates, Australia).

3.2 Experiment One: Effect of heat on gas exchange

3.2.1 Experimental Design

All *Gossypium* and *Nicotiana* species (see Table 1) were included for gas exchange measurements, however due to poor germination for the Tasmanian and Northern Territory populations, only individuals from the Sydney population of *Themeda* were included in this part of the study. See Section 3.1.2 for information on the germination, growth and maintenance of the species. All *Gossypium* plants were between six and seven months old, while the individual plants of *Nicotiana* and the Sydney *Themeda* population were between three and four months old.

Following a small pilot study to establish the sub-lethal temperature threshold (data not shown), heat tolerance was tested at 38^{o}C. This temperature regime met the definition of a heat wave as five days where the maximum temperature was at least 5^{o}C above the average growth conditions of 30^{o}C (Frich *et al.*, 2002).

To heat treat plants, mature plants were transferred from glasshouses to large growth cabinets (Conviron, Model no.: PGR15) for a 3-d pre-treatment prior to the 5-d treatment. The replicates in this study (n = 6 for all species) were divided evenly between two large growth cabinets to account for any effect of individual cabinets.

There were four measurement periods in total: *Gossypium* species at (1) 30^{o}C and (2) 38^{o}C; *Nicotiana* species and *Themeda* at (3) 30^{o}C and (4) 38^{o}C. There was a five-day acclimation window between each measurement period. The maximum temperature of the growth cabinets for each

measurement period was reached over a two-hour ramping period at the beginning of the photoperiod. The same plants were used across both within-genus measurement periods. Night-time temperature was always 22ºC and humidity was maintained at 70% throughout the course of the measurement periods. Light levels in the growth cabinets were set to 1000 µmol photons m^{-2} s^{-1} with ramping to peak irradiance over the course of two hours.

3.2.2 Gas Exchange Measurements and Tissue Collection

Gas exchange was measured using two Licor 6400 portable systems (Li-Cor, Lincoln, USA). One Licor was used per growth cabinet, with the machines randomly assigned to each growth cabinet at each measurement period. A PPFD of 2000 µmol m^{-2} s^{-1} was used for all measurements, which were taken within the interior of the growth chambers. Measurements began shortly after the commencement of the photoperiod. The leaves examined were all healthy and fully expanded. The CO_2 in the reference chamber was set to 400 µmol mol^{-1} and the block temperature was set to either 30ºC or 38ºC. Three leaves per plant (technical replicates) were measured for a total of six plants per species (biological replicates) at each measurement period. One of the three technical replicate leaves from each plant was harvested, weighed, and immediately frozen in liquid nitrogen. These samples were stored at -80ºC for protein extraction.

3.3 Experiment Two: Effect of heat on the abundance of RCA in leaf tissue

3.3.1 Species Selection

Not all species could be included in this proteomics experiment. I was interested in testing differences in RCA abundance in thermotolerant wild species and their domestic counterparts. So, it was determined that *G. robinsonii* (n = 3) would be selected for comparison to *G. hirsutum* (n = 3), and *N. megalosiphon* (n = 3) would be selected for comparison to *N. tabacum* (n = 3; see gas exchange results in section 4.2). Only individuals from the Sydney population of *Themeda* were included in this preliminary proteomics study.

3.3.2 Total Protein Extraction and Sample Preparation

Solvent washing

A 50-100 mg subset of each frozen leaf sample was disrupted with sharp silicon carbide particles 4 mm in diameter and with a density = 3.2 g cc^{-1}, or carbide beads (Daintree Scientific, Australia) in sterile 2-mL stainless steel microvials with a TissueLyser II (Qiagen, Hilden, Germany) set to 20 hz

until a fine powder was produced. The resulting powder was suspended in -20°C 98% 1:1 chloroform:ethanol, 50 mM ammonium acetate, 1% 2-mercaptoethanol, 5 mM diethyldithiocarbamate (DEDTC), 2 mM dithionite. Tubes were then centrifuged at 4300 g, 5 min, 0°C, supernatants discarded, and the same solvent extraction repeated. Pellets were then suspended in -20°C hexane, centrifuged as above, and supernatants discarded. Pellets were then suspended in -20°C 98% ethanol, 50 mM ammonium acetate, 1% 2-mercaptoethanol, 5 mM DEDTC, 2 mM dithionite solution, centrifuged, and supernatants discarded. Pellets were suspended in 0.4 M boric acid, 0.3 M trisodium citrate, 1 M LiCl, 50 mM glycine, 50 mM EDTA, 1% Lithium dodecyl sulfate (LiDS), pH 9.0 with NaOH, 20 µM E-64, 1 mM benzamidine, 2% 2-mercaptoethanol, 2 mM Na_2SO_3 and stored at -80°C.

Phenol extraction, urea resuspension, and alkylation

To each resuspended frozen pellet 5 µL of ovalbumin from chicken egg white (0.5 mg mL^{-1}) was added as an internal standard per cm^2 of starting leaf material. 800 uL water-saturated phenol was added and samples thawed while mixing. The suspensions were then sonicated at 80°C, 10 min. Cooled samples were centrifuged at 14,500 g for 5 min at room temperature. The top phenol phases were transferred to 15 mL centrifuge tubes. The remaining aqueous phases were re-extracted twice with 800 uL phenol and the phenol phases combined with the first phenol phases. 4 mL ice cold 50 mM glycine, pH 9.0 with NaOH, 1 M LiCl, 5 mM EDTA, 1% 2-mercaptoethanol, 1 mM Na_2SO_3 was added to the phenol phases, the samples were mixed by inversion, and centrifuged at 1500 g, 5 min, at room temperature. The aqueous top layer was discarded, the phenol phases re-extracted with with LiCl-free cold glycine buffer, centrifuged again, and the aqueous phases discarded. Phenol phases were stored at -80°C.

Protein was precipitated from thawed phenol phases overnight on ice with 1:1 ethanol:diethyl ether, 0.1 M ammonium acetate. Samples were centrifuged in -20°C buckets at 1500 g for 10 min. Supernatants were discarded and pellets were resuspended twice in cold 98% 1:1 ethanol:diethyl ether, 1% water, 1% glycerol by vortexing and rasping, followed by centrifugation and disposal of the supernatants. Partially air-dried pellets were resuspended in 8 M urea, 50 mM lactic acid, 0.2 M LiCl, 1 mM EDTA, 2% LiDS, 10 mM TCEP, 5 mM cysteine, 1 mM Na_2SO_3 and mixed by vortexing. 8 uL N-methylmorpholine was added, followed by sonication at room temperature for 10 min. Then samples were alkylated by the addition of 10 uL 50% 2-vinylpyridine in methanol followed by a 1 hr incubation. Reactions were stopped by the addition of 5 uL 2-mercaptoethanol and samples

stored at -80°C. Proteins were quantified using a FluoroProfile protein assay kit (Sigma); results are reported in ovalbumin equivalents.

Sample preparation for enzyme digestion

Aliquots of 50 ug total protein from each sample were spiked with 3.3 pmol of ^{13}C- and ^{15}N-labelled lysine and arginine peptide standards for ovalbumin, the large subunit of Rubisco, and RCA (see Supplementary Material table 7.1.3 for labelled target peptides). Samples were then methanol-chloroform re-extracted using a modification of Wessel and Flugge (1984). 250 uL 67% methanol, 25% chloroform, and 8% water solution were added to each sample, followed by mixing, 500 uL ice cold 10 M ammonium acetate was added, mixing, and then centrifugation at 15000 g, room temperature, 1 min. The top aqueous phases were discarded, then 500 uL ice cold water-saturated diethyl ether was added to the interphases and lower chloroform phases. Suspensions were mixed, then 100 uL cold 25% TFA in ethanol was added to make the phases miscible, resulting in protein precipitate. Samples were then centrifuged at 15000 g for 10 minutes at 4°C and supernatants discarded. Pellets were washed with ice cold 0.1 M triethanolamine, 0.1 M acetic acid, 1% water, and 1% DMSO in 1:1 ethanol:diethyl ether, centrifuged as above, supernatants discarded, and pellets stored at -20°C.

Sample digestion

Sample proteins were digested with Lys-C and trypsin. Partially air-dried pellets were suspended in 25 µL 0.3% Rapigest (Waters, Rydalmere, Australia), 0.2 M N-methylmorpholine, 40 ng/µL Lys-C (Wako Pure Chemical Industries, Osaka, Japan). Samples were incubated on a Thermomixer (Eppendorf, Hamburg, Germany) at 45°C and 1200 rpm for 15 minutes, then in a sonicating water bath for 45 minutes, 40-45°C. Then 5 µL of 0.1 µg/µL trypsin (Sigma) was added to each sample followed by overnight incubation at 37°C. Digestions were stopped with 6 uL 20% (v/v) HFBA in water and incubated at 37°C for 45 minutes, according to the Rapigest manufacturer's protocol, and centrifuged at 4°C and 15000 g for 10 minutes. Peptide containing supernatants were used for mass spectrometry.

3.3.3 Mass Spectrometry

An initial Independent Data Acquisition (IDA) dataset was acquired at the University of Sydney, which, along with a secondary IDA dataset collected at the Australian Proteomics Analysis Facility

(APAF), allowed for the identification of target peptides for later scheduled High Resolution Multiple Reaction Monitoring (sMRM-HR) analysis.

Initial IDA Analyses – Target Peptide Identification

A 6600 TripleTOF mass spectrometer (AB Sciex) coupled to an Eksigent LC with an autosampler (Ekspert nanoLC 400; Eksigent, Dublin, CA) was used for IDA analysis of test samples for the identification of target peptides. 10 µL (1 µg) of sample was injected a reverse phase peptide C18 self-packed trap for pre-concentration and desalted for six minutes the with loading buffer (3% (v/v) acetonitrile, 0.1% (v/v) formic acid) at a flow rate of 630 nl/second. The peptide trap was switched to a 75 um x 10 cm, 1.9 um ReproSil-Pur C18-AQ column (Dr. Maisch GmbH, Ammerbuch, Germany). Peptides were eluted and separated from the column using the buffer B (80% (v/v) acetonitrile, 0.1% (v/v) formic acid) gradient: 3% at 0 minutes, 8% at 5 minutes, 20% at 45 minutes, 30% at 60 minutes, 100% at 61 minutes (for seven minutes), and 3% at 68 minutes, at a flow rate of 300 nL/minute. The column was re-equilibrated with buffer A (0.1% (v/v) formic acid in water) after each sample run. The peptides were analysed in the positive ion nanoflow electrospray mode in an information dependent acquisition (IDA) mode.

TOF-MS survey scan was acquired at m/z 350-1500 with 0.25 second accumulation time, with 20 most intense precursor ions (2+ - 5+; >10 cps) in the survey scan consecutively isolated for subsequent automated measurement of their corresponding product ion analysis. Dynamic accumulation was used with a 10 ppm mass tolerance. Product ion spectra were accumulated for 100 milliseconds in the mass range m/z 100-1500 with rolling collision energy.

Secondary IDA and MRM-HR Analyses

The secondary IDA run was identical to the initial IDA run, with the following changes: an Eksigent Ultra-nanoLC-1D system (Eksigent, Dublin, CA) was employed for IDA and sMRM-HR analysis, loading buffer was 2% (v/v) acetonitrile with 0.1% (v/v) formic acid, an in-house packed analytical column (150 µm × 10 cm) with solid core Halo C18, 160 Å, 2.7 µm (Bruker) was used in place of the C18-AQ column, buffer B was 99.9% (v/v) acetonitrile, the gradient starting from 2% and increasing to 10% for 10 minutes and to 35% over another 78 minutes at a flow rate of 600 nL per minute, and the TOF-MS survey utilised dynamic exclusion with a 30 second and 4 Da window.

For sMRM-HR acquisition, a TOF-MS scan was acquired (m/z 350 – 1500, 0.25 second) with a maximum number of twelve candidate precursor ions (2+ - 5+; counts > 200,000,000) with no dynamic exclusion. Product ion scans for 37 pre-selected precursor ions with their corresponding retention time (± 300 second, intensity > 0) were chosen for inclusion list.

Data processing

Protein databases were constructed for each genus. The *Gossypium* protein database contained 87,998 sequences collected from Phytozome v1.1 (US Department of Energy, California, USA), Uniprot (The UniProt Consortium, 2017), the translated PCR sequence data from the present study (including N-terminal and C-terminal fragments that were missed during PCR; these were appended onto the truncated PCR data from highly related sequences following BLAST searches of PCR data), and the common Repository of Adventitious Proteins (cRAP) through the Global Proteome Machine Organisation (GPM) database. The *Nicotiana* database contained 113,599 protein sequences collected from SolGenomics release 1.0.1 (Fernandez-Pozo *et al.*, 2015), Uniprot, translated PCR data from the present study (including missing N- and C- terminals as above), and the cRAP through GPM. The *Themeda* database was based on the *Sorghum* proteome, these two genera being closely related taxonomically. The database contained 47,331 protein sequences collected from Phytozome v3.1.1, Uniprot (mitochondrial and chloroplastic proteomes), PCR data from the present study (including N- and C- terminals as above), and the cRAP through the GPM. ProteinPilot was used to analyse IDA data, including the possibility of amino acid substitutions, which allowed for the identification of mutant peptides, against the constructed databases. Supplementary Material table 7.1.3 illustrates the targeted proteins, the corresponding peptides, the retention times of the peptides and the fragment ions used for quantitation of the samples. The peptide retention time variation window was 5 min and the product ion variation window was ± 0.05 Da for peak integration. All the data collected were processed by MultiQuant (AB Sciex v 2.1.1) software.

3.4 Experiment One and Two Data Analysis

3.4.1 Physiology Data

The design of the physiology experiment required the use of a mixed-effect model ANOVA. The lme4 (Bates *et al.*, 2015) and lmerTest (Kuznetsova *et al.*, 2016) packages in the R environment (R Core Team, 2017) were used for these analyses. Analyses were conducted on the level of the

Genus, allowing for comparisons between each *Gossypium* species, and each *Nicotiana* species, and the effect of high temperature on the Sydney population of *Themeda triandra.* Models with and without an interaction term between species and temperature were generated. Log likelihood ratio tests were used to assess the goodness of fit of each model to each parameter for each Genus. The results of these analyses were corroborated with the generation of AIC values. These tests were used to decide whether an interaction term should be included in downstream models and, thus, whether an interaction existed between species and temperature. In all models, pot ID was treated as a random factor to account for the variability introduced to the model by re-measuring the same pot at control and treatment temperature. Once a model was decided upon (i.e. whether to include the interaction term or not) models were re-fitted without intercepts so that individual coefficients were tested against zero. Additionally, the glht function of the multcomp package (Hothorn *et al.*, 2008) was used to compare the means of groups in a biologically interesting manner, in a manner similar to that of a post-hoc Tukey test. Outliers in the data set were identified with Cleveland Plots through the ggplot2 package (Wickham, 2009). Any identified outliers were scrutinized in the context of the raw data to decide whether they were erroneous measurements or a result of inter- and/or intra-species variability. The assumption of equal variances of residuals was assessed by plotting the residuals against the fitted values of the model. The assumption of normality of residuals was assessed with the qqplot function of the car package (Fox and Weisberg, 2011).

3.4.2 RCA and RbcL Abundance Data

See Supplementary Material table 7.1.3 for a list of target peptides. The top three transitions per peptide were used for all calculations of absolute and relative abundance. Initially, the inclusion of all transitions per peptide (the number of which varied across peptides) resulted in large coefficients of variance for the summed peak areas of the peptides. Using only the top three transitions per peptide substantially reduced the variance of the summed peak area for 722 (78.6%) of the 919 transitions obtained through MRM-HR.

Absolute Abundance

To find the absolute abundance (µmol target m^{-2} leaf tissue) of RCA and RbcL in *G. hirsutum*, *G. robinsonii*, *N. tabacum*, *N. megalosiphon*, and *T. triandra* at control and high temperatures, the peak areas of the top three transitions of the isotope-labelled and corresponding non-labelled native peptides were summed and corrected using internal labelled and non-labelled ovalbumin

peptides (see Supplementary Materials table 7.1.3 for peptides, and Fig. 11 for peptide locations in target proteins). RCA, RbcL, and ovalbumin were each quantified using four target peptides – two native peptides and their isotope labelled counterparts. In detail, summed peak areas for each peptide were corrected using the equation:

$$\frac{UT}{Av\ UO} * \frac{Av\ LO}{LT} * n_O\ cm^{-2}\ leaf$$

Where, UT is the summed peak area of the native un-labelled target peptide, Av UO is the average summed peak area of the two un-labelled ovalbumin peptides, Av LO is the average peak area of the two labelled ovalbumin peptides, LT is the summed peak area of the isotope labelled target peptides, and $n_O\ cm^{-2}$ leaf is the moles of ovalbumin per leaf area spiked into each sample ($5.64*10^{-11}$). This equation was applied to the two target peptides of RCA and RbcL and an average absolute abundance was obtained for RCA and RbcL per sample. These values were used to also obtain a ratio between RCA and RbcL for each sample.

Due to the repeated measures design of the experiment, mixed effect models were fit to detect any differences in the absolute abundance of RCA and the ratio of RCA to RbcL within each genus and across temperature groups. See section 3.4.1 for a detailed description of the mixed effect model approach.

Relative Abundance

Following initial IDA analyses, a number of peptides unique to different RCA isoforms, including mutant forms of existing known peptides, across the species of interest were identified. During MRM-HR, these peptides were ascribed to one of four RCA isoforms present in the *Gossypium* protein database, one of four isoforms in the *Nicotiana* database, and a single RCA isoform in the *Sorghum*-based *Themeda* database (see Supplementary Materials table 7.1.3 for peptides, and Fig. 8 for peptide locations in target proteins). These peptides were used to estimate how the abundance of the identified RCA isoforms change within a species and across temperature groups, relative to the internal ovalbumin un-labelled peptide standards. The data analysis for this section was generated using the Real Statistics Resource Pack software (Release 5.1). Copyright (2013 – 2017) Charles Zaiontz. www.real-statistics.com.

Peak areas of the top three transitions of the unique peptides were summed, and the summed peak areas were corrected using the average summed peak areas of the two un-labelled

ovalbumin peptide internal standards. An average corrected sum peak area was obtained for each sample. Due to the repeated measures design of this experiment, paired T-tests were performed for each species, using temperature as the independent variable, to test for changes in the relative abundance of each identified isoform across temperatures.

The response of individual peptides to temperature was estimated by calculating a fold-change ratio of the relative abundance of each peptide between the 30°C and 38°C groups within each species. This was done in an attempt to detect any alternative splicing events of the RCA C-terminus domain that may have been taking place across temperature groups. One-way ANOVAs were used to test for differences in mean fold change ratios for each isoform in each species.

3.5 Experiment Three: Rubisco activase sequencing

3.5.1 Tissue Collection and RNA Extraction

Leaf discs with an area of approximately 2.74 cm^2 for cotton or 3.08 cm^2 for tobacco were excised from healthy, mature leaves of each species with sharpened brass pipes and immediately frozen in liquid nitrogen. Discs of these sizes yielded enough leaf tissue (50-100 mg) for RNA extraction. It was determined that a 12.8-cm long piece of leaf tissue with an area of 4.1 cm^2 from *Themeda triandra* would yield the required 50-100 mg of tissue for RNA extraction. Frozen leaf tissue was ground as in section 3.3.2.

Initially, total RNA was extracted from homogenised *T. triandra*, *Gossypium,* and *Nicotiana* tissue with the ISOLATE II RNA Plant Kit (Bioline, Alexandria, Australia) following the manufacturer's instructions. RNA concentration and purity was estimated with a NanoDrop 2000c (Thermo Scientific, Waltham, United States). Each sample was measured four times. RNA integrity was assessed on 1% agarose gels. From these gels and measurements, it was determined that RNA was readily extracted from the leaves of *T. triandra* and *Nicotiana* tissue, however no *Gossypium* sample yielded good quality RNA with the Bioline kit. Instead, RNA from *Gossypium* tissue was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich, St. Louis, United States), according to manufacturer's instructions.

3.5.2 cDNA Synthesis

cDNA was synthesised non-specifically from the total extracted RNA pool using the Bioline Tetro cDNA Synthesis kit (Bioline, Alexandria, Australia). Briefly, an aliquot of RNA was taken from each sample so that 2 - 5 ng of RNA was included in each cDNA synthesis reaction. The master mix for these reactions was made up of an oligo $(dT)_{18}$ mix, a 10 mM dNTP mix, 5× reverse transcription buffer, 10 u μl^{-1} Ribosafe RNase Inhibitor, and 200 u μl^{-1} Tetro Reverse Transcriptase. Reaction mixes were made up to a total volume of 20 µl with DEPC-treated water. To synthesise cDNA, the reaction mixes were incubated at 45°C for 30 min in a Px2 Thermocycler (Thermo Scientific, Waltham, United States). Reactions were terminated after this period by incubating the samples at 85°C for five minutes. cDNA samples were stored in a -20°C freezer until PCR reactions were carried out.

3.5.3 Primer Design and PCR Conditions

Gene-specific primers were designed for the amplification of different Rubisco activase (RCA) isoforms and alcohol dehydrogenase (ADH) cDNA fragments in the total cDNA pool from all species (Supplementary Material section 7.1.1). ADH was included to act as a 'control' gene that should be under little or no selective pressure in arid environments and thus serve as a baseline to which SNPs in RCA could be compared between species.

To design primers specific for RCA 1, RCA 2 and ADH in wild, non-sequenced species, the Uniprot (UniProt Consortium, 2017) protein identification numbers of previously sequenced isoforms in domestic counterparts were obtained (Table 2).

Table 2. Protein identification numbers of known peptide sequences in domestic cotton, tobacco and rice. These numbers were used to search translated nucleotide databases through NCBI's BLAST algorithm. *Gossypium hirsutum* sequences were used as a database probe for all wild cottons. *Nicotiana tabacum* sequences were used as a database probe for all wild tobacco. *Oryza sativa* sequences were used as a database probe for all populations of *Themeda triandra*.

Species	Target Protein/Gene	Protein Identification No.
G. hirsutum	RCA 1 (beta)	Q9AXG0
G. hirsutum	RCA 2 (alpha)	Q9AXG1
G. hirsutum	ADH	Q42763
N. tabacum	RCA 1 (beta)	Q40460
N. tabacum	RCA 2 (beta)	Q40565
N. tabacum	ADH	Q42953
O. sativa	RCA	P93431
O. sativa	ADH	Q2R8Z5

These protein ID numbers were submitted to NCBI's protein-nucleotide BLAST algorithm (tblastn) to search translated nucleotide databases using a protein query. The BLAST output was limited to

nucleotide sequences that fell within the taxonomic family (i.e. Malvaceae for cottons, Solanaceae for tobaccos, and Panicoid for kangaroo grass) of the query protein sequence. Genomic DNA sequences were also excluded. The output of these searches was used to obtain a number of sequences for the genes of interest within each taxonomic Family. These were then aligned in UGENE (Okonechnikov *et al.*, 2012) and a consensus sequence was derived (Table 3).

Table 3. Within-Family consensus sequences derived from the alignment of previously-sequenced genes-of-interest of species closely related to this study's species.

Species	Family	Target Gene	No. Aligned Sequences	Consensus Sequence Size (bp)
Wild cotton	Malvaceae	RCA 1	74	1357
		RCA 2	11	1354
		ADH	242	1183
Wild tobacco	Solanaceae	RCA 1	92	1351
		RCA 2	31	1344
		ADH	147	1162
T. triandra	Panicoid	RCA	9	1435
		ADH	135	1131

"Terminal primers" (primers designed to cover the full consensus sequence) and "Middle primers" (primers designed to amplify the front and back halves of the consensus sequences independently) were designed using default parameters in Primer3 (Untergasser *et al.*, 2012). Primer pairs generated through Primer3 were submitted to the Operon Oligo Analysis Tool (Eurofins Genomics, Ebersberg, Germany) to check the quality of the designed primers. Once quality of primers was assessed, two Terminal and two Middle primer pairs for each gene of each Family were chosen for synthesis (see Supplementary Material Table 7.1.1). Oligo DNA primers were synthesised by IDT (Integrated DNA Technologies Pte. Ltd., Baulkham Hills, Australia). After a number of weak or failed PCR amplifications, a second batch of primers was designed using a more defined target region in the Primer3 algorithm (see Supplementary Material Table 7.1.2). For *Gossypium* and *Nicotiana,* RCA sequence obtained using the first set of primers were used to inform design of new primers, while for *Themeda* highly conserved regions within the RCA genes of several *Oryza* species and several species within the Panicoid family were used for primer re-design.

Lyophilised primers were re-suspended in DEPC-treated water to a final concentration of 100 mM. From this solution, a 20 µM working mix was created and the 100 µM stock was kept at -20°C.

The PCRs were run as per the manufacturer's instructions for the MyTaq Red Mix PCR kit (Bioline, Alexandria, Australia). Starting RNA concentrations were used as a proxy for final cDNA concentrations in the calculation of how much cDNA stock template to add for the PCR reactions. All reactions contained 5 pM forward and reverse primers, 25 µl of the 2× MyTaq Red Mix

(containing all reagents, including stabilisers, for PCRs), 200 ng of cDNA template, and a volume of ddH_2O to make up the total 50 µl. The PCR conditions used were an initial denaturation phase of 95°C for one min, followed by 35 cycles of denaturation at 95°C for 15 s, annealing at 55°C for 15 s, extension at 72°C for 10 s and a final extension step of 72°C for 10 s. The PCR products were visualised by gel electrophoresis. From each sample, 1 µl was run alongside 5 µl of Hyperladder 1kbp (Bioline, Alexandria, Australia) at 100 V for 55 min on a 1% agarose gel stained with 2 µl/100 mL GelRed. Gels were visualised under ultra-violet light and examined using GeneSnap software (SynGene, Bangalore, India). Successful PCR products, as determined by a strong band present between the 1000 and 1500 bp marks, were sent to Macrogen for purification and sequencing.

3.5.4 PCR Product Purification and Sequencing

Initially, PCR product purifications were attempted using an Isolate II PCR and Gel Purification kit (Bioline, Alexandria, Australia). This was unsuccessful for every sample. The resulting eluent yielded no DNA, as determined by spectrophotometry with a NanoDrop 2000c (Thermo Scientific, Waltham, United States). As a result, all raw PCR products were sent to Macrogen for purification. Briefly, each sample was treated with 0.4 µl ExoSAP-IT (Thermo Scientific, Waltham, United States) per µl of sample. This mixture was incubated at 37^{o}C for 15 minutes to degrade remaining primers and nucleotides. The reaction was terminated by incubating the mixtures at 80^{o}C for 15 min.

Sequencing reactions were performed by Macrogen in the DNA Engine Tetrad 2 Peltier Thermal Cycler (BIO-RAD, California, United States) using the ABI BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, California, United States) following the protocols supplied by the manufacturer. Single-pass sequencing was performed on each template using the forward and reverse primers outlined in Supplementary Material Tables 7.1.1 and 7.1.2. The fluorescent-labelled fragments were purified from the unincorporated terminators with the BigDye XTerminator® Purification Kit (Applied Biosystems, California, United States). The samples were injected for electrophoresis in an ABI 3730xl DNA Analyzer (Applied Biosystems, California, United States).

3.5.5 In silico RCA Sequence Clean-up, and Alignment

Sequences were aligned to reference RCA sequences (*Gossypium* RCA 1 reference: AF329934; *Gossypium* RCA 2 reference: NM_001327460; *Nicotiana* RCA 1 reference: NM_001326055.1;

Nicotiana RCA 2 reference: Z14980; *Zea mays* RCA [surrogate for *Themeda*] reference: NM_001111451) in UGENE using the Sanger Sequencing tools. All alignments included Forward and Reverse "Terminal" and "Middle" sequences. The chromatograms for each base call were manually interrogated to check the quality of the base calls between the Forward and Reverse sequences. All sequences were truncated by ~100 bp from their 5' and 3' ends to remove poor quality sequence information. Double base calls were present in the chromatograms which may have substantial implications for functional RCA peptides (Supplementary Material Table 4 section 7.2.6). A double base call is defined here as any point in the Forward or Reverse sequences that displays a peak for two bases in one position (for example, see Fig. 2).

Double base calls were further scrutinised in the R coding environment using the packages sangerseqR (Hill *et al.*, 2014) and seqinr (Gouy *et al.*, 1984). The .abi files containing sequence information obtained from Macrogen were read into the sangerseqR package and analysed using default settings. The 'makebasecalls' function of sangerseqR was used to identify likely double base calls above an intensity ratio of 0.33.

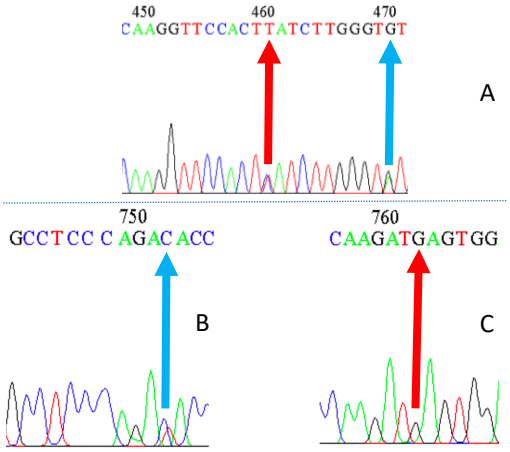


Figure 2. Snippets of Forward and Reverse strand Sanger chromatograms from the sequencing reactions of *N. africana* RCA 1. [A] shows clear double base calls at positions 461 bp and 471 bp in the Forward sequence. [B] shows a double base call at position 753 in the Reverse sequence that corresponds to and complements the double base call at positon 471 in the Forward Sequence, lending supporting the double base call of the Forward sequence. On the other hand, [C] shows a single base call at position ~763 in the Reverse sequence (G) that, in the Forward sequence, was identified as a double base call (T or C) at position ~461. This is evidence against a double base call at position ~461 in the Forward sequence, however the peak for G in [C] is small and there may be a signal for the A base hidden by the prominent A base call of the adjacent nucleotide in the sequence.

After discrepancies between the Forward and Reverse sequences were resolved, a consensus sequence between the Forward sequence and the reverse complement of the Reverse sequence was generated for each sample. Cleaned sequences for each genus (i.e. *Gossypium, Nicotiana, and*

Themeda) were aligned with the MUSCLE algorithm (EMBL, Cambridgeshire, United Kingdom) in UGENE to detect any nucleotide polymorphisms between the species of interest.
The nucleotide sequences were converted to peptide sequences using the UGENE amino acid translation workflow tool and aligned using the MUSCLE algorithm. The correct codon reading frames were selected based on the previously published RCA 1 and RCA 2 isoform sequences from *G. hirsutum*, *N. tabacum*, and *Z. mays* (for *Themeda*), mentioned previously.

4. Results

4.1 Biogeography

The wild Australian species of the present study occupy different thermal niches (Table 4). Within *Gossypium*, *G. robinsonii* occupies the hottest niche, though *G. bickii* occupies the narrowest niche. Of the wild Australian *Nicotiana* species, *N. benthamiana* occupies the hottest niche, though *N. megalosiphon* occupies the narrowest thermal niche. The maximum Summer maximum temperature increase predictably between the locations of the three *T. triandra* populations.

Table 4. Thermal niche information for Australian wild *Gossypium* and *Nicotiana* species, and the three populations of *T. triandra*. Niche estimates are based on gridded interpolations of maximum temperature during summer (December, January, February) from 1950-2000. Minimum and maximum values are the lowest and highest temperatures experienced by the corresponding species within their natural ranges. Breadth values are the differences between maximum and minimum values. All values are in units of ^{o}C.

Species	Summer Tmax (min)	Summer Tmax (max)	Summer Tmax (breadth)	Summer Tmax (mean)
G. robinsonii	34.4	40.3	5.9	37.9
G. bickii	34.2	38.5	4.3	36.6
G. australe	27.1	40.1	13	36.4
G. sturtianum	25.2	38.4	13.2	34.5
N. benthamiana	33	40.1	7.1	37.4
N. gossei	31.5	37.7	6.2	35.3
N. megalosiphon	31.7	37.3	5.6	34.5
***T. triandra* Rainbow Valley, NT**	--	36	--	--
***T. triandra* Glenbrook, NSW**	--	28	--	--
***T. triandra* Mount Wellington, TAS**	--	14	--	--

4.2 Effect of Heat on Gas Exchange Parameters and Leaf Temperature

For each genus, models were fitted for photosynthetic gas exchange rate (µmol CO_2 m^{-2} s^{-1}), stomatal conductance (mol H_2O m^{-2} s^{-1}) and leaf temperature (^{o}C). For *Gossypium* and *Nicotiana*, models were fitted with and without interaction terms and y-intercepts and goodness-of-fit was evaluated using log likelihood ratio tests. For *Themeda triandra* (Sydney population), a single model was fitted as no interaction was possible within this dataset.

Photosynthetic rate and leaf temperature differed between control and heat treatment in *Themeda triandra* from the Sydney region. Stomatal conductance and intercellular [CO_2] data were omitted due to Licor calibration issues. That is, there was a significant decrease in the photosynthetic rate of *Themeda triandra* from the Sydney region at high temperature (38°C), relative to the 30°C control group (F_1 = 39.1, p < 0.001; Fig. 3a) as well as a significant increase in leaf temperature at 38°C, relative to the control group (F_1 = 642.4, p < 0.001; Fig. 3b).

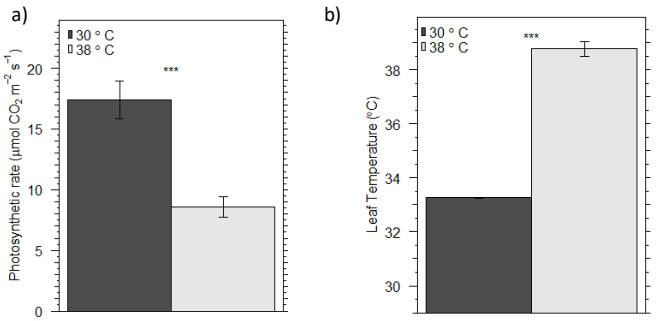


Figure 3. Mean (± SE, n = 6) gas exchange parameters of control and heat-treated leaves of the Sydney population of *Themeda triandra*: (a) net photosynthetic rate, and (b) leaf temperature (note: scale starts at 29°C). Plants were exposed to control conditions (30°C) for five days, then exposed to high temperature (38°C) for five days and re-measured. The significance of differences between the temperature treatments were determined by individual z-tests via the glht function of the multcomp package: '.', p<0.1; '*', p<0.05; '**', p<0.01; '***', p<0.001.

The best-fit models for the effect of heat on the photosynthetic rates, stomatal conductance, leaf temperature, and intercellular [CO_2] of *Gossypium* and *Nicotiana* species included interaction terms between species and temperature, except for intercellular [CO_2] in *Nicotiana* (χ^2_4 = 27.9 p < 0.001; χ^2_4 = 74.6, p < 0.001; χ^2_4 = 25.2, p < 0.001; χ^2_4 = 48.8, p < 0.001; χ^2_4 = 13.6, p = 0.009; χ^2_4 = 25.6, p < 0.001; χ^2_4 = 27.199, p < 0.001; χ^2_4 = 5.1108, p = 0.27, respectively). Therefore, the effect of the heat treatment on photosynthetic rates, stomatal conductance and leaf temperature varied between species in both *Gossypium* and *Nicotiana* (Fig. 4 & Fig. 5).

4.2.1 Gossypium: species by temperature interaction

Photosynthesis

At control temperature (30°C) there were no significant differences in the photosynthetic rates of *G. hirsutum*, *G. bickii* or *G. australe*. However, the photosynthetic rates of two species (*G. robinsonii* and *G. sturtianum*) were significantly higher than all other *Gossypium* species studied (Table 5; Fig. 4a). There was no significant change in the photosynthetic rate of *G. hirsutum* or *G. australe* at 38°C, relative to their controls (z = -0.8, p = 0.998; z = -1.5, p = 0.9, respectively), and there was a non-significant trend of declining photosynthetic rate in *G. bickii* at 38°C (z = 2.9, p =

0.08). There were, however, significant decreases in the photosynthetic rates of *G. robinsonii* and *G. sturtianum* at 38°C ($z = 3.4$, $p = 0.01$; $z = 4.7$, $p < 0.001$).

Despite this decrease, photosynthetic rates in *G. robinsonii* were marginally higher at 38°C than in *G. hirsutum* ($z = -2.8$, $p = 0.09$). The reduction in photosynthetic rate seen in *G. sturtianum* at 38°C relative to the control meant that there was no statistical difference between *G. sturtianum* and *G. hirsutum* at high temperature ($z = -1.7$, $p = 0.71$). Decreased photosynthetic performance at 38°C relative to 30°C in *G. bickii* rendered rates significantly lower than those of *G. hirsutum* at 38°C ($z = 3.1$, $p = 0.03$).

Table 5. Grouping information for mean ($n = 6$) photosynthetic rates (A_n) of the focal *Gossypium* species at control (30°C) and high (38°C) temperatures, expressed relative to *G. hirsutum* A_n values. Groups that do not share a letter are significantly different.

Species	Temperature (°C)	$A_n/A_{n,GH}$	Grouping					
G. hirsutum	30	1	A		C			
G. australe	30	0.89	A		C			
G. bickii	30	0.84	A	B				
G. robinsonii	30	1.66						F
G. sturtianum	30	1.65					E	F
G. hirsutum	38	1		B	C	D		
G. australe	38	0.98	A			D		
G. bickii	38	0.6	A					
G. robinsonii	38	1.35	A			D	E	
G. sturtianum	38	1.23			C	D		

Stomatal conductance

At 30°C there was no significant difference in stomatal conductance between *G. hirsutum*, *G. bickii* and *G. australe*. However, the stomatal conductance of two species (*G. robinsonii* and *G. sturtianum*) were significantly higher than all other *Gossypium* species studied (Table 6; Fig. 4b). There was a significant increase in the stomatal conductance of *G. hirsutum* and a significant decrease in *G. sturtianum* at 38°C, relative to their controls ($z = -3.6$, $p < 0.01$; $z = 3.4$, $p = 0.02$, respectively). Heat did not affect the stomatal conductance of *G. australe, G. bickii* or *G. robinsonii* relative to their controls ($z = -1.3$, $p = 0.9$; $z = 0.8$, p 9981; $z = -0.6$, $p = 0.99$, respectively). At 38°C the stomatal conductance of *G. hirsutum* was not statistically different to any of the wild species, though the conductance of *G. robinsonii* and *G. sturtianum* were both significantly higher than *G. bickii* ($z = 5.1$, $p < 0.01$; $z = 4.2$, $p < 0.01$, respectively), while only *G. robinsonii* was significantly higher than *G. australe* ($z = 3.9$, $p < 0.01$).

Table 6. Grouping information for mean (*n* = 6) stomatal conductance rates (g_s) of the focal *Gossypium* species at control (30°C) and high (38°C) temperatures, expressed relative to *G. hirsutum* g_s values. Groups that do not share a letter are significantly different.

Species	Temperature (°C)	$g_s/g_{s,GH}$	Grouping				
G. hirsutum	30	1	A				
G. australe	30	0.88	A	B			
G. bickii	30	0.94	A	B			
G. robinsonii	30	2.53				D	E
G. sturtianum	30	3					E
G. hirsutum	38	1		B		D	
G. australe	38	0.69	A	B	C		
G. bickii	38	0.41	A	B			
G. robinsonii	38	1.52				D	E
G. sturtianum	38	1.38			C	D	

Leaf temperature

At 30°C the leaves of *G. sturtianum* were significantly cooler than *G. hirsutum* leaves (z = 3.3, p = 0.02), however there were no significant differences between any of the other species (Table 7; Fig. 4c). The leaves of all species increased significantly at 38°C, and there were no significant differences between any of the species at 38°C.

Table 7. Grouping information for mean (*n* = 6) leaf temperatures (L_T) of the focal *Gossypium* species at control (30°C) and high (38°C) temperatures, expressed relative to *G. hirsutum* L_T values. Groups that do not share a letter are significantly different.

Species	Temperature (°C)	$L_T/L_{T,GH}$	Grouping		
G. hirsutum	30	1		B	
G. australe	30	0.99	A	B	
G. bickii	30	0.97	A	B	
G. robinsonii	30	0.94	A	B	
G. sturtianum	30	0.93	A		
G. hirsutum	38	1			C
G. australe	38	0.99			C
G. bickii	38	1.01			C
G. robinsonii	38	0.97			C
G. sturtianum	38	0.96			C

Intercellular [CO_2]

At 30°C *G. hirsutum* had significantly less internal [CO_2] than *G. robinsonii* and *G. sturtianum* (z = -4.7, p < 0.001; z = -5.3, p < 0.001; Table 8). However, at 38°C there was no significant difference in the intercellular [CO_2] of any *Gossypium* species.

Table 8. Grouping information for mean (n = 6) intercellular [CO_2] (C_i) of the focal *Gossypium* species at control (30°C) and high (38°C) temperatures, expressed relative to *G. hirsutum* C_i at 30°C. Groups that do not share a letter are significantly different.

Species	Temperature (°C)	$C_i/C_{i,GH}$	Grouping		
G. hirsutum	30	1	A		
G. australe	30	1.17	A	B	
G. bickii	30	1.2	A		C
G. robinsonii	30	1.32		B	C
G. sturtianum	30	1.36		B	C
G. hirsutum	38	1		B	C
G. australe	38	0.92		B	C
G. bickii	38	0.94		B	C
G. robinsonii	38	1.06			C
G. sturtianum	38	1.03		B	C

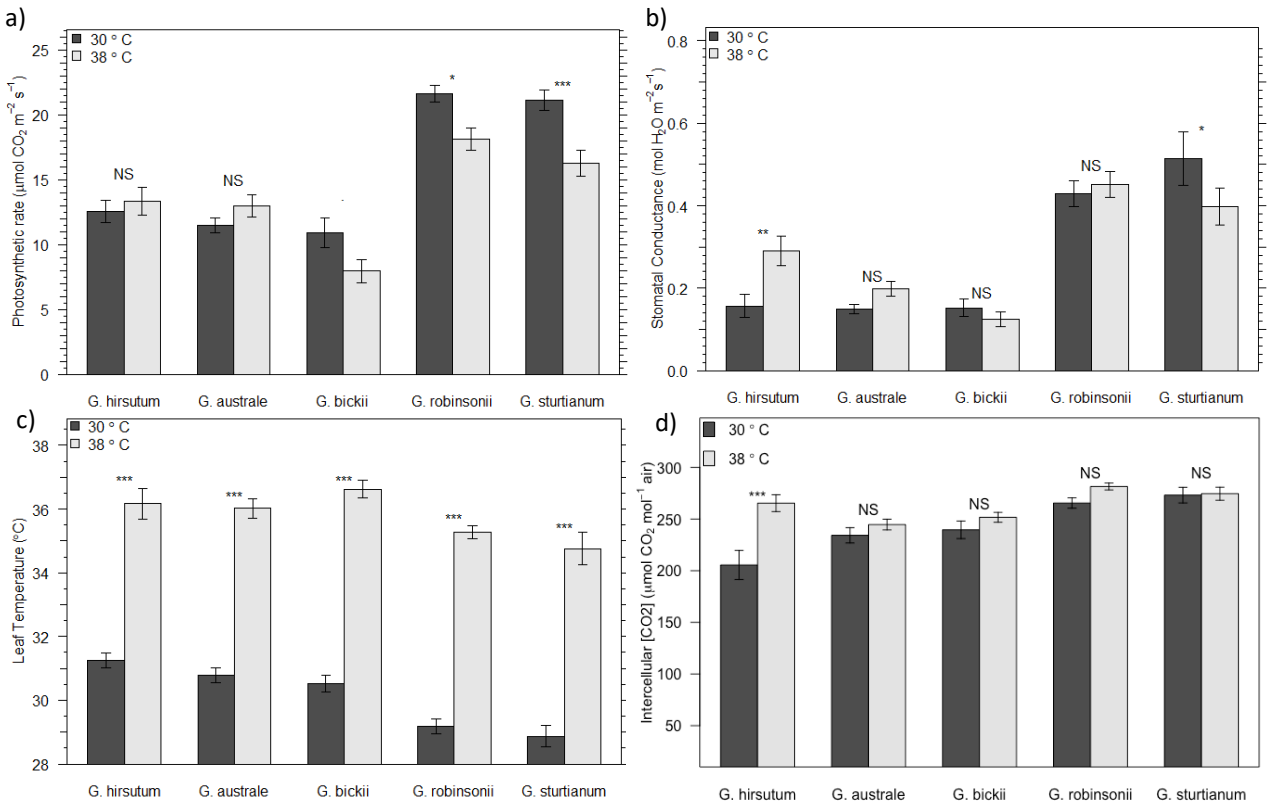


Figure 4. Mean (± SE, n = 6) gas exchange parameters of control and heat-treated leaves of five *Gossypium* species that occupy different thermal niches: (a) net photosynthetic rate, (b) stomatal conductance, (c) leaf temperature (note: scale starts at 28°C), and (d) intercellular [CO_2]. Plants were exposed to control conditions (30°C) for five days, then exposed to high temperature (38°C) for five days and re-measured. The significance of differences between the temperature treatments for each species were determined by individual z-tests: '.', p<0.1; '*', p<0.05; '**', p<0.01; '***', p<0.001.

4.2.2 Nicotiana: species by temperature interaction

Photosynthesis

At control temperatures (30°C) there were no significant differences in the photosynthetic rates of the five *Nicotiana* species (Table 9; Fig. 5a). The photosynthetic rates of *N. tabacum* and *N.*

benthamiana decreased significantly at 38°C, relative to rates at control temperatures (z = 4.4, p <0.001; z = 4.7, p < 0.001). However, there was a significant increase in the photosynthetic rate of *N. africana* at 38°C, relative to control temperatures (z = -6.5, p < 0.001). These changes meant that at 38°C the photosynthetic rates of *N. africana* and *N. megalosiphon* were significantly greater than those of *N. tabacum, N. benthamiana* and *N. gossei* (z = -7.5, p < 0.001; z = 8.9, p < 0.001; z = 6.7, p < 0.001, respectively for *N. africana*; z = -4.9, p < 0.001; z = 6.3, p < 0.001; z = -4.1, p < 0.01, respectively for *N. megalosiphon*). Additionally, at 38°C there were no significant differences between the photosynthetic rates of *N. tabacum, N. benthamiana* and *N. gossei*, nor between *N. africana* and *N. megalosiphon*.

Table 9. Grouping information for mean (*n* = 6) photosynthetic rates (A_n) of the focal *Nicotiana* species at control (30°C) and high (38°C) temperatures, expressed relative to *N. tabacum* A_n values. Groups that do not share a letter are significantly different.

Species	Temperature (°C)	$A_n/A_{n,NT}$	Grouping				
N. tabacum	30	1			C	D	
N. africana	30	1.19			C	D	
N. benthamiana	30	0.77		B	C		
N. gossei	30	0.98		B	C	D	
N. megalosiphon	30	1.2			C	D	
N. tabacum	38	1	A	B			
N. africana	38	3.49					E
N. benthamiana	38	0.51	A				
N. gossei	38	1.32	A		C		
N. megalosiphon	38	2.71				D	E

Stomatal conductance

At control temperatures (30°C), there were no significant differences in stomatal conductance of the five *Nicotiana* species (Table 10; Fig. 5b). There was no significant change in stomatal conductance at 38°C for *N. tabacum, N. benthamiana, N. gossei* and *N. megalosiphon* relative to their controls, though there was a significant increase seen in *N. africana* (z = -6.4, p < 0.001). At 38°C stomatal conductance in *N. africana* was significantly higher than in *N. tabacum*, *N. benthamiana* and *N. gossei* (z = -5.3, p < 0.001; z = 6.3, p < 0.001; z = 4.9, p < 0.001, respectively), though there was no significant difference to *N. megalosiphon*.

Table 10. Grouping information for mean (*n* = 6) stomatal conductance rates (g_s) of the focal *Nicotiana* species at control (30°C) and high (38°C) temperatures, expressed relative to *N. tabacum* g_s values. Groups that do not share a letter are significantly different.

Species	Temperature (°C)	$g_s/g_{s,NT}$	Grouping		
N. tabacum	30	1	A	B	
N. africana	30	1.14	A	B	
N. benthamiana	30	0.52	A	B	
N. gossei	30	0.95	A	B	
N. megalosiphon	30	1.1	A	B	
N. tabacum	38	1	A	B	
N. africana	38	4.36			C
N. benthamiana	38	0.36	A		
N. gossei	38	1.36	A	B	
N. megalosiphon	38	2.73		B	C

Leaf temperature

There were no significant differences in leaf temperature between the *Nicotiana* species at 30°C (Table 11; Fig. 5c). Leaf temperature of all species increased significantly at 38°C, however the leaves of *N. africana* were significantly cooler than those of *N. tabacum, N. benthamiana* and *N. gossei* ($z = 3.6$, $p < 0.01$; $z = -4.1$, $p < 0.01$; $z = -3.3$, $p = 0.02$, respectively), though there was no significant difference between *N. africana* and *N. megalosiphon*.

Table 11. Grouping information for mean (*n* = 6) leaf temperatures (L_T) of the focal *Nicotiana* species at control (30°C) and high (38°C) temperatures, expressed relative to *N. tabacum* L_T values. Groups that do not share a letter are significantly different.

Species	Temperature (°C)	$L_T/L_{T,NT}$	Grouping		
N. tabacum	30	1	A		
N. africana	30	0.99	A		
N. benthamiana	30	1.03	A		
N. gossei	30	1.01	A		
N. megalosiphon	30	0.99	A		
N. tabacum	38	1			C
N. africana	38	0.94		B	
N. benthamiana	38	1.01			C
N. gossei	38	0.98			C
N. megalosiphon	38	0.97		B	C

Intercellular [CO_2]

At 30°C, *N. tabacum* intercellular [CO_2] was significantly higher than *N. benthamiana* ($z = 3.1$, $p = 0.03$; Table 12). At 38°C *N. benthamiana* intercellular [CO_2] was significantly lower than *N. africana* ($z = -3.4$, $p = 0.018$), though all other species had similar internal [CO_2].

Table 12. Grouping information for mean (n = 6) intercellular [CO_2] (C_i) of the focal *Nicotiana* species at control (30°C) and high (38°C) temperatures, expressed relative to *N. tabacum* C_i values. Groups that do not share a letter are significantly different.

Species	Temperature (°C)	$C_i/C_{i,NT}$	Grouping		
N. tabacum	30	1			C
N. africana	30	0.98		B	C
N. benthamiana	30	0.76	A	B	
N. gossei	30	0.91	A		C
N. megalosiphon	30	0.88	A		C
N. tabacum	38	1	A		C
N. africana	38	1.08			C
N. benthamiana	38	0.8	A		
N. gossei	38	0.91	A		C
N. megalosiphon	38	0.95	A		C

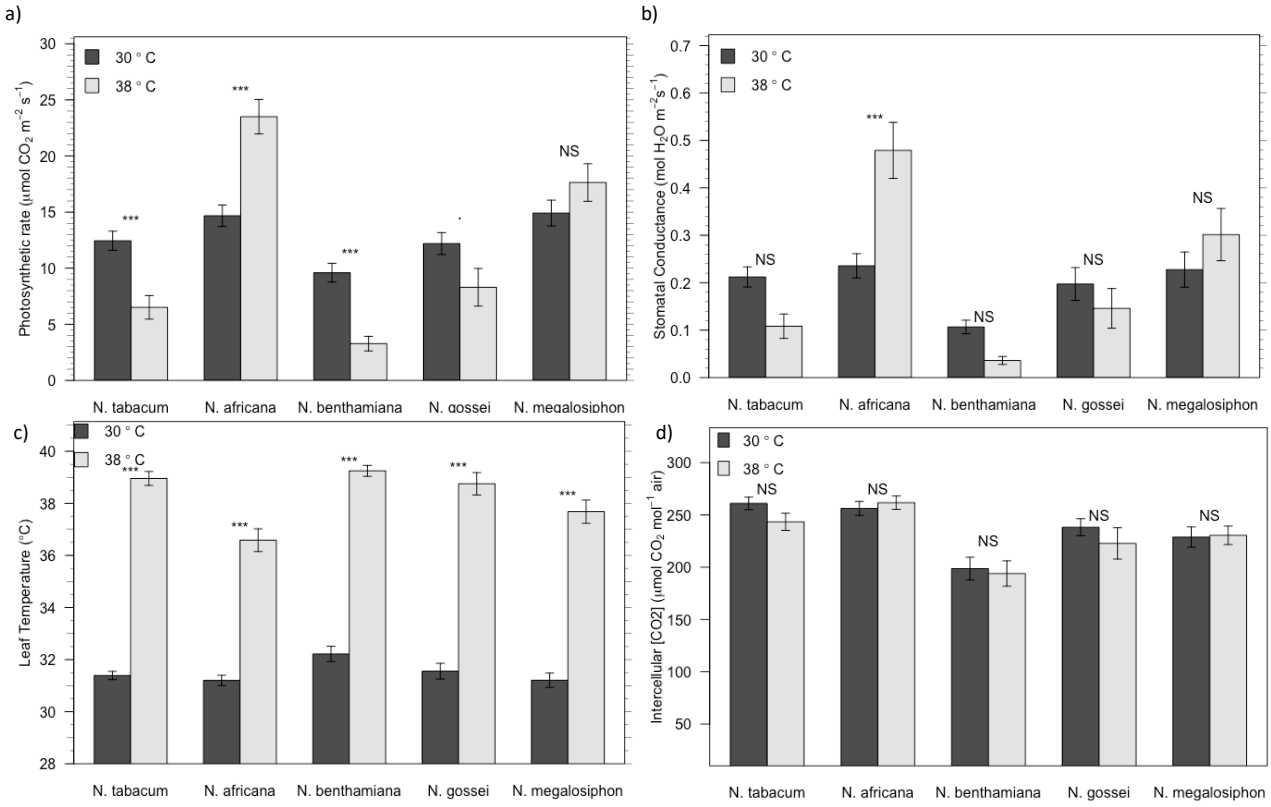


Figure 5. Mean (± SE, n = 6) gas exchange parameters of control and heat-treated leaves of five *Nicotiana* species that occupy different thermal niches: (a) net photosynthetic rate, (b) stomatal conductance, (c) leaf temperature (note: scale starts at 28°C), and (d) intercellular [CO_2]. Plants were exposed to control conditions (30°C) for five days, then exposed to high temperature (38°C) for five days and re-measured. The significance of differences between the temperature treatments for each species were determined by individual z-tests: '.', p<0.1; '*', p<0.05; '**', p<0.01; '***', p<0.001.

4.3 Effect of Heat on the Abundance of Rubisco Activase in Leaf Tissue

4.3.1 Absolute RCA Abundance

For each genus, mixed effect models were fit for total abundance of RCA and the ratio of total RCA (µmol RCA m^{-2} leaf) to the Rubisco large subunit (µmol RbcL m^{-2} leaf). For *Gossypium* and

Nicotiana, models were fitted with and without interaction terms and y-intercepts and goodness-of-fit was evaluated using log likelihood ratio tests. For *Themeda triandra* (Sydney population), a single model was fit as no interaction was possible within this dataset.

Total RCA Pool Abundance

For both *Gossypium* and *Nicotiana*, there was no significant difference in the goodness of fit of models that did and did not include an interaction between species and temperature (*Gossypium*: $\chi^2_1 = 2.2$, p = 0.14; *Nicotiana*: $\chi^2_1 = 2.2$, p = 0.14). This implies that the effect of the heat treatment on the absolute abundance of the total RCA pool was the same for both the domestic and wild species of *Gossypium* and *Nicotiana* (Fig. 6a-b). This was confirmed in *post-hoc* tests of biologically interesting means.

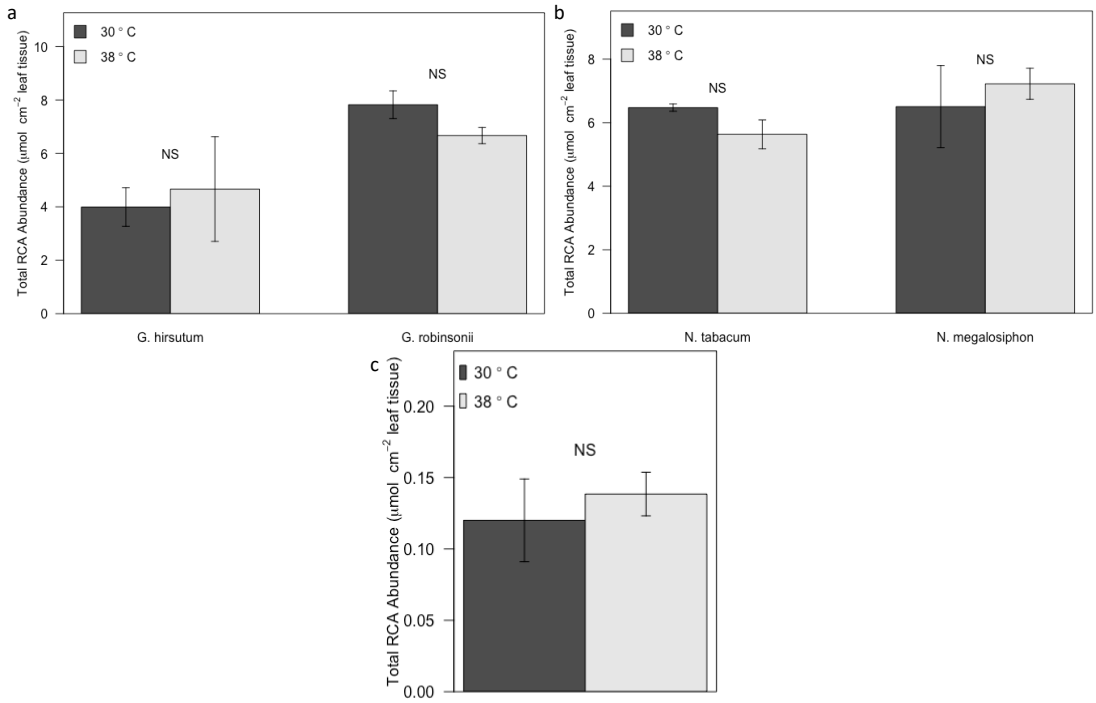


Figure 6. Mean (± SE, n = 3) concentration of the total RCA pool in (a) *G. hirsutum* and *G. robinsonii*, (b) *N. tabacum* and *N. megalosiphon*, and (c) the Sydney population of *T. triandra* at control (30°C) and treatment (38°C) temperatures. Absolute abundance values were obtained by spiking highly conserved isotope labelled RCA and ovalbumin (internal standard) peptides into each sample. 'NS' indicates a p-value >0.05.

At 30°C *G. robinsonii* had significantly more RCA than *G. hirsutum* (z = -2.5, p = 0.046), though there was no significant difference in RCA abundance at 38°C, owing largely to an increase in the variability of measurements in *G. hirsutum* (z = -1.3, p = 0.51). There was no significant difference in the abundance of RCA in *N. tabacum* and *N. megalosiphon* at 30°C (z = -0.12, p = 0.999) or at 38°C (z = -1.6, p = 0.34). Finally, there was no effect of heat on the abundance of the total RCA

pool in *G. hirsutum*, *G. robinsonii*, *N. tabacum*, *N. megalosiphon*, nor in the Sydney population of *T. triandra* (post-hoc tests: z = -0.67, p = 0.88; z = 1.18, p = 0.6, z = 0.9, p = 0.77; z = -0.92, p = 0.76; One-way ANOVA: $F_{1,4} = 0.45$, p = 0.57, respectively).

RCA/Rubisco Ratio

For both *Gossypium* and *Nicotiana*, there was no significant difference in the goodness of fit of models that did and did not include an interaction between species and temperature (*Gossypium*: $\chi^2_1 = 2.972$, p = 0.085; *Nicotiana*: $\chi^2_1 = 0.626$, p = 0.429). This implies that the effect of the heat treatment on the ratio of RCA and RbcL was the same for the wild and domestic *Gossypium* and *Nicotiana* species (Fig. 7a-b). However, a *post-hoc* comparison of biologically interesting means in the *Gossypium* dataset (i.e. GH 30°C vs GH 38°C, GR 30°C vs GR 38°C, GH 30°C vs GR 30°C, GH 38°C vs GR 38°C) shows that RCA/RbcL increases marginally significantly in *G. robinsonii* at high temperature (z = -2.474, p = 0.047).

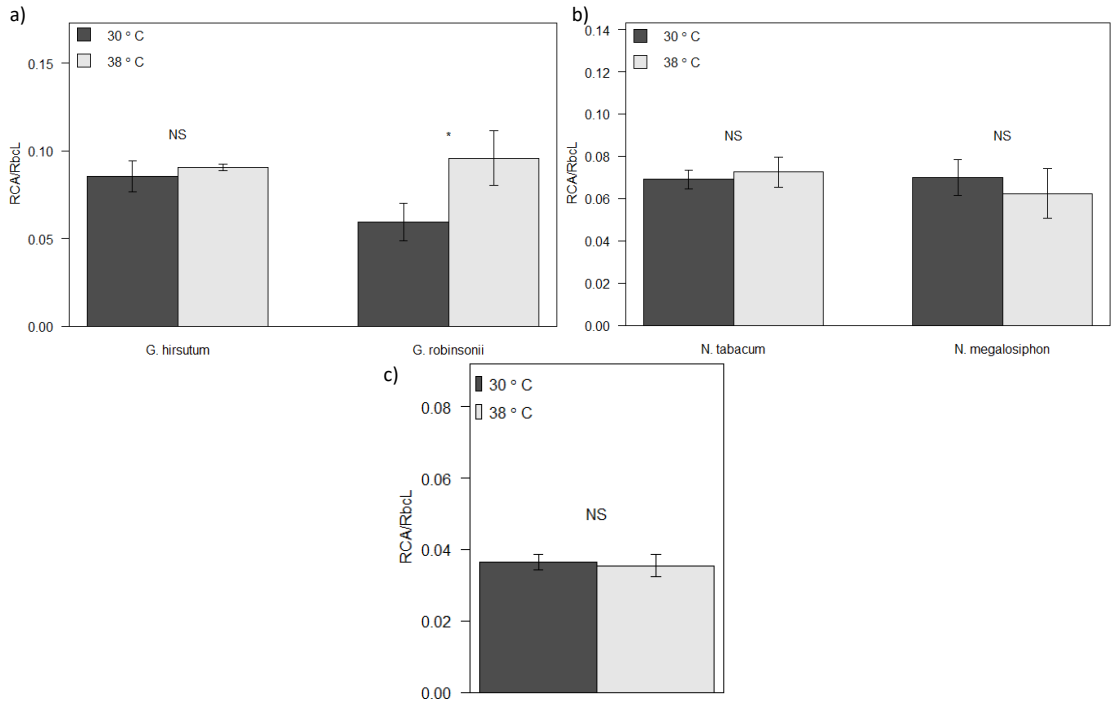


Figure 7. Mean (± SE, n = 3) RCA to RbcL ratios of domestic and wild (a) cotton and (b) tobacco species, and (c) *Themeda triandra* (Sydney population) leaves at control (30°C) and treatment (38°C) temperatures. Ratios (*n* = 3) were calculated from absolute RCA and RbcL abundance data each in units of µmol protein m^{-2} leaf tissue. Plants were exposed to control conditions (30°C) for five days, then exposed to high temperature (38°C) for five days. Total leaf protein was extracted and the abundance of RCA and RbcL determined through the use of heavy-isotope labelled peptides. The significance of differences between the temperature treatments for each species were determined by individual z-tests: 'NS', p>0.1; '.', p<0.1; '*', p<0.05; '**', p<0.01; '***', p<0.001.

There was no effect of heat on the RCA-RbcL ratio in *G. hirsutum*, *N. tabacum*, *N. megalosiphon* or *T. triandra* (Fig. 7a-c; post-hoc tests: z = -0.354, p = 0.98; z = -0.291, p = 0.989; z = 0.634, p = 0.9; One-way ANOVA: $F_{1,16}$ = -0.279, p = 0.794, respectively), and there was no overall effect of species on the RCA-RbcL ratio in *Gossypium* or *Nicotiana* (F_1 = 1.001, p = 0.374; F_1 = 0.292, p = 0.616, respectively). At 38^{o}C there was no significant difference between RCA/RbcL ratios in *G. hirsutum* and *G. robinsonii* or between *N. tabacum* and *N. megalosiphon* (z = -0.353, p = 0.98; z = 0.858, p = 0.786).

4.3.2 Relative Abundance of Identified RCA Isoforms

Following initial SWATH analyses of the samples, unique peptides within the RCA sequences of *Nicotiana*, *Gossypium* and *Themeda* species were identified. ProteinPilot ascribed the unique target peptides of both *Gossypium* and *Nicotiana* to four different isoforms of RCA in their respective databases and identified mutant peptides, while those of *Themeda* were ascribed to a single isoform of RCA from the Sorghum-based database (Fig. 8a-c; Supplementary Materials Table 3 section 7.1.3). Mutant peptides are those which did not occur in any of the RCA sequences in the constructed databases but which were nonetheless detected during mass spectrometry. These isoforms will be referred to below.

There was no significant effect of heat on the relative abundance of *Themeda* RCA Isoform 1 in the Sydney population of *T. triandra* (Fig. 9; t_8 = 0.5, p = 0.32).

No significant change was seen in the relative abundance of *Gossypium* RCA Isoform 1 or its mutant form in either *G. hirsutum* or *G. robinsonii* at control and treatment temperatures (Fig. 10a-b; Isoform 1: t_8 = -1.24, p = 0.12; t_8 = 1.4, p = 0.1, respectively; Isoform 1 Mutant: t_2 = -1.1, p = 0.2; t_2 = 1.7, p = 0.12, respectively). No significant change in Isoform 2 was detected in *G. hirsutum* at high temperature (Fig. 9c; t_5 = -0.3, p = 0.38) though there was a significant decline in the relative abundance of this isoform in *G. robinsonii* at high temperature (t_5 = 2.41, p = 0.03). There was no effect of heat on the relative abundance of the Isoform 2 mutant in *G. hirsutum* (Fig. 9d; t_2 = 0.08, p = 0.5) or *G. robinsonii*, though there appears to be a trend in *G. robinsonii* of decreasing relative abundance at high temperature (t_2 = 2.31, p = 0.074). *Gossypium* RCA Isoform 3 and its mutant form did not respond to high temperature in both *G. hirsutum* and *G. robinsonii* (Fig. 10e-f; Isoform 3: t_8 = -1.2, p = 0.14; t_8 = -0.5, p = 0.32, respectively; Isoform 3 Mutant: t_2 = -1, p = 0.2; t_2 = 1.3, p = 0.17, respectively). There was no effect of heat on the relative abundance of RCA Isoform

4 in *G. hirsutum* at high temperature (Fig. 10g; $t_8 = -0.6$, $p = 0.28$), though there was a significant decrease in *G. robinsonii* ($t_8 = 2.8$, $p = 0.012$).

There was no significant effect of heat on the relative abundance of *Nicotiana* Isoform 1 in *N. tabacum* or *N. megalosiphon* (Fig. 11a; $t_5 = 0.3$, $p = 0.4$; $t_5 = -0.97$, $p = 0.2$). The relative abundance of the Isoform 1 mutant did not change in response to heat in *N. tabacum* (Fig. 11b; $t_2 = -0.3$, $p = 0.4$), though there was a significant increase in *N. megalosiphon* at high temperature ($t_2 = -6.4$, $p = 0.01$). While there was no significant effect of heat on the relative abundance of RCA Isoform 2 in either *N. tabacum* or *N. megalosiphon* (Fig. 11c; $t_5 = 1.5$, $p = 0.09$; $t_5 = -1.7$, $p = 0.07$) there appears to be a trend of decreasing relative abundance in *N. tabacum* at high temperature but an increase in *N. megalosiphon*. There was a significant increase in the relative abundance of *Nicotiana* RCA Isoform 3 at high temperature in *N. tabacum* and *N. megalosiphon* (Fig. 11d; $t_5 = -2.1$, $p = 0.05$; $t_5 = -2.1$, $p = 0.046$). There was no significant effect of heat on the relative abundance of Isoform 4 in either *N. tabacum* or *N. megalosiphon* (Fig. 11e; $t_5 = -0.07$, $p = 0.47$; $t_5 = 1.61$, $p = 0.08$), though there may be a trend of decreasing abundance in *N. megalosiphon*.

Figure 8. Diagrammatic representation of (a) the four identified *Gossypium* RCA isoforms, (b) the four identified *Nicotiana* RCA isoforms, and (c) the single RCA isoform identified for *Themeda*. Isoforms were identified for each genus in their respective databases constructed for the MRM-HR pipeline. Each identified RCA isoform is numbered for its residues. Grey boxes indicate highly conserved peptides used for the absolute quantification of RCA. Blue boxes indicate peptides unique to the corresponding RCA isoform within a genus. Peptides in red text indicate those containing at least one residue change (mutation). Refer to Supplementary Table 3 for details on each RCA isoform. Note: the position of each peptide in each isoform is approximate and not to scale.

Peptide Fold-Change at High Temperature - Evidence of Alternative Splice Variants

See Supplementary Material Table 3 (7.1.3) for details of the RCA isoforms and their unique peptides. See also Figure 8 for a diagrammatic representation of the identified RCA isoforms and the position of each targeted peptide discussed below.

There was no significant difference in the fold changes of the peptides ascribed to *Themeda* RCA Isoform 1 (Supplementary Material section 7.2.1; $F_{2,6} = 0.05$, $p = 0.95$). This implies there is no evidence that this isoform is the product of alternative splicing at 38ºC.

There was no significant difference between the fold changes of the peptides ascribed to *Gossypium* RCA Isoform 1 (including its mutant form), Isoform 2 (including its mutant form), Isoform 3 (including its mutant form) or Isoform 4 in *G. hirsutum* or *G. robinsonii* (Supplementary Materials section 7.2.2). This indicates that there is no evidence that these isoforms are the products of alternative splicing in these species at 38°C.

There was no significant difference between the fold changes of the peptides ascribed to *Nicotiana* RCA Isoform 1 (including its mutant form), Isoform 2, Isoform 3, or Isoform 4 in *N. tabacum* (Supplementary Materials section 7.2.3). This indicates that there is no evidence that these isoforms are the products of alternative splicing at 38°C. For *N. megalosiphon* there was a significant difference in the fold change of the peptides ascribed to Isoform 3 ($F_{1,4}$ = 42.905, p = 0.003). These results indicate that there may be evidence of alternative splicing of Isoform 3 at 38°C.

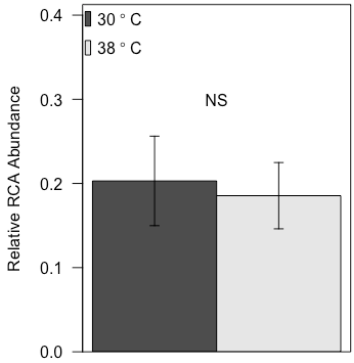


Figure 9. Mean (± SE, n = 3) relative abundance of *Themeda* RCA Isoform 1 at control (30°C) and treatment (38°C) temperatures in the Sydney population of *T. triandra*. Three peptides were used for relative quantification. Summed peak intensities were corrected using summed peak intensities of the ovalbumin-spike peptides. 'NS' = non-significant.

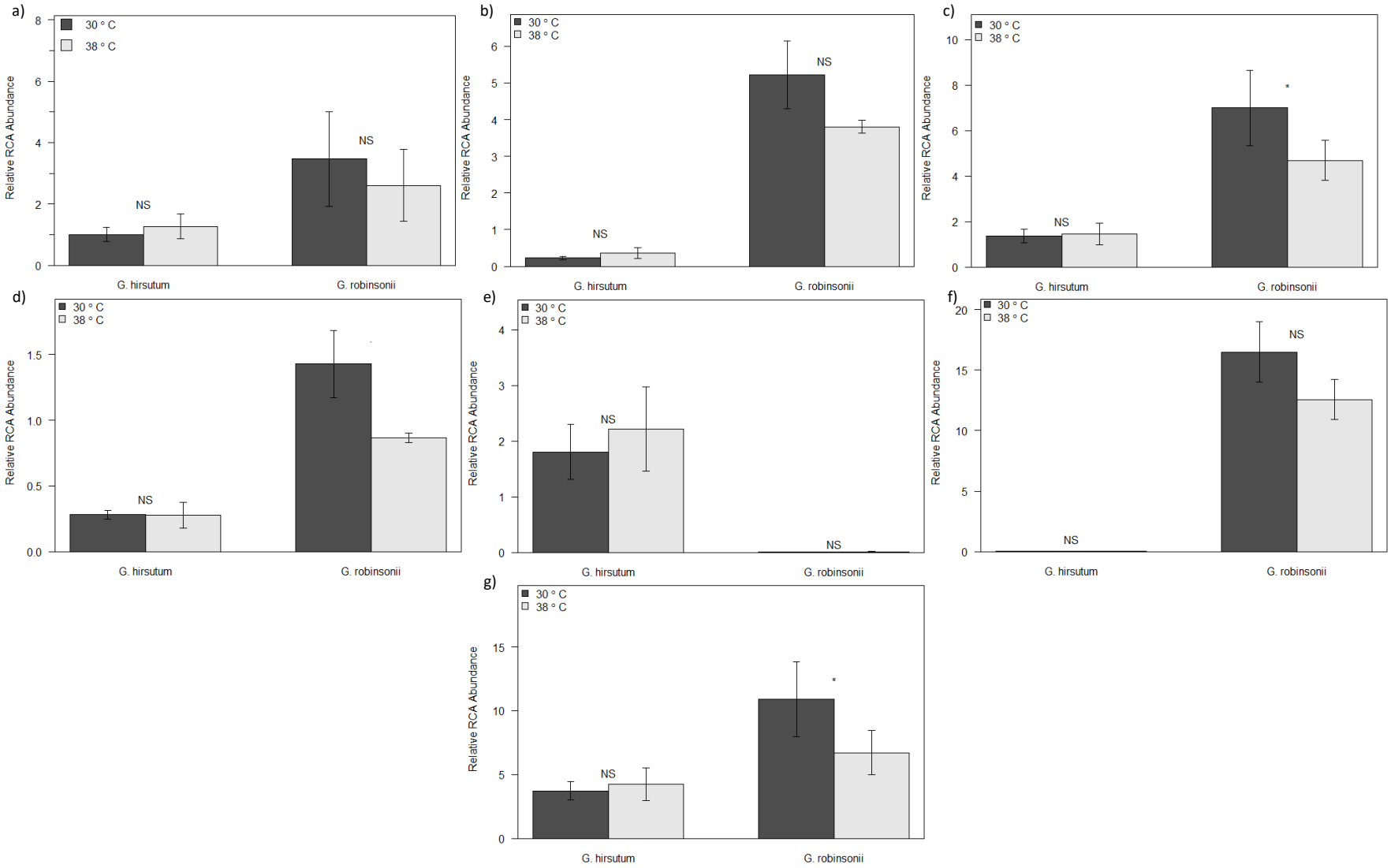


Figure 10. Mean (± SE, n = 3) relative abundance of *Gossypium* RCA (a) Isoform 1, (b) Isoform 1 mutant, (c) Isoform 2, (d) Isoform 2 mutant, (e) Isoform 3, (f) Isoform 3 mutant, and (g) Isoform 4 in *G. hirsutum* and *G. robinsonii* at control (30°C) and treatment (38°C) temperatures. The number of peptides used varies across isoforms (see Figure 11). Summed peak intensities for each peptide for each isoform were corrected using the summed peak intensities of the peptides comprising the spiked ovalbumin standard. Individual t-tests: 'NS', p>0.1; '.', p<0.1; '*', p<0.05; '**', p<0.01; '***', p<0.001. Note the different scales in the *y*-axes.

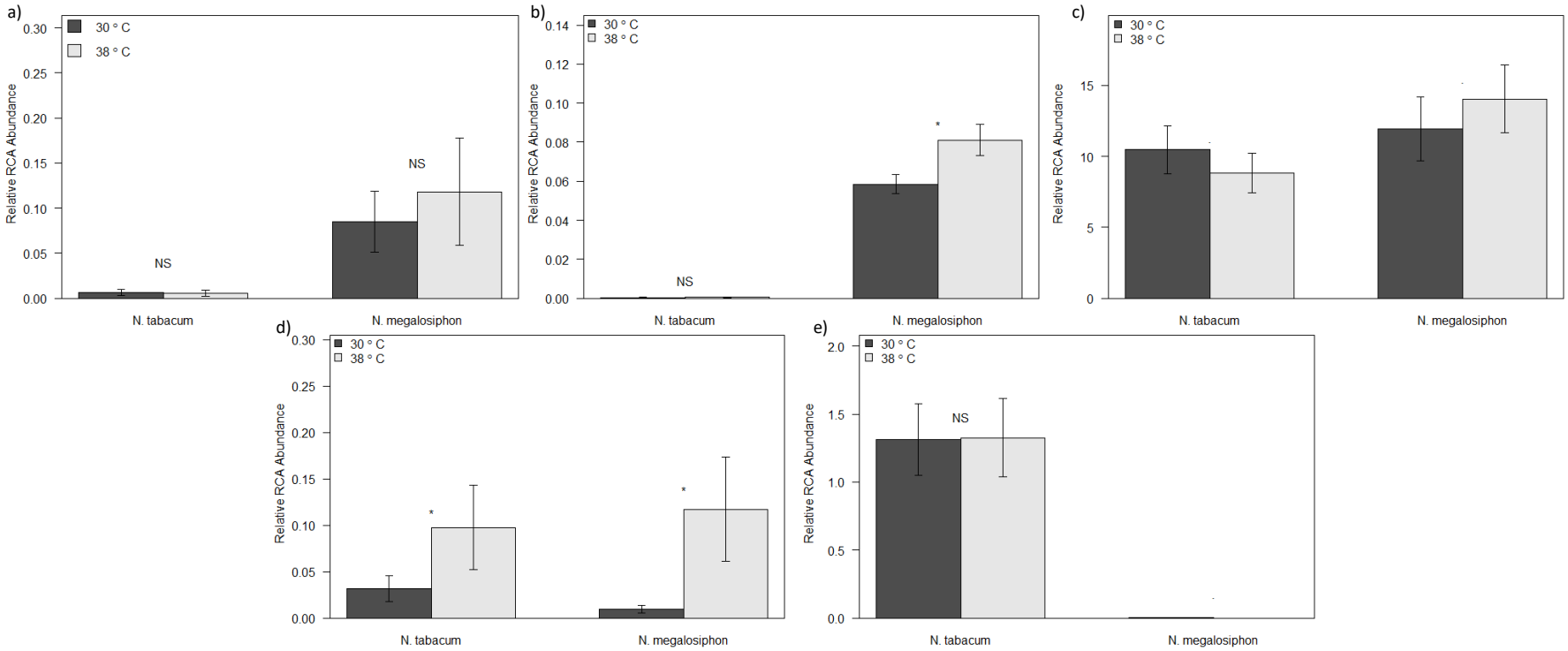


Figure 11. Mean (± SE, n = 3) relative abundance of *Nicotiana* RCA (a) Isoform 1, (b) Isoform 1 mutant, (c) Isoform 2, (d) Isoform 3, and (e) Isoform 4 in *N. tabacum* and *N. megalosiphon* at control (30°C) and treatment (38°C) temperatures. The number of peptides used varies across isoforms (see Figure 11). Summed peak intensities for each peptide for each isoform were corrected using the summed peak intensities of the peptides comprising the spiked ovalbumin standard. Individual t-tests: 'NS', p>0.1; '.', p<0.1; '*', p<0.05; '**', p<0.01; '***', p<0.001. Note the different scales in the *y*-axes.

4.4 RCA Sequences

Primers designed against the RCA 1 and RCA 2-specific Malvaceae, Solanaceae, and Panicoid consensus mRNA sequences enabled amplification and sequencing of these isoforms synthesised from leaf-extracted mRNA of *G. hirsutum*, *G.* robinsonii, *G. bickii, N. tabacum*, *N. africana*, *N. benthamiana*, *N. gossei*, *N. megalosiphon*, and a single RCA isoform from three populations of *T. triandra*. Although these sequences were variably truncated, aligning them with previously published *G. hirsutum* or *N. tabacum* RCA 1 and 2 isoform sequences, or RCA 1 from *Zea mays* for *Themeda*, respectively, highlighted SNPs (Supplementary Material sections 7.2.4-7.2.8). Translation of the DNA sequences revealed differences in the primary sequence of RCA 1 and RCA 2 between the species (Fig. 12-14; Tables 13-16).
There was no successful amplification of ADH cDNA in any sample. This gene was therefore excluded from further analyses.

4.4.1 Gossypium Sequences

RCA 1 was sequenced in *G. hirsutum*, *G. bickii,* and *G. robinsonii* (Supplementary Material 7.2.4). Issues with high amounts of polyphenolic compounds in the leaf material appeared to interfere with RNA extraction from *G. australe* and *G. sturtianum* despite the use of a kit designed to remove polyphenols from leaf material or the addition of PVPP to a standard RNA extraction kit. There were seven recorded single nucleotide polymorphisms (SNPs) between the reference *G. hirsutum* sequence and commercial *G. hirsutum* cv. Siokra used in the present study. In *G. bickii*, there were 14 SNPs relative to the reference sequence, while only six SNPs were detected in RCA 1 from *G. robinsonii*.

Of the seven SNPs detected in *G. hirsutum* cv. Siokra RCA 1, only two resulted in residue changes in the primary peptide sequence relative to the reference sequence (Fig. 12a; Table 13). For *G. bickii* RCA 1, there were five residue changes in the primary peptide sequence. Two residue changes were detected in *G. robinsonii* RCA 1.

Table 13. Location of residue changes in the functional domains of *Gossypium* RCA 1 peptide sequences relative to a previously published RCA 1 sequence (AF329934).

Species	N-terminal Domain	N-terminal AAA+ module	ATPase Core	C-terminal AAA+ Module	C-terminal Domain
G. hirsutum	1	-	-	1	-
G. bickii	-	3	-	-	2
G. robinsonii	-	2	-	-	-

RCA 2 was sequenced in *G. hirsutum* and *G. bickii* (Supplementary Material 7.2.5). Despite a re-design of RCA-2-specific primers following this sequencing event, no amplification could be made of RCA 2 from *G. robinsonii*. There was a single SNP recorded between the reference RCA 2 *G. hirsutum* sequence and the *G. hirsutum* cv. Siokra from the present study. There were three nucleotide insertions at positions 8, 876, and 877 in *G. bickii* RCA 2, along with 24 SNPs relative to the reference sequence.

The single SNP detected in *G. hirsutum* RCA 2 did not result in residue changes in the primary peptide sequence (Fig. 12b; Table 14). In contrast, *G. bickii* RCA 2 had 18 residue changes relative to the reference sequence, including an insertion.

Table 14. Location of residue changes in the functional domains of *Gossypium* RCA 2 peptide sequences relative to a previously published RCA 2 sequence (NM_001327460).

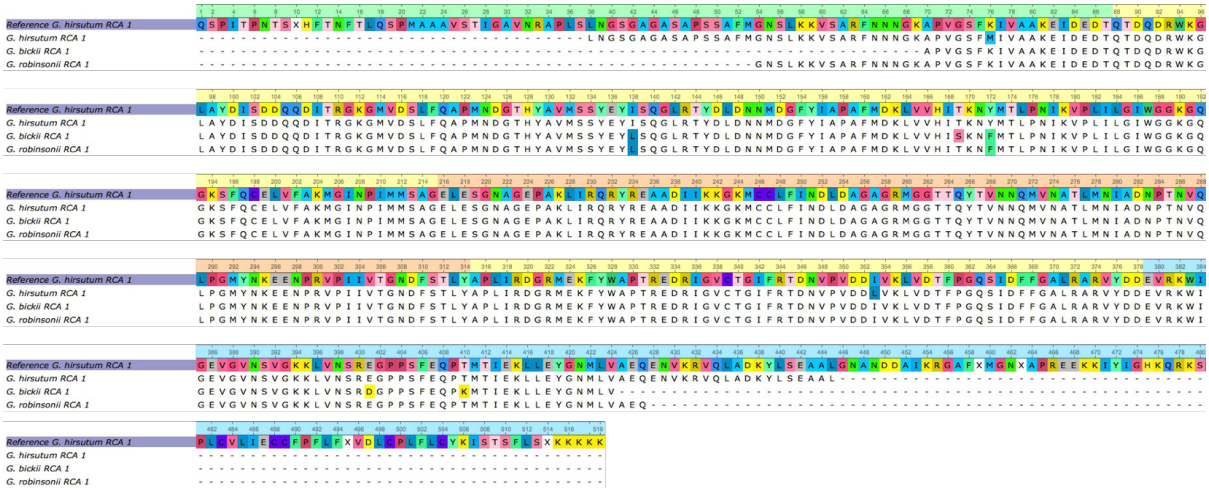
Species	N-terminal Domain	N-terminal AAA+ module	ATPase Core	C-terminal AAA+ Module	C-terminal Domain
G. hirsutum	-	-	-	-	-
G. bickii	-	5	4	5	4

4.2.2 Nicotiana Sequences

RCA 1 was sequenced in all focal *Nicotiana* species (Supplementary Material 7.2.6). There were 25 SNPs between the reference *N. tabacum* sequence and the *N. tabacum* from the present study, resulting in three residue changes (Fig. 13a; Table 15). In *N. africana*, there were 36 SNPs relative to the reference sequence which resulted in seven residue changes in the primary peptide sequence. There were 34 SNPs were detected in RCA 1 from *N. benthamiana*, which translated to five residue changes. In *N. gossei* 60 SNPs were detected relative to the reference *N. tabacum* sequence, 15 of which resulted in residue changes.

Finally, there were 39 SNPs detected in RCA 1 from *N. megalosiphon*, which resulted in six residue changes in the primary peptide sequence of RCA 1.

(a)



(b)

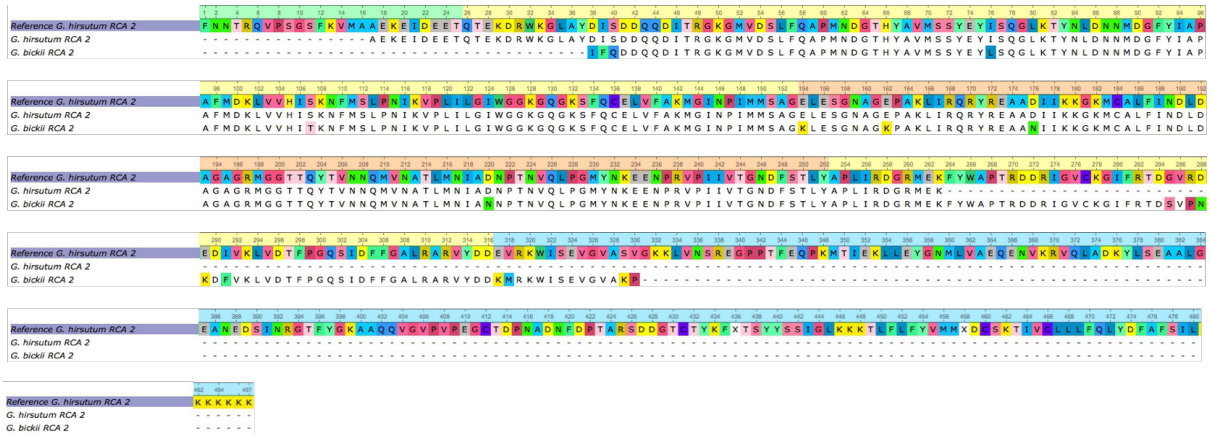


Figure 12. Peptide sequences of (a) RCA 1 from *G. hirsutum*, *G. bickii*, and *G. robinsonii* and (b) RCA 2 from *G. hirsutum* and *G. bickii*. Sequences generated during the present study are presented alongside previously published *G. hirsutum* RCA 1 (AF329934) and RCA 2 (NM_001327460), which served as the reference sequences for alignment and as indicators of sequencing coverage. Residue numbers above the peptide sequences are coloured according to RCA functional domains, based on the crystal structure of tobacco RCA (Stotz *et al.*, 2011): green = N-terminus domain, yellow = AAA^+ module, orange = ATPase core, and blue = C-terminus domain.

Table 15. Location of residue changes in the functional domains of *Nicotiana* RCA 1 peptide sequences relative to a previously published RCA 1 sequence (NM_001326055.1).

Species	N-terminal Domain	N-terminal AAA+ module	ATPase Core	C-terminal AAA+ Module	C-terminal Domain
N. tabacum	1	-	-	1	1
N. africana	3	1	-	2	1
N. benthamiana	4	1	-	-	-
N. gossei	7	2	-	2	4
N. megalosiphon	3	1	-	1	-

RCA 2 was sequenced in all *Nicotiana* species (Supplementary Material 7.2.7). There were 51 SNPs detected between the reference *N. tabacum* RCA 2 sequence and the *N. tabacum* line of the present study, resulting in 13 residue changes (Fig. 13b; Table 16). In *N. africana* RCA 2, there were 18 SNPs relative to the reference sequence. These SNPs resulted in six residue changes in the primary peptide sequence. There were 19 recorded SNPs in RCA 2 from *N. benthamiana*, which translated to five residue changes. In *N. gossei* RCA 2, there were 16 SNPs relative to the reference sequence, resulting in six residue changes in the peptide sequence. Finally, there were 70 recorded SNPs in RCA 2 from *N. megalosiphon*, resulting in 19 residue changes.

Table 16. Location of residue changes in the functional domains of *Nicotiana* RCA 2 peptide sequences relative to a previously published RCA 2 sequence (Z14980).

Species	N-terminal Domain	N-terminal AAA+ module	ATPase Core	C-terminal AAA+ Module	C-terminal Domain
N. tabacum	1	6	1	5	-
N. africana	1	4	1	-	-
N. benthamiana	-	2	2	-	1
N. gossei	1	3	2	-	-
N. megalosiphon	4	7	1	5	2

4.2.3 Themeda Sequences

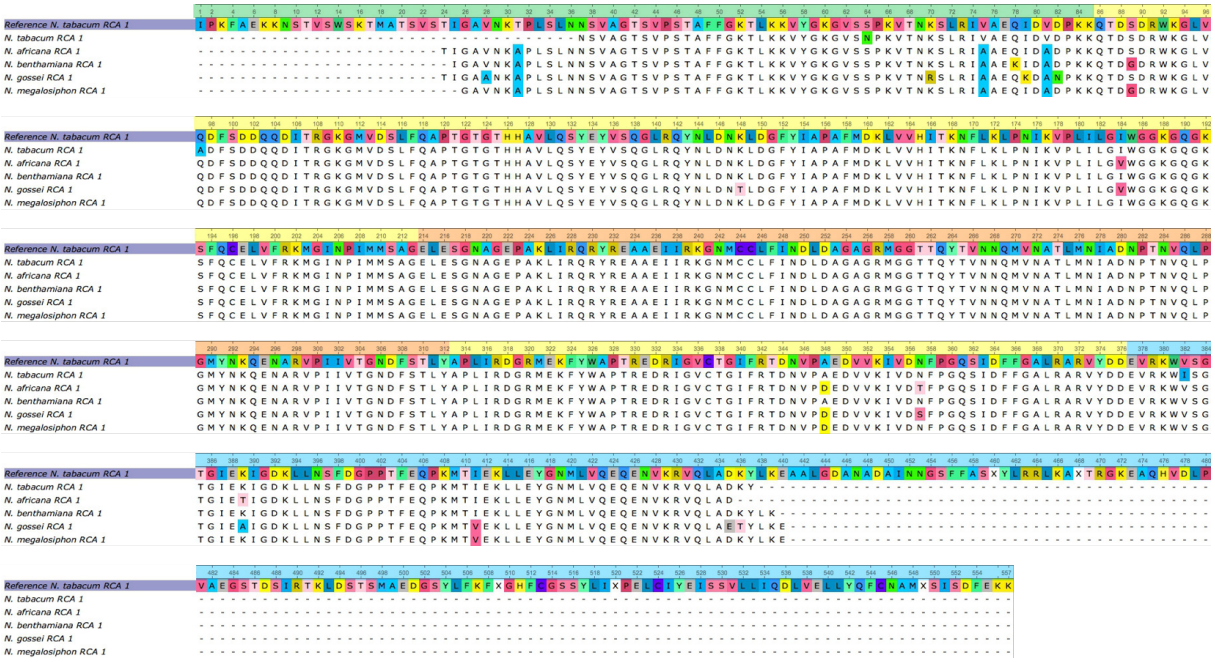
Here, the Sydney population RCA acts as the reference sequence, though the previously published sequence of RCA 1 from *Zea mays* was included in the alignment to serve as a representation of RCA from a relatively heat-adapted Panicoid species. The sequences presented here for *Themeda* are truncated relative to the *Z. mays* RCA 1 sequence.

There were no SNPs detected between the Sydney *T. triandra* population and the Tasmanian population (Fig. 14; Supplementary Material 7.2.8). However, there were 14 SNPs recorded between the Sydney population and the Northern Territory population, resulting in one residue change in the primary peptide sequence. Within the *Themeda* genus, these SNPs were unique to the Northern Territory population. There were 57 SNPs detected between the Sydney population RCA and RCA 1 from *Z. mays*. Interestingly, 10 (~71%) of the SNPs in *T. triandra* from the Northern Territory population were shared with RCA 1 from *Z. mays*, while four (29%) are unique to itself.

4.2.4 Double Base Calls

A number of ambiguous base calls were generated during the sequencing reactions. Most of these base calls were made on both the forward and reverse strands, reducing the probability that they were erroneous. A total of 668 unique double base calls were manually detected in the Sanger chromatograms of all species and genes considered in this study. Parsing the chromatograms with the R packages sangerseqR and seqinR revealed that 510 of these double base calls were above the 0.33 threshold, indicating that the signal intensities of these secondary base calls were significant. Following further manual scrutiny of the 510 valid double base calls, it became apparent that if the double base calls were real observations they could result in a total of 118 residue changes across all RCA isoforms and species (Supplementary Material Table 4 section 7.2.9).

(a)



(b)

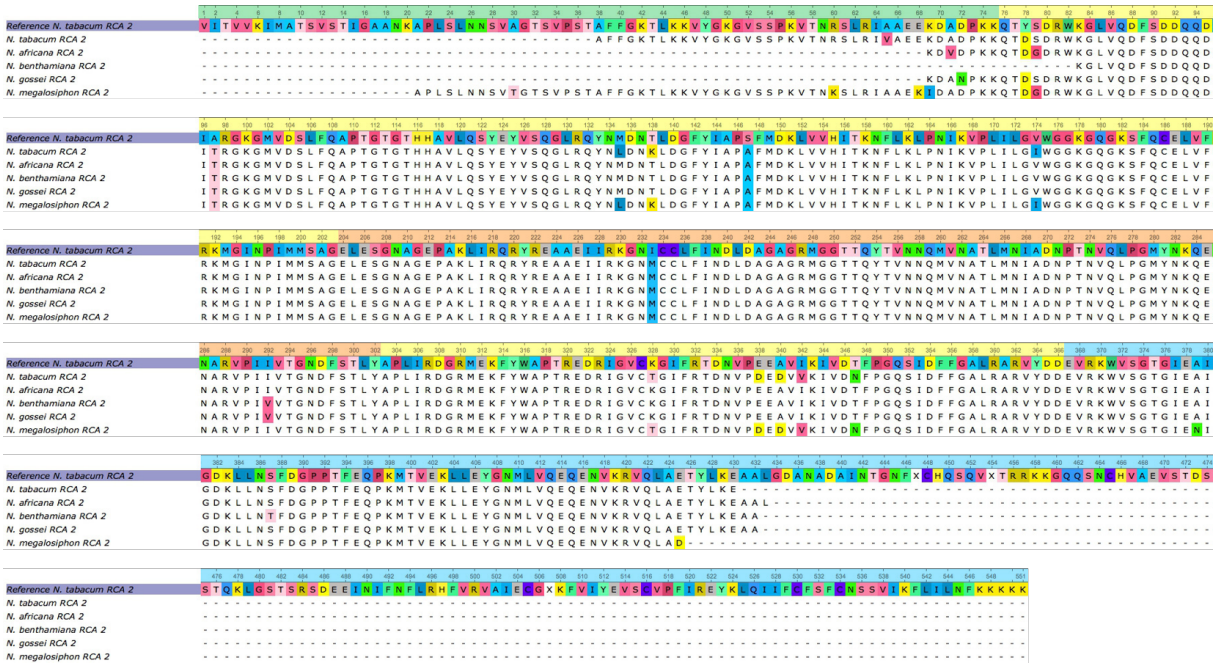


Figure 13. Peptide sequences of (a) RCA 1 and (b) RCA 2 from *N. tabacum*, *N. africana*, *N. benthamiana, N. gossei, and N. megalosiphon*. Sequences generated during the present study are presented alongside previously published *N. tabacum* RCA 1 (NM_001326055.1) and RCA 2 (Z14980), which served as the reference sequences for alignment and as indicators of sequencing coverage. Residue numbers above the peptide sequences are coloured according to RCA functional domains, based on the crystal structure of tobacco RCA (Stotz *et al.*, 2011): green = N-terminus domain, yellow = AAA^+ module, orange = ATPase core, and blue = C-terminus domain.

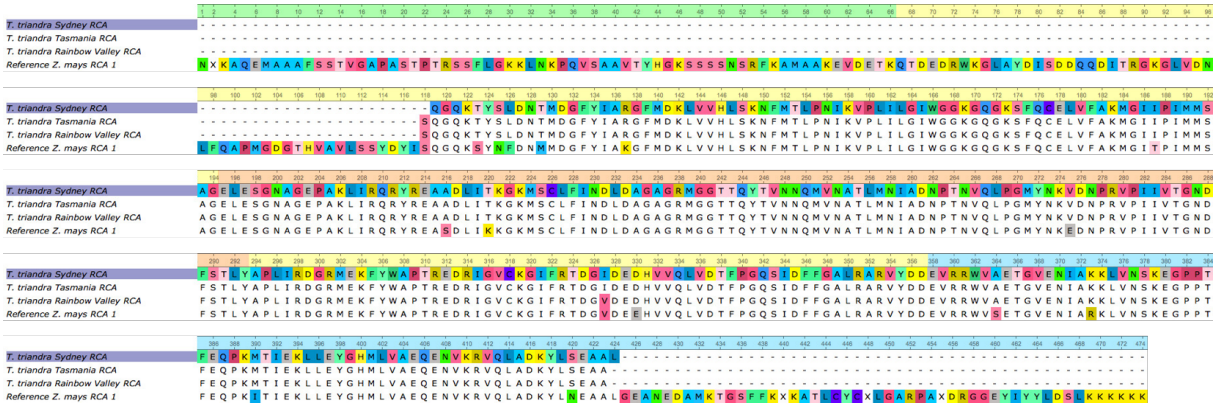


Figure 14. Peptide sequences of RCA of populations of *T. triandra* from Mount Wellington (Tasmania), Glenbrook (New South Wales), and Rainbow Valley (Northern Territory). Sequences generated during the present study are presented alongside previously published *Z. mays* RCA 1 (NM_001111451), which served as the reference sequence for alignment and as an indicator of sequencing coverage. Residue numbers above the peptide sequences are coloured according to RCA functional domains, based on the crystal structure of tobacco RCA (Stotz *et al.*, 2011): green = N-terminus domain, yellow = AAA^+ module, orange = ATPase core, and blue = C-terminus domain.

5. Discussion

In this study, I tested the hypothesis that endemic relatives of domestic cotton and tobacco adapted to the hot Australian arid-zone would exhibit photosynthetic thermotolerance through selective pressure on RCA. Specifically, I tested whether thermotolerance could be attributed to changes in the abundance of different RCA isoforms, changes in the RCA peptide sequences, or the combined effect of both. Our analyses also included three populations of Kangaroo grass (*Themeda triandra*) which occupy distinct thermal niches in the landscape.

Unexpectedly, I found that domestic cotton, along with some wild species, exhibited photosynthetic thermotolerance equivalent to its wild relatives, although some wild species had higher rates of photosynthesis at 30°C and 38°C. On the other hand, the photosynthetic rate of domestic tobacco declined at high temperature, while wild species exhibited a range of photosynthetic responses to heat. The abundance of RCA mirrored photosynthetic rates except in the C4 *Themeda* (see Figures 3-6), reflecting the importance of this enzyme as one of the key factors controlling the rate of carbon gain in C3 plants, even at optimal temperatures. The total RCA pool was unresponsive to temperature stress across all the selected species, even though there was evidence of species-specific regulation of unique RCA isoforms as temperature increased. I also found evidence of alternative splicing of an RCA isoform in *N. megalosiphon* at high temperature, though this was not observed in any

other species or isoform. Finally, translation of cDNA sequences revealed amino acid residue changes in two RCA isoforms between the five *Gossypium* and five *Nicotiana* species, and one isoform in the *Themeda* populations.

An unavoidable complication of this study is that the *precise* provenance of seed material for each wild *Gossypium* and *Nicotiana* species is unknown. The wide distribution of some species (e.g. *Gossypium australe*), and the likelihood of highly localised adaptation to micro-climates, may result in intraspecific variation in thermotolerance which has not been captured (Atwell *et al.*, 2014). To address this, I present comparisons across a range of species occupying different thermal niches – from relatively narrow (e.g. *G. robinsonii*) to widespread (e.g. *G. australe*) – and by including an explicit intra-species comparison with the inclusion of three distinct populations of *Themeda* from distinct and non-overlapping thermal environments (see Table 4).

5.1 Photosynthetic Thermotolerance in Heat-Adapted Species

Species from regions characterised by hot climates are typically thought to maintain robust photosynthetic capacity at high temperatures (Salvucci and Crafts-Brandner, 2004b). Surprisingly, there was a range of photosynthetic responses to day temperatures of 38°C in *Gossypium* and *Nicotiana* (Fig. 6 and Fig. 7). Importantly, the comparisons between wild and domestic *Gossypium* populations challenge current ideas about photosynthetic thermotolerance. Despite a clear interaction term between species and temperature, I did not observe uniformly greater photosynthetic thermotolerance in the wild *Gossypium* species relative to *G. hirsutum* (domestic cotton). Instead, this interaction arose from significant *decreases* in the normally high photosynthetic rates of *G. robinsonii* and *G. sturtianum* at 38°C, while *G. hirsutum* photosynthetic rates were lower but remained stable 38°C. This contrasts with reports of reduced photosynthetic capacity and increased Rubisco inhibition in *G. hirsutum* at 35^{o}C (Feller *et al.*, 1998; Law and Crafts-Brandner, 1999) but supports the absence of suppressed photosynthetic rates in *G. hirsutum* at 38^{o}C reported by Crafts-Brandner and Salvucci (2000).

While the sensitivity of photosynthesis in *G. sturtianum* to heat is consistent with its broad distribution and unknown seed provenance, the decline in the photosynthetic rate of

G. robinsonii at high temperature is more puzzling. *G. robinsonii* occurs across a narrow geographical range in the Gibson and Great Sandy Deserts of Western Australia and is exposed to a correspondingly narrow range of high summer temperatures (34.4 - 40.3°C; Table 4). Given the temperature-dependency of photosynthesis (Yamori *et al.*, 2014), it was expected that photosynthetic rates would be more resilient in *G. robinsonii*. However, the photosynthetic rate of *G. robinsonii* declined under heat. Notwithstanding, this species maintained high absolute rate even at 38°C and should therefore be pursued as a source of heat-tolerance genes.

A more predictable relationship between photosynthesis and heat was observed in the *Nicotiana* species which are all fast-growing perennial or annual plants (Horton, 1981). Hence, a species by temperature interaction was observed, arising from wild species adapted to hot environments either maintaining rapid photosynthesis at high temperature or even increasing, while rates declined in domestic tobacco and other wild tobacco species (Fig. 5a).

5.2 Stomatal Conductance, Leaf Cooling and Intercellular CO_2

The varied effects of heat on photosynthesis are partly explained by differences in stomatal physiology (Fig. 4b & 5b). In the presence of adequate water supply, transpiring plants can largely avoid heat stress through evaporative cooling as expressed by stomatal conductance (Matsui *et al.*, 2007; Wright *et al.*, 2017). In this study, plants had constant access to water. Where stomatal conductance in the *Gossypium* species did not change or decreased, there was either no change or a decline in photosynthesis at 38°C. Conversely, an increase in stomatal conductance in *G. hirsutum* might have supported photosynthetic capacity at high temperature. However, if differences in stomatal conductance did influence photosynthetic rates in *Gossypium* species then this was not obviously due to changes to leaf temperature, which were, on average, uniform across the species (Fig. 4c).

This contrasts with observations made on *Nicotiana* species, where changes in photosynthetic rates were linked to changes in stomatal conductance. Most starkly, the increase in stomatal conductance in *N. africana* cooled its leaves, which may have ameliorated the effect of heat on photosynthesis (Fig. 5). Interestingly, *N. megalosiphon* did not exhibit increased stomatal conductance and its leaves were generally the same

temperature as domestic tobacco in spite of a higher photosynthetic rate at 38°C. This combination of high photosynthetic rate and leaf temperature maintenance suggests that thermotolerance in *N. megalosiphon* was independent of stomatal physiology.

Intercellular [CO_2] correlates with photosynthetic rate and variations have been linked to changes in photosynthesis at high temperature (Bernacchi *et al.*, 2002; Scafaro *et al.*, 2011; Carmo-Silva *et al.*, 2012). In this study, there was an increase in intercellular [CO_2] in *G. hirsutum* at high temperature, though there was no effect of heat on intercellular [CO_2] in any other species (Fig. 4d; Table 8). This is interesting because while *G. hirsutum* has access to more CO_2 at high temperature relative to its control, its photosynthetic rate remains constant, which indicates vulnerability to heat at the chloroplast level. By contrast, photosynthetic rates of *G. robinsonii* were greater than *G. hirsutum* at both temperatures (Fig. 5a) despite the leaves of *G. robinsonii* not being cooler than *G. hirsutum* leaves and both species having the same internal [CO_2]. I speculate that *G. robinsonii* has photosynthetic thermotolerance at the chloroplast level.

There was no effect of temperature on intercellular [CO_2] in most *Nicotiana* species (Fig. 5d; Table 12), with all species except *N. benthamiana* having the same intercellular CO_2 at control and high temperature while photosynthetic rates varied widely under heat. Again, this points to biochemical variation in thermotolerance among *Nicotiana* species. Such photosynthetic thermotolerance has been extensively explored in other genera, where high temperature causes losses in photosynthetic efficiency due to the inactivation of Rubisco (Feller *et al.*, 1998; Crafts-Brandner and Salvucci, 2000; Crafts-Brandner and Salvucci, 2002; Salvucci and Crafts-Brandner, 2004b; Salvucci and Crafts-Brandner, 2004a; Carmo-Silva *et al.*, 2012; Scafaro *et al.*, 2012).

5.3 Responses of RCA Isoforms to High Temperature

The physiological data presented here (section 5.1) suggest that *G. robinsonii* and *N. megalosiphon* exhibit thermotolerance at the chloroplast level, relative to their domestic counterparts. Hence, I tested the hypothesis that these species regulated the accumulation of RCA isoforms differently from domestic species. Previous studies have shown that the expression of RCA is tightly linked to rates of photosynthesis (Saeed *et al.*, 2016), that RCA is

especially susceptible to heat (Weston *et al.*, 2007; Scafaro *et al.*, 2010) and that distinct RCA isoforms can perform different roles and be expressed independently (Weston *et al.*, 2007; Wang *et al.*, 2010; Bayramov and Guliyev, 2014). Further, it is known that alternate splicing of RCA mRNA is capable of producing separate RCA isoforms (Werneke *et al.*, 1989; Jurczyk *et al.*, 2015), and recent work recognises the role of alternate splicing in plant responses to stresses, including changes in temperature (Jurczyk *et al.*, 2016; Laloum *et al.*, In press).

Therefore, I determined the absolute abundance of the total RCA pool in *G. hirsutum*, *G. robinsonii*, *N. tabacum*, *N. megalosiphon*, and *T. triandra* using highly conserved isotope labelled RCA peptides. This showed that the *total* RCA content of each of these species did not change in response to heat (Fig. 6), contrasting with the findings of DeRidder and Salvucci (2007), Scafaro *et al.* (2010) and Chen *et al.* (2015) who showed that the abundance of various RCA isoforms increased in response to instantaneous temperature stress in cotton, and wild and domestic rice species, respectively. Imposition of heat without an acclimation period might account for the results from previous studies (Maire *et al.*, 2012; Walker *et al.*, 2014) because plants in this study were exposed to high temperature for four days, whereas Scafaro *et al.* (2010) and DeRidder and Salvucci (2007) imposed heat for 24 hours and 1 hour, respectively.

Importantly, the ratio of absolute RCA to the large subunit of Rubisco (RbcL) remained constant in all species across temperatures except in *G. robinsonii*, where RCA:RbcL increased at high temperature (Fig. 7). This change was not due to changes in RCA abundance but instead by a lower abundance of RbcL (data not shown). Rubisco demands a significant energy and nitrogen investment (Carmo-Silva *et al.*, 2015) so the reduction in RbcL may reflect an acclimation in resource use in *G. robinsonii* at high temperature. The maintenance of a high photosynthetic rate in *G. robinsonii* at high temperature may result from more copies of RCA servicing fewer copies of Rubisco (Wei *et al.*, 2017). Conceivably, *G. robinsonii* maintains photosynthesis by 'prioritising' Rubisco activation while saving nitrogen and energy. In short, photosynthetic rates are not directly dependent on the abundance of Rubisco *per se*, but rather the abundance of *active* Rubisco enzymes (Thanh *et al.*, 2011).

While there was no change in the absolute abundance of RCA in any species at high temperature, there was species-specific regulation of unique RCA isoforms (Figures 10 and 11; see Fig. 8 and Supplementary Material Table 3 for isoform details). Species from hotter climates are thought to have evolved thermostable RCA isoforms and/or regulate the expression of RCA isoforms differentially to maximise photosynthetic efficiency at high temperature (Salvucci and Crafts-Brandner, 2004b; Chao *et al.*, 2014; Scafaro *et al.*, 2016; Shivhare and Mueller-Cajar, 2017). I identified four isoforms of RCA in both *Gossypium* and *Nicotiana* species, and one isoform in *Themeda* as well as three RCA mutant isoforms in *Gossypium* and one in *Nicotiana*. Mutant isoforms were more abundant in the wild *Gossypium* and *Nicotiana* species.

G. hirsutum maintained an unchanged abundance of RCA 1 and RCA 2 at high temperature while *G. robinsonii* had stable RCA 1 abundance and decreasing RCA 2 abundance relative to its control. By contrast, *N. tabacum* exhibited signs of increasing RCA 1 abundance, and signs of decreasing RCA 2 abundance whereas *N. megalosiphon* showed signs of increased RCA 1 *and* RCA 2 at high temperature. I also identified evidence of a possible splice variant of RCA 1 (Isoform 3) arising in *N. megalosiphon* at high temperature (Fig. 14). The importance of splice variants of RCA in the tolerance of ryegrass to changes in temperature has been highlighted (Jurczyk *et al.*, 2016).

In *Themeda* there was no change in the absolute abundance of RCA at high temperature, no shift in RCA:RbcL, and no evidence of alternative splicing. The abundance of RCA and RbcL were low in *Themeda* relative to *Gossypium* and *Nicotiana*. These results reflect the reduced reliance on Rubisco activity, and therefore on RCA, in *Themeda*, which is a C4 species (Raven, 2013).

5.4 Divergence in Primary Structure of RCA Isoforms Between Species

Even if RCA expression changes in response to heat, every species must reach a temperature threshold above which heat-labile isoforms do not contribute to photosynthetic thermotolerance. At this threshold, RCA isoforms that possess mutations for increased rigidity become critical for thermotolerance (for example Rathi *et al.*, 2015). Within species, these RCA isoforms make differential contributions to heat tolerance

(Salvucci and DeRidder, 2006) and may explain changes in abundance at high temperature. Shivhare and Mueller-Cajar (2017) compared the activities of chimeric agave (*Agave tequilana*) and domestic rice (*Oryza sativa*) RCA isoforms to elucidate regions of the two proteins that contribute to thermotolerance of agave relative to heat-labile rice RCA isoforms. They showed that RCA thermostability is conferred by the first 250 residues of the enzyme. In most species, increases in RCA stability comes at a cost to function (D'Amico *et al.*, 2003), though this was not observed in agave, suggesting multiple routes to thermotolerant RCA (Shivhare and Mueller-Cajar, 2017).

Here, I identify several amino acid residue changes in the sequences of RCA 1 and RCA 2 in *Gossypium* and *Nicotiana* and a single mutation in RCA between three populations of *T. triandra* (Fig. 15-17). The two residue changes detected in *G. robinsonii* RCA 1 were shared with *G. bickii* RCA 1 and both occurred within the proposed thermostability-determining region of the sequence (Shivhare and Mueller-Cajar, 2017). However, despite its response to temperature, *G. robinsonii* maintains a high rate of photosynthesis at high temperature, while *G. bickii* has consistently low photosynthetic rates at both control and high temperatures. I hypothesise that the specific residue changes in *G. robinsonii* reduce its rigidity much as in domestic rice (Scafaro *et al.*, 2016): increasing the 'floppiness' would allow *G. robinsonii* RCA 1 to act on Rubisco quickly, especially at control temperature, but would also render it somewhat susceptible to high temperature. Likewise, the C-terminal of *G. bickii* RCA, containing multiple residue changes, may contribute to increasing its rigidity, and may therefore explain the differences in photosynthetic performance between *G. bickii* and *G. robinsonii*. However, photosynthesis is a complex process so it likely that there are multiple factors affecting the rates at which different species accrue CO_2.

There is no evidence that the thermostability-determining region of agave is seen in the tobacco sequence. For example, the photosynthetic rate of *N. africana* and *N. megalosiphon* increased uniquely at high temperature even though they share every mutation with at least one other species (except for A71V in RCA 2). This does not mean categorically that the mutations in the primary structure of these RCA isoforms are not contributing to thermostability; it may simply indicate that the model of RCA thermostability proposed by Shivhare and Mueller-Cajar (2017) does not apply to RCA in cotton or tobacco. For example,

Scafaro *et al.* (2016) conclude that the considerable differences in the thermostability of RCA from domestic and wild rice can be attributed to 19 amino acid differences distributed across the entire sequence in the wild RCA isoform, and not due to changes in the first 250 residues alone.

Interestingly, only one residue change was detected between the three populations of *Themeda*, and it occurred in the Rainbow Valley population. Aligning the *Themeda* gene and peptide sequences with that of RCA 1 from *Z. mays*, a closely related species with recognised thermotolerance (Frova, 1996), shows that the Rainbow Valley *Themeda* shares many SNPs with *Z. mays,* though only one of the SNPs results in a shared residue change. While the contribution of a single residue change to overall thermotolerance of RCA may be negligible, the apparent convergence of the gene sequences suggests some kind of movement towards thermotolerance in *Themeda* from the tropical Rainbow Valley accession.

5.5 Concluding Remarks

This study reports on some of the physiological and biochemical changes that occur in wild and domestic species of cotton and tobacco, and in kangaroo grass, adapted to different thermal niches under high temperature. I found evidence of photosynthetic thermotolerance in some wild species *Gossypium* and *Nicotiana*, while some were unexpectedly susceptible to heat relative to the domestic species. I find that differences in photosynthetic thermotolerance in *G. hirsutum* and *G. robinsonii* may be explained by changes in the abundance not only of individual RCA isoforms but also in the abundance of RCA relative to Rubisco in *G. robinsonii.* Likewise, the difference in the thermotolerance of *N. tabacum* and *N. megalosiphon* may be explained by subtle differences in the abundance of different RCA isoforms at control and high temperature. Of particular note is the abundance of mutant RCA isoforms present in wild, but not domestic, species. Shifting isoform expression patterns may reflect the relative importance of these isoforms across species and temperatures, and highlight differences in thermotolerance of the RCA isoforms. Differences in isoform thermotolerance may be driven by the combined contribution of all residues changes detected in the primary peptide sequences or, less

likely, may be attributed to a few residue changes in a proposed thermostability-determining region of RCA (Shivhare and Mueller-Cajar, 2017).

While this study did not extend to test catalytic differences in the thermostability of RCA isoforms, future studies could identify the contribution of the polymorphisms in the thermostability of RCA across species using a Rubisco activation assay (Scales *et al.*, 2014). Additionally, it would be beneficial to monitor changes in photosynthesis following the over-expression or knock-out of RCA, and/or track changes in the yield components of modified plants. The inclusion of more *Nicotiana* and *Gossypium* species from hot-arid locations around the world is recommended, as well as a wider range of the many *T. triandra* populations available. Finally, the double base calls identified in Supplementary Materials section 7.2.9 can be resolved by cloning the cDNA in bacterial vectors and re-sequencing the isoforms from different colonies. This would also allow for full coverage of the RCA sequences, and would therefore identify any other residue changes that might exist between species.

The present study provides further evidence that wild crop relatives show diverse responses to abiotic stresses. Wild crop relatives represent a vast reservoir of untapped genetic diversity and characterising the function of RCA from diverse sources is expected to enhance the development of domesticated lines of major crop species with improved thermotolerance.

References

Alvarez-Ponce D, Sabater-Muñoz B, Toft C, et al. 2016. Essentiality is a strong determinant of protein rates of evolution during mutation accumulation experiments in *Escherichia coli*. Genome Biology and Evolution **8**, 2914-2927.

Ashraf M, Harris PJC. 2013. Photosynthesis under stressful environments: An overview. Photosynthetica **51**, 163-190.

Asseng S, Ewert F, Martre P, et al. 2014. Rising temperatures reduce global wheat production. Nature Climate Change **5**, 143-147.

Atwell BJ, Wang H, Scafaro AP. 2014. Could abiotic stress tolerance in wild relatives of rice be used to improve *Oryza sativa*? Plant Science **215-216**, 48-58.

Bates D, Mächler M, Bolker B, et al. 2015. Fitting linear mixed-effects models using lme4. 2015 **67**, 48.

Bayramov S, Guliyev N. 2014. Changes in Rubisco activase gene expression and polypeptide content in Brachypodium distachyon. Plant Physiology and Biochemistry **81**, 61-66.

Benomar L, Lamhamedi MS, Rainville A, et al. 2016. Genetic adaptation vs. ecophysiological plasticity of photosynthetic-related traits in young *Picea glauca* trees along a regional climatic gradient. Frontiers in Plant Science doi: 10.3389/fpls.2016.00048.

Bernacchi CJ, Portis AR, Nakano H, et al. 2002. Temperature response of mesophyll conductance. Implications for the determination of Rubisco enzyme kinetics and for limitations to photosynthesis in vivo. Plant Physiology **130**, 1992-1998.

Bishop J, Potts SG, Jones HE. 2016. Susceptibility of faba bean (*Vicia faba* L.) to heat stress during floral development and anthesis. Journal of Agronomy and Crop Science **202**, 508-517.

Bita C, Gerats T. 2013. Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. Frontiers in Plant Science doi: 10.3389/fpls.2013.00273.

Bortesi L, Fischer R. 2014. The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnology Advances **33**, 41-52.

Carmo-Silva EA, Gore MA, Andrade-Sanchez P, et al. 2012. Decreased CO2 availability and inactivation of Rubisco limit photosynthesis in cotton plants under heat and drought stress in the field. Environmental and Experimental Botany **83**, 1-11.

Carmo-Silva E, Scales JC, Madgwick PJ, et al. 2015. Optimizing Rubisco and its regulation for greater resource use efficiency. Plant, Cell & Environment **38**, 1817-1832.

Cavanagh AP, Kubien DS. 2014. Can phenotypic plasticity in Rubisco performance contribute to photosynthetic acclimation? Photosynthesis Research **119**, 203-214.

Chao M, Yin Z, Hao D, et al. 2014. Variation in Rubisco activase (RCAβ) gene promoters and expression in soybean (Glycine max L. Merr.). Journal of Experimental Botany **65**, 47-59.

Chen Y, Wang XM, Zhou L, et al. 2015. Rubisco activase is also a multiple responder to abiotic stresses in rice. PLoS One doi: 10.1371/journal.pone.0140934.

Cowtan K, Hausfather Z, Hawkins E, et al. 2015. Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. Geophysical Research Letters **42**, 6526-6534.

Crafts-Brandner SJ, Salvucci ME. 2000. Rubisco activase constrains the photosynthetic potential of leaves at high temperature and CO2. PNAS **97**, 13430-13435.

Crafts-Brandner SJ, Salvucci ME. 2002. Sensitivity of photosynthesis in a C4 plant, maize, to heat stress. Plant Physiology **129**, 1773-1780.

Criddle RS, Hopkin MS, McArthur ED, et al. 1994. Plant distribution and the temperature coefficient of metabolism. Plant, Cell & Environment **17**, 233-243.

D'Amico S, Marx J-C, Gerday C, et al. 2003. Activity-stability relationships in extremophilic enzymes. Journal of Biological Chemistry **278**, 7891-7896.

Dempewolf H, Guarino L. 2015. Reaching back through the domestication bottleneck: tapping wild plant biodiversity for crop improvement. Acta Horticulturae 10.17660/ActaHortic.2015.1101.25.

DeRidder BP, Salvucci ME. 2007. Modulation of Rubisco activase gene expression during heat stress in cotton (*Gossypium hirsutum* L.) involves post-transcriptional mechanisms. Plant Science **172**, 246-254.

Evans JR. 2013. Improving photosynthesis. Plant Physiology **162**, 1780-1793.

Farquhar GD, von Caemmerer S, Berry JA. 1980. A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta **149**, 78-90.

Feller U, Crafts-Brandner SJ, Salvucci ME. 1998. Moderately high temperatures inhibit ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) activase-mediated activation of Rubisco. Plant Physiology **116**, 539-546.

Fernandez-Pozo N, Menda N, Edwards JD, et al. 2015. The Sol Genomics Network (SGN)—from genotype to phenotype to breeding. Nucleic Acids Research 10.1093/nar/gku1195.

Fitzgerald TL, Shapter FM. 2011. Genome diversity in wild grasses under environmental stress. PNAS **108**, 21140-21145.

Fox J, Weisberg S. 2011. An R Companion to Applied Regression. Thousand Oaks, CA: Sage.

Frich P, Alexander LV, Della-Marta P, et al. 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. Climate Research **19**, 193-212.

Frova C. 1996. Genetic dissection of thermotolerance in maize. In: Grillo S, Leone A, eds. Physical Stresses in Plants: Genes and Their Products for Tolerance. Berlin, Heidelberg: Springer Berlin Heidelberg, 31-38.

Fukayama H, Ueguchi C, Nishikawa K, et al. 2012. Overexpression of Rubisco activase decreases the photosynthetic CO2 assimilation rate by reducing Rubisco content in rice leaves. Plant and Cell Physiology **53**, 976-986.

Gouy M, Milleret F, Mugnier C, et al. 1984. ACNUC: a nucleic acid sequence data base and analysis system. Nucleic Acids Research **12**, 121-127.

Gutteridge S, Gatenby AA. 1995. Rubisco synthesis, assembly, mechanism, and regulation. The Plant Cell **7**, 809-819.

Hasanuzzaman M, Nahar K, Alam MM, et al. 2013. Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. International Journal of Molecular Sciences **14**, 9643-9684.

Hatfield JL, Boote KJ, Kimball BA, et al. 2011. Climate impacts on agriculture: implications for crop production. Agronomy Journal **103**, 351-370.

Hauser T, Popilka L, Hartl UF, et al. 2015. Role of auxiliary proteins in Rubisco biogenesis and function. Nature Plants **1**, 15065.

Hendrickson L, Sharwood R, Ludwig M, et al. 2007. The effects of Rubisco activase on C4 photosynthesis and metabolism at high temperature. Journal of Experimental Botany **59**, 1789-1798.

Hill JT, Demarest BL, Bisgrove BW, et al. 2014. Poly Peak Parser: method and software for identification of unknown indels using Sanger Sequencing of PCR products. Developmental Dynamics **243**, 1632-1636.

Horton P. 1981. A taxonomic revision of *Nicotiana* (Solanaceae) in Australia. Journal of the Adelaide Botanic Garden **3**, 1-56.

Hothorn T, Bretz F, Westfall P. 2008. Simultaneous inference in general parametric models. Biometrical Journal **50**, 346-363.

Hsu C-H, Chiang AWT, Hwang M-J, et al. 2016. Proteins with highly evolvable domain architectures are nonessential but highly retained. Molecular Biology and Evolution **33**, 1219-1230.

Hüner N, Dahal K, Bode R, et al. 2016. Photosynthetic acclimation, vernalization, crop productivity and 'the grand design of photosynthesis'. Journal of Plant Physiology **203**, 29-43.

Hüner NPA, Dahal K, Kurepin LV, et al. 2014. Potential for increased photosynthetic performance and crop productivity in response to climate change: role of CBFs and gibberellic acid. Frontiers in Chemistry doi: 10.3389/fchem.2014.00018.

Hutchinson M, Evans B, Stein J, et al. 2014. Monthly daily maximum temperature: ANUClimate 1.0, 0.01 degree, Australian Coverage, 1970-2012. National Computational Infrastructure.

Hyten DL, Song Q, Zhu Y, et al. 2006. Impacts of genetic bottlenecks on soybean genome diversity. PNAS **103**, 16666-16671.

IPCC. 2014a. Part A: global and sectoral aspects – Working Group II contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change. In: Field CB, Barros VR, Mach K, Mastrandrea M, eds. Climate change 2014: impacts, adaptation, and vulnerability. Cambridge University Press, Cambridge, United Kingdom and New York, NY , USA: Cambridge University Press.

IPCC. 2014b. Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In: Pachauri RK, Meyer LA, eds. Climate Change 2014. Geneva, Switzerland: IPCC, 151.

Ishikawa C, Hatanaka T, Misoo S, et al. 2011. Functional incorporation of sorghum small subunit increases the catalytic turnover rate of Rubisco in transgenic rice. Plant physiology **156**, 1603-1611.

Jiang Y, Wang J, Tao X, et al. 2013. Characterization and expression of Rubisco activase genes in *Ipomoea batatas*. Molecular Biology Reports **40**, 6309-6321.

Jing P, Wang D, Zhu C, et al. 2016. Plant physiological, morphological and yield-related responses to night temperature changes across different species and plant functional types. Frontiers in Plant Science doi: 10.3389/fpls.2016.01774.

Jurczyk B, Hura K, Trzemecka A, et al. 2015. Evidence for alternative splicing mechanisms in meadow fescue (*Festuca pratensis*) and perennial ryegrass (*Lolium perenne*) Rubisco activase gene. Journal of Plant Physiology **176**, 61-64.

Jurczyk B, Pociecha E, Grzesiak M, et al. 2016. Enhanced expression of Rubisco activase splicing variants differentially affects Rubisco activity during low temperature treatment in *Lolium perenne*. Journal of Plant Physiology **198**, 49-55.

Kaluthota S, Pearce DW, Evans LM, et al. 2015. Higher photosynthetic capacity from higher latitude: foliar characteristics and gas exchange of southern, central and northern populations of Populus angustifolia. Tree Physiology **35**, 936-948.

Kaur R, Chakraborty A, Bhunia RK, et al. 2017. Tolerance to soil water stress by *Oryza sativa* cv. IR20 was improved by expression of Wsi18 gene locus from *Oryza nivara*. Biologia Plantarum doi: 10.1007/s10535-017-0742-7.

Kaushal N, Bhandari K, Siddique K, et al. 2016. Food crops face rising temperatures: An overview of responses, adaptive mechanisms, and approaches to improve heat tolerance. Cogent Food and Agriculture doi: 10.1080/23311932.2015.1134380.

Kromdijk J, Long SP. 2016. One crop breeding cycle from starvation? How engineering crop photosynthesis for rising CO2 and temperature could be one important route to alleviation. Proceedings of the Royal Society of London B Biological Sciences doi: 10.1098/rspb.2015.2578.

Kumar A, Li C, Portis AR. 2009. *Arabidopsis thaliana* expressing a thermostable chimeric Rubisco activase exhibits enhanced growth and higher rates of photosynthesis at moderately high temperatures. Photosynthesis Research **100**, 143-153.

Kuznetsova A, Brockhoff B, Christensen HB. 2016. lmerTest: tests in linear mixed effect models. Journal of Statistical Software doi: 10.18637/jss.v082.i13.

Laloum T, Martín G, Duque P. In press. Alternative Splicing Control of Abiotic Stress Responses. Trends in Plant Science doi: 10.1016/j.tplants.2017.09.019.

Law DR, Crafts-Brandner SJ. 1999. Inhibition and acclimation of photosynthesis to heat stress is closely correlated with activation of ribulose-1, 5-bisphosphate carboxylase/oxygenase. Plant Physiology **120**, 173-182.

Long SP, Bernacchi CJ. 2003. Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. Journal of Experimental Botany **54**, 2393-2401.

Lopes MS, El-Basyoni I, Baenziger PS, et al. 2015. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. Journal of Experimental Botany **66**, 3477-3486.

Lozano-Juste J, Cutler SR. 2014. Plant genome engineering in full bloom. Trends in Plant Science **19**, 284-287.

Luo Q. 2011. Temperature thresholds and crop production: a review. Climatic Change **109**, 583-598.

Maire V, Martre P, Kattge J, et al. 2012. The coordination of leaf photosynthesis links C and N fluxes in C(3) plant species. PLoS One doi: 10.1371/journal.pone.0038345.

Matsui T, Kobayasi K, Yoshimoto M, et al. 2007. Stability of rice pollination in the field under hot and dry conditions in the Riverina region of New South Wales, Australia. Plant Production Science **10**, 57-63.

Maurino VG, Peterhansel C. 2010. Photorespiration: current status and approaches for metabolic engineering. Current Opinion in Plant Biology **13**, 248-255.

Maurino VG, Weber APM. 2013. Engineering photosynthesis in plants and synthetic microorganisms. Journal of Experimental Botany **64**, 743-751.

McGrath CL, Gout J-F, Johri P, et al. 2014. Differential retention and divergent resolution of duplicate genes following whole-genome duplication. Genome Research **24**, 1665-1675.

McGrath CL, Lynch M. 2012. Evolutionary significance of whole-genome duplication. In: Soltis PS, Soltis DE, eds. Polyploidy and Genome Evolution. Berlin: Springer.

Mehta SC, Rice K, Palzkill T. 2015. Natural variants of the KPC-2 carbapenemase have evolved increased catalytic efficiency for ceftazidime hydrolysis at the cost of enzyme stability. PLoS Pathogens doi: 10.1371/journal.ppat.1004949.

Mercer KL, Perales HR. 2010. Evolutionary response of landraces to climate change in centers of crop diversity. Evolutionary Applications **3**, 480-493.

Morita K, Hatanaka T, Misoo S, et al. 2014. Unusual small subunit that is not expressed in photosynthetic cells alters the catalytic properties of Rubisco in rice. Plant Physiology **164**, 69-79.

Morrell PL, Buckler ES, Ross-Ibarra J. 2011. Crop genomics: advances and applications. Nature Reviews Genetics **13**, 85-96.

Mueller-Cajar O, Stotz M, Bracher A. 2014. Maintaining photosynthetic CO2 fixation via protein remodelling: the Rubisco activases. Photosynthesis Research **119**, 191-201.

Nölke G, Houdelet M, Kreuzaler F, et al. 2014. The expression of a recombinant glycolate dehydrogenase polyprotein in potato (*Solanum tuberosum*) plastids strongly enhances photosynthesis and tuber yield. Plant Biotechnology Journal **12**, 734-742.

Ohno S. 1970. Evolution by gene duplication. New York: Springer Science & Business Media.

Okonechnikov K, Golosova O, Fursov M. 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics **28**, 1166-1167.

Orr D, Alcântara A, Kapralov MV, et al. 2016. Surveying Rubisco diversity and temperature response to improve crop photosynthetic efficiency. Plant Physiology **172**, 707-717.

Ort DR, Merchant SS, Alric J, et al. 2015. Redesigning photosynthesis to sustainably meet global food and bioenergy demand. PNAS **112**, 8529-8536.

Palmgren MG, Edenbrandt A, Vedel S, et al. 2015. Are we ready for back-to-nature crop breeding? Trends in Plant Science **20**, 155-164.

Parry MAJ, Andralojc PJ, Mitchell RAC, et al. 2003. Manipulation of Rubisco: the amount, activity, function and regulation. Journal of Experimental Botany **54**, 1321-1333.

Parry MAJ, Andralojc PJ, Scales JC. 2013. Rubisco activity and regulation as targets for crop improvement. Journal of Experimental Botany **64**, 717-730.

Parry MAJ, Madgwick PJ, Carvalho JFC, et al. 2006. Prospects for increasing photosynthesis by overcoming the limitations of Rubisco. Journal of Agricultural Science **145**, 31-43.

Phillips AS, Deser C, Fasullo J. 2014. Evaluating modes of variability in climate models. EOS, Transactions American Geophysical Union **95**, 453-455.

Portis AR. 2003. Rubisco activase–Rubisco's catalytic chaperone. Photosynthesis Research **75**, 11-27.

Prins A, Orr DJ, Andralojc JP, et al. 2016. Rubisco catalytic properties of wild and domesticated relatives provide scope for improving wheat photosynthesis. Journal of Experimental Botany **67**, 1827-1838.

R Core Team. 2017. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rathi PC, Jaeger K-E, Gohlke H. 2015. Structural rigidity and protein thermostability in variants of lipase A from *Bacillus subtilis*. PLoS One doi: 10.1371/journal.pone.0130289.

Raven JA. 2013. Rubisco: still the most abundant protein of Earth? New Phytologist **198**, 1-3.

Reddy KR, Hodges HF, McKinion JM. 1993. Temperature effects on Pima cotton leaf growth. Agronomy Journal **85**, 681-686.

Reddy KR, Hodges HF, Reddy VR. 1992. Temperature effects on cotton fruit retention. Agronomy Journal **84**, 26-30.

Saeed I, Bachir DG, Chen L, et al. 2016. The expression of RaRCA2-α gene associated with net photosynthesis rate, biomass and grain yield in bread wheat (*Triticum aestivum* l.) under field conditions. PLoS One doi: 10.1371/journal.pone.0161308.

Salvucci ME, Crafts-Brandner SJ. 2004a. Inhibition of photosynthesis by heat stress: the activation state of Rubisco as a limiting factor in photosynthesis. Physiologia Plantarum **120**, 179-186.

Salvucci ME, Crafts-Brandner SJ. 2004b. Relationship between the heat tolerance of photosynthesis and the thermal stability of Rubisco activase in plants from contrasting thermal environments. Plant Physiology **134**, 1460-1470.

Salvucci ME, Crafts-Brandner SJ. 2004. Mechanism for deactivation of Rubisco under moderate heat stress. Physiologia Plantarum **122**, 513-519.

Salvucci ME, DeRidder BP. 2006. Effect of activase level and isoform on the thermotolerance of photosynthesis in *Arabidopsis*. Journal of Experimental Botany **57**, 3793-3799.

Scafaro AP, Gallé A, Rie J, et al. 2016. Heat tolerance in a wild *Oryza* species is attributed to maintenance of Rubisco activation by a thermally stable Rubisco activase ortholog. New Phytologist doi: 10.1111/nph.13963.

Scafaro AP, Haynes PA, Atwell BJ. 2010. Physiological and molecular changes in *Oryza meridionalis* Ng., a heat-tolerant species of wild rice. Journal of Experimental Botany **61**, 191-202.

Scafaro AP, Von Caemmerer S, Evans JR, et al. 2011. Temperature response of mesophyll conductance in cultivated and wild *Oryza* species with contrasting mesophyll cell wall thickness. Plant, Cell & Environment **34**, 1999-2008.

Scafaro AP, Yamori W, Carmo-Silva EA, et al. 2012. Rubisco activity is associated with photosynthetic thermotolerance in a wild rice (*Oryza meridionalis*). Physiologia Plantarum **146**, 99-109.

Scales JC, Parry MAJ, Salvucci ME. 2014. A non-radioactive method for measuring Rubisco activase activity in the presence of variable ATP: ADP ratios, including modifications for measuring the activity and activation state of Rubisco. Photosynthesis Research **119**, 355-365.

Scharf KD, Berberich T, Ebersberger I. 2012. The plant heat stress transcription factor (Hsf) family: structure, function and evolution. Biochimica et Biophysica Acta **1819**, 104-119.

Sharkey TD. 2005. Effects of moderate heat stress on photosynthesis: importance of thylakoid reactions, Rubisco deactivation, reactive oxygen species, and thermotolerance provided by isoprene. Plant, Cell & Environment **28**, 269-277.

Sharkey TD, Badger MR, von Caemmerer S. 2001. Increased heat sensitivity of photosynthesis in tobacco plants with reduced Rubisco activase. Photosynthesis Research **67**, 147-156.

Sharkey TD, Zhang R. 2010. High temperature effects on electron and proton circuits of photosynthesis. Journal of Integrative Plant Biology **52**, 712-722.

Shivhare D, Mueller-Cajar O. 2017. Characterization of thermostable CAM Rubisco activase reveals a Rubisco interacting surface loop. Plant Physiology **174**, 1505-1516.

Sicher R. 2015. Temperature shift experiments suggest that metabolic impairment and enhanced rates of photorespiration decrease organic acid levels in soybean leaflets exposed to supra-optimal growth temperatures. Metabolites **5**, 443-454.

Stotz M, Mueller-Cajar O, Ciniawsky S, et al. 2011. Structure of green-type Rubisco activase from tobacco. Nature Structural & Molecular Biology **18**, 1366-1370.

Studer RA, Christin P-A, Williams MA, et al. 2014. Stability-activity tradeoffs constrain the adaptive evolution of Rubisco. PNAS **111**, 2223-2228.

Tabita FR, Satagopan S, Hanson TE, et al. 2008. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. Journal of Experimental Botany **59**, 1515-1524.

Tcherkez GGB, Farquhar GD. 2006. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. PNAS **103**, 7246-7251.

Thanh T, Chi VTQ, Abdullah MP, et al. 2011. Cloning and characterization of ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit (RbcS) cDNA from green microalga *Ankistrodesmus convolutus*. Molecular Biology Reports **38**, 5297-5305.

The UniProt Consortium. 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Research **45**, D158-D169.

Tilman D, Balzer C, Hill J, et al. 2011. Global food demand and the sustainable intensification of agriculture. PNAS **108**, 20260-20264.

Truongvan N, Jang S-H, Lee C. 2016. Flexibility and stability trade-off in active site of cold-adapted Pseudomonas mandelii esterase EstK. Biochemistry **55**, 3542-3549.

Untergasser A, Cutcutache I, Koressaar T, et al. 2012. Primer3—new capabilities and interfaces. Nucleic Acids Research **40**, e115-e115.

Vincent H, Wiersema J, Kell S, et al. 2013. A prioritized crop wild relative inventory to help underpin global food security. Biological Conservation **167**, 265-275.

von Caemmerer S, Farquhar GD. 1981. Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves. Planta **153**, 376-387.

Wachter RM, Henderson NJ. 2015. Photosynthesis: Rubisco rescue. Nature Plants doi: 10.1038/nplants.2014.10.

Wachter RM, Salvucci ME, Carmo-Silva EA, et al. 2013. Activation of interspecies-hybrid Rubisco enzymes to assess different models for the Rubisco–Rubisco activase interaction. Photosynthesis Research **117**, 557-566.

Walker AP, Beckerman AP, Gu L, et al. 2014. The relationship of leaf photosynthetic traits – Vcmax and Jmax – to leaf nitrogen, leaf phosphorus, and specific leaf area: a meta-analysis and modeling study. Ecology and Evolution **4**, 3218-3235.

Wang D, Heckathorn SA, Mainali K, et al. 2016. Timing effects of heat-stress on plant ecophysiological characteristics and growth. Frontiers in Plant Science doi: 10.3389/fpls.2016.01629.

Wang D, Li X-F, Zhou Z-J, et al. 2010. Two Rubisco activase isoforms may play different roles in photosynthetic heat acclimation in the rice plant. Physiologia Plantarum **139**, 55-67.

Warschefsky E, Penmetsa VR, Cook DR, et al. 2014. Back to the wilds: tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. American Journal of Botany **101**, 1791-1800.

Wei H, Liu J, Wang Y, et al. 2013. A dominant major locus in chromosome 9 of rice (*Oryza sativa* L.) confers tolerance to 48°C high temperature at seedling stage. Journal of Heredity **104**, 287-294.

Wei L, Wang Q, Xin Y, et al. 2017. Enhancing photosynthetic biomass productivity of industrial oleaginous microalgae by overexpression of RuBisCO activase. Algal Research **27**, 366-375.

Werneke JM, Chatfield JM, Ogren WL. 1989. Alternative mRNA splicing generates the two ribulosebisphosphate carboxylase/oxygenase activase polypeptides in spinach and *Arabidopsis*. The Plant Cell **1**, 815-825.

Wessel D, Flugge UI. 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Annals of Biochemistry **138**, 141-143.

Weston DJ, Bauerle WL, Swire-Clark GA, et al. 2007. Characterization of Rubisco activase from thermally contrasting genotypes of *Acer rubrum* (Aceraceae). American Journal of Botany **94**, 926-934.

Whitney SM, Kane HJ, Houtz RL. 2009. Rubisco oligomers composed of linked small and large subunits assemble in tobacco plastids and have higher affinities for CO2 and O2. Plant Physiology **149**, 1887-1895.

Wickham H. 2009. ggplot2: Elegant Graphics for Data Analysis: Springer Publishing Company, Incorporated.

Wright IJ, Dong N, Maire V, et al. 2017. Global climatic drivers of leaf size. Science **357**, 917-921.

Yamori W, Hikosaka K, Way DA. 2014. Temperature response of photosynthesis in C3, C4, and CAM plants: temperature acclimation and temperature adaptation. Photosynthesis Research **119**, 101-117.

Zhu Y, Yan H, Wang Y, et al. 2016. Genome duplication and evolution of heat shock transcription factor (Hsf) gene family in four model angiosperms. Journal of Plant Growth Regulation doi: 10.1007/s00344-016-9590-5.

7. Supplementary Material

7.1 Methods

7.1.1 Supplementary Table 1: Initial primers designed for the amplification of RCA and ADH.

Species	*Gene Name*	*Terminal/Middle*	*Pair no.*	*Primer ID (start site)*	*Primer Sequence (5'->3')*	*Tm (°C)*	*GC%*
G. hirsutum	RCA 1 (beta)	Terminal	1	F37-GhRCA1	CGAGCACCGTTGAGTTTGAA	59.07	50
				R1286-GhRCA1	TTAGCATTTCCAAGGGCAGC	58.81	50
			2	F55-GhRCA1	AATGGATCAGGTGCCGGAG	59.47	57.89
				R1282-GhRCA1	CATTTCCAAGGGCAGCTTCA	58.74	50
		Middle	1	F759-GhRCA1	GAACGCCACACTCATGAACA	58.77	50
				R860-GhRCA1	CCGGTGACAATGATCGGAAC	58.99	55
			2	F778-GHRCA1	ATCGCTGATAACCCCACCAA	59.08	50
				R877-GHRCA1	GCGTCGAGAAATCGTTACCG	59.44	55
	RCA 2 (alpha)	Terminal	1	F32-GhRCA2	AGTTTTAAGGTGATGGCGGC	58.83	50
				R1250-GhRCA2	CATTTGGATCAGTGCACCCT	58.16	50
			2	F64-GhRCA2	AGACGAAGAGACACAGACCG	59.12	55
				R1281-GhRCA2	TCACTTCTAGCCGTTGGATCA	58.82	47.62
		Middle	1	F718-GhRCA2	CCCCATCATCGTCACTGGTA	58.88	55
				R856-GhRCA2	GCCGTCGGTCCTGAAAATAC	59	55
			2	F756-GhRCA2	ATGCTCCTCTCATCCGTGAC	59.25	55
				R893-GhRCA2	AGGTGTCGACGAGCTTAACA	59.04	50
Wild Cottons	RCA 1 (beta)	Terminal	1	F37-WiGRCA1	CGAGCACCGTTGAGTTTGAA	59.07	50
				R1324-WiGRCA1	ATCTCCAAGGGCAGCTTCAC	60.03	55
			2	F49-WiGRCA1	AGTTTGAATGGCTCTGGTGC	58.75	50
				R1329-WiGRCA1	TTAGCATCTCCAAGGGCAGC	60.11	55
		Middle	1	F641-WiGRCA1	GGCAAAGGTATCGTGAAGCT	58.26	50
				R740-WiGRCA1	CCACCAAATCTACCAGCACC	58.54	55
			2	F653-WiGRCA1	GTGAAGCTGCCGACATGATC	59.35	55
				R757-WiGRCA1	CCGTCATTTGAGTATTTCCACCA	58.99	43.48
	RCA 2 (alpha)	Terminal	1	F31-WiGRCA2	AGTTTCAAGGTTGTGGCTGC	59.25	50
				R1266-WiGRCA2	ACACCCAACTTGTTGTGCTG	59.18	50
			2	F8-WiGRCA2	ACAACAGCAAGGTTCCTTCG	58.69	50
				R1323-WiGRCA2	CTGGCCGTTGGATCAAAGTT	58.75	50
		Middle	1	F601-WiGRCA2	GTGCTGGTAGAATGGGAGGA	58.8	55
				R712-WiGRCA2	CCATACCGGGGATGCTAACT	58.94	55
			2	F602-WiGRCA2	TGCTGGTAGAATGGGAGGAAAT	59.15	45.45
				R724-WiGRCA2	CGCCAATCTTGTCCATACCG	59.06	55
N. tabacum	RCA beta 1	Terminal	1	F3-NtRCAB1	GGCTACCTCTGTCTCAACCA	58.73	55
				R1280-NtRCAB1	CCAAGTGCAGCCTCTTTGAG	59.12	55
			2	F73-NtRCAB1	ACTTCAGTTCCAAGCACAGC	58.69	50
				R1282-NtRCAB1	CACCAAGTGCAGCCTCTTT	58.28	52.63
		Middle	1	F611-NtRCAB1	GAGAGCCAGCCAAGTTGATT	58.17	50
				R729-NtRCAB1	TCCACCCATTCTACCAGCTC	58.8	55
			2	F613-NtRCAB1	GAGCCAGCCAAGTTGATTAGG	58.98	52.38
				R726-NtRCAB1	ACCCATTCTACCAGCTCCTG	58.79	55
	RCA beta 2	Terminal	1	F73-NtRCAB2	ACTTCAGTTCCAAGCACAGC	58.69	50
				R1284-NtRCAB2	ATCTCCAAGTGCTGCCTCTT	59.01	50
			2	F3-NtRCAB2	GGCTACCTCAGTGTCAACCA	59.31	55
				R1289-NtRCAB2	TTTGCATCTCCAAGTGCTGC	59.4	50
		Middle	1	F631-NtRCAB2	AGGCAAAGGTACAGAGAGGC	59.38	55
				R730-NtRCAB2	TTCCACCCATTCTACCAGCT	58.33	50
			2	F636-NtRCAB2	AAGGTACAGAGAGGCAGCAG	59.1	55
				R786-NtRCAB2	AGCAATGTTCATGAGAGTGGC	58.91	47.62
Wild Tobaccos	RCA beta 1	Terminal	1	F13-WiNRCAB1	GTCTCAACCATTGGAGCTGC	59.19	55
				R1309-WiNRCAB1	ATCTCCAAGTGCTGCCTCTT	59.01	50
			2	F3-WiNRCAB1	GGCTACCTCTGTCTCAACCA	58.73	55
				R1314-WiNRCAB1	TTTGCATCTCCAAGTGCTGC	59.4	50

Species	*Gene Name*	*Terminal/Middle*	*Pair no.*	*Primer ID (start site)*	*Primer Sequence (5'->3')*	*Tm (°C)*	*GC%*
		Middle	1	F609-WiNRCAB1	AAGTGGAAATGCAGGAGAGC	58.17	50
				R741-WiNRCAB1	TCCACCCATTCTACCAGCTC	58.8	55
			2	F648-WiNRCAB1	AAGGTACAGAGAGGCAGCAG	59.1	55
				R798-WiNRCAB1	AGCAATGTTCATGAGAGTGGC	58.91	47.62
	RCA beta 2	Terminal	1	F13-WiNRCAB2	GTCTCAACCATTGGAGCTGC	59.19	55
				R1311-WiNRCAB2	ATCTCCAAGTGCTGCCTCTT	59.01	50
			2	F3-WiNRCAB2	GGCTACCTCTGTCTCAACCA	58.73	55
				R1316-WiNRCAB2	TTTGCATCTCCAAGTGCTGC	59.4	50
		Middle	1	F643-WiNRCAB2	AGGCAAAGGTACAGAGAGGC	59.38	55
				R744-WiNRCAB2	TCCACCCATTCTACCAGCTC	58.8	55
			2	F609-WiNRCAB2	AAGTGGAAATGCAGGAGAGC	58.17	50
				R741-WiNRCAB2	ACCCATTCTACCAGCTCCTG	58.79	55
T. triandra	RCA	Terminal	1	F60-TtRCA	AGCAGCTTCCTCGGGATAAA	58.8	50
				R1412-TtRCA	TTCTGGCGGTAGGGAAGAAG	59.1	55
			2	F11-TtRCA	CCTTCTCCTCCACTCCGTC	58.81	63.16
				R1399-TtRCA	GAAGAAGTTGCCGGTCTTGG	59.12	55
		Middle	1	F665-TtRCA	ACCTCATCAAGAACCAGGGC	59.67	55
				R764-TtRCA	TTGTTCACCGTGTACTGGGT	59.17	50
			2	F689-TtRCA	TGTGCTGCCTGTTCATCAAC	59.05	50
				R794-TtRCA	ATGTTCATCAGGGTGGCGTT	59.96	50
Cottons	ADH	Terminal	1	F53-CottonADH	CAGGGAAGCCACTGGTGATA	59.09	55
				R1025-CottonADH	AAGGTCAGATCGGGGTTTGT	58.94	50
		Middle	1	F613-CottonADH	TTTTGGACTGGGAGCTGTTG	58.31	50
				R734-CottonADH	TCTTGGCTGCAACCTTACCT	59.23	50
Tobaccos	ADH	Terminal	1	F160-TobaccoADH	AAGCCAAGGGTGAAACTCCT	59.15	50
				R1145-TobaccoADH	CGCAACGTAGGCATTTTCC	57.69	52.63
		Middle	1	F639-TobaccoADH	GCAGAAGGAGCAAGGATTGG	58.9	55
				R741-TobaccoADH	TCACAAACTCAGTCACTCCAA	57.1	42.86
Themeda	ADH	Terminal	1	F143-PanicoidADH	ACACCGACGTCTACTTCTGG	59.12	55
				R1043-PanicoidADH	AACTTCTCCACCTCCAGCTC	59.02	55
		Middle	1	F601-PanicoidADH	TTCGGTCTAGGAGCTGTTGG	59.1	55
				R704-PanicoidADH	GCTTCTTCGAATCTGCTGGG	58.99	55

7.1.2 Supplementary Table 2: Re-designed primers.

Primers were re-designed for the following species-by-gene combinations due to poor amplification using the initial primers outlined in Supplementary Table 1.

Species	*Gene Name*	*Terminal/Middle*	*Pair no.*	*Primer ID (start site)*	*Primer Sequence (5'->3')*	*Tm (oC)*	*GC%*
Wild Cotton	RCA 2	Terminal	1	F63-GossRCA2	AGACGAAGAGACACAGACCG	59.12	55
				R1220-GossRCA2	GAACCCCAACTTGTTGTGCT	58.89	50
		Middle	1	F717-GossRCA2	CCCCATCATCGTCACTGGTA	58.88	55
				R855-GossRCA2	GCCGTCGGTCCTGAAAATAC	59	55
Wild Tobacco	RCA 2	Forward	1	F24-NicoRCA2	TGGAGCTGTCAACAAAGCAC	58.98	50
				R-1294-NicoRCA2	ATCTCCAAGTGCTGCCTCTT	59.01	50
		Middle	1	F716-NicoRCA2	GAGCTGGTAGAATGGGTGGA	58.8	55
				R835-NicoRCA2	GCTTGTTGTACATACCGGGG	58.62	55
Themeda	RCA	Forward	1	F305-ThemRCA	CCTACGAGTACCTCAGCCAG	58.97	60
				R1252-ThemRCA	GTTAGCGTCACCAAGAGCAG	58.93	55
		Middle	1	F57-ThemRCA	CCTGGGGAAGAAGCTGAAGA	59.01	55
				R946-ThemRCA	GTTGTCGGTGCGGAAGATAC	59	55
			2	F695-ThemRCA	ACCCAGTACACGGTGAACAA	59.17	50
				R846-ThemRCA	TACAGCGTGGAGAAGTCGTT	59.04	50

7.1.3 Supplementary Table 3: Protein IDs and corresponding target peptides

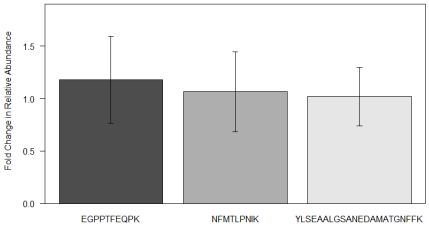
Protein ID In Databases	Protein Label	Protein Name	Peptides	Precursor (m/z)	Retention time (min)	Product ions used (Da)
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	DTDILAAFR	511.2693	54.4	577.3456
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	DTDILAAFR	511.2693	54.4	690.4297
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	DTDILAAFR	511.2693	54.4	805.4567
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	DTDILAAFR[+10]	516.2734	54.4	587.3539
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	DTDILAAFR[+10]	516.2734	54.4	700.438
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	DTDILAAFR[+10]	516.2734	54.4	815.4649
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	TFQGPPHGIQVER	489.2564	29.6	609.3229
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	TFQGPPHGIQVER	489.2564	29.6	701.394
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	TFQGPPHGIQVER	489.2564	29.6	935.5057
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	TFQGPPHGIQVER[+10]	492.5925	29.6	614.3271
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	TFQGPPHGIQVER[+10]	492.5925	29.6	711.4023
|Symbols:RBCL|ribulose-bisphosphate carboxylases|chrC:54958-56397	RbcL	Rubisco Large Subunit (RbcL)	TFQGPPHGIQVER[+10]	492.5925	29.6	945.514
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	SDDGTC[Pyr]TYK	547.7266	17.6	619.2908
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	SDDGTC[Pyr]TYK	547.7266	17.6	777.36
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	SDDGTC[Pyr]TYK	547.7266	17.6	892.3869
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	WIGEVGVATVGK	608.3402	42.4	730.4457
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	WIGEVGVATVGK	608.3402	42.4	916.5098
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	WIGEVGVATVGK	608.3402	42.4	1029.5939
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	WISEVGVASVGK	616.3377	40.4	845.4727
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	WISEVGVASVGK	616.3377	40.4	932.5048
Gohir.A06G166200.1.p_plus-mutant	RCA 1	Gossypium RCA Isoform 1	WISEVGVASVGK	616.3377	40.4	1045.5887
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	EGPPSFEQPTMTIEK	845.9033	41.4	819.428
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	EGPPSFEQPTMTIEK	845.9033	41.4	1076.5293
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	EGPPSFEQPTMTIEK	845.9033	41.4	1407.6825
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	WIGEVGTNSVGK	623.825	37.7	585.3031
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	WIGEVGTNSVGK	623.825	37.7	614.8196
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	WIGEVGTNSVGK	623.825	37.7	947.4792
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	WIGEVGVNSVGK	622.8353	39.2	759.4359
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	WIGEVGVNSVGK	622.8353	39.2	945.5
Gohir.A10G221700.1.p_plus-mutant	RCA 2	Gossypium RCA Isoform 2	WIGEVGVNSVGK	622.8353	39.2	1058.5841
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	SDDGTC[Pyr]TYQF	621.2426	32	948.3404
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	SDDGTC[Pyr]TYQF	621.2426	32	1039.4189
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	SDDGTC[Pyr]TYQF	621.2426	32	1076.3989
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	TDGIPDEDIVK	601.301	29.8	387.1874
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	TDGIPDEDIVK	601.301	29.8	815.4146
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	TDGIPDEDIVK	601.301	29.8	985.5201
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	TDGVPDEDIVK	594.2932	27.9	373.1718
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	TDGVPDEDIVK	594.2932	27.9	815.4146
Gohir.D06G171700.1.p_plus-mutants	RCA 2	Gossypium RCA Isoform 3	TDGVPDEDIVK	594.2932	27.9	971.5044
Gohir.D10G234000.1.p	RCA 2	Gossypium RCA Isoform 4	TDNVPVDDLVK	607.8168	34.9	785.4404
Gohir.D10G234000.1.p	RCA 2	Gossypium RCA Isoform 4	TDNVPVDDLVK	607.8168	34.9	884.5088
Gohir.D10G234000.1.p	RCA 2	Gossypium RCA Isoform 4	TDNVPVDDLVK	607.8168	34.9	998.5517
NA	Total RCA	Total RCA	FYWAPTR	470.7374	39.3	444.2565
NA	Total RCA	Total RCA	FYWAPTR	470.7374	39.3	630.3358
NA	Total RCA	Total RCA	FYWAPTR	470.7374	39.3	793.3992
NA	Total RCA	Labelled RCA	FYWAPTR[+10]	475.7415	39.3	454.2648
NA	Total RCA	Labelled RCA	FYWAPTR[+10]	475.7415	39.3	640.3441
NA	Total RCA	Labelled RCA	FYWAPTR[+10]	475.7415	39.3	803.4074
NA	Total RCA	Total RCA	VYDDEVR	448.2114	18.6	518.2569
NA	Total RCA	Total RCA	VYDDEVR	448.2114	18.6	633.2839
NA	Total RCA	Total RCA	VYDDEVR	448.2114	18.6	796.3472
NA	Total RCA	Labelled RCA	VYDDEVR[+10]	453.2156	18.6	528.2652
NA	Total RCA	Labelled RCA	VYDDEVR[+10]	453.2156	18.6	643.2921
NA	Total RCA	Labelled RCA	VYDDEVR[+10]	453.2156	18.6	806.3555
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	EAALGDANADAINNGSFFTR	685.324	50.9	828.3999
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	EAALGDANADAINNGSFFTR	685.324	50.9	942.4428
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	EAALGDANADAINNGSFFTR	685.324	50.9	999.4378
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	EAALGDANADAINNGSFFTS	992.9478	52.2	999.4378

Protein ID In Databases	Protein Label	Protein Name	Peptides	Precursor (m/z)	Retention time (min)	Product ions used (Da)
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	EAALGDANADAINNGSFFTS	992.9478	52.2	1112.5219
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	EAALGDANADAINNGSFFTS	992.9478	52.2	1226.5648
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	IVDSFPGQSIDFFGALR	934.9807	75.1	825.4254
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	IVDSFPGQSIDFFGALR	934.9807	75.1	1025.5415
Niben101Scf01653g02013.1_plus-mutant	RCA 1	Nicotiana RCA Isoform 1	IVDSFPGQSIDFFGALR	934.9807	75.1	1307.6743
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	DGPPTFEQPK	558.272	25.9	472.2478
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	DGPPTFEQPK	558.272	25.9	648.3351
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	DGPPTFEQPK	558.272	25.9	846.4356
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	VPLILGVWGGK	569.8528	65.2	716.409
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	VPLILGVWGGK	569.8528	65.2	829.493
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	VPLILGVWGGK	569.8528	65.2	942.5771
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	VQLAETYLK	532.803	37.9	653.3505
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	VQLAETYLK	532.803	37.9	724.3876
Niben101Scf05368g03015.1	RCA 2	Nicotiana RCA Isoform 2	VQLAETYLK	532.803	37.9	837.4716
Niben101Scf28395g00001.1	RCA 1	Nicotiana RCA Isoform 3	LLEYGNLLVQEQENVK	945.0044	58.2	874.4265
Niben101Scf28395g00001.1	RCA 1	Nicotiana RCA Isoform 3	LLEYGNLLVQEQENVK	945.0044	58.2	1086.579
Niben101Scf28395g00001.1	RCA 1	Nicotiana RCA Isoform 3	LLEYGNLLVQEQENVK	945.0044	58.2	1370.7274
Niben101Scf28395g00001.1	RCA 1	Nicotiana RCA Isoform 3	LLNSIDGPPTFEQPK	828.4356	44.1	428.2504
Niben101Scf28395g00001.1	RCA 1	Nicotiana RCA Isoform 3	LLNSIDGPPTFEQPK	828.4356	44.1	943.4883
Niben101Scf28395g00001.1	RCA 1	Nicotiana RCA Isoform 3	LLNSIDGPPTFEQPK	828.4356	44.1	1228.6208
tr|A0A1S4AW82|A0A1S4AW82_TOBAC	RCA 1	Nicotiana RCA Isoform 4	GLVADFSDDQQDITR	840.3972	43.7	1339.576
tr|A0A1S4AW82|A0A1S4AW82_TOBAC	RCA 1	Nicotiana RCA Isoform 4	GLVADFSDDQQDITR	840.3972	43.7	1224.5491
tr|A0A1S4AW82|A0A1S4AW82_TOBAC	RCA 1	Nicotiana RCA Isoform 4	GLVADFSDDQQDITR	840.3972	43.7	1410.6132
tr|A0A1S4AW82|A0A1S4AW82_TOBAC	RCA 1	Nicotiana RCA Isoform 4	WISGAGIEK	480.7611	30.2	774.4356
tr|A0A1S4AW82|A0A1S4AW82_TOBAC	RCA 1	Nicotiana RCA Isoform 4	WISGAGIEK	480.7611	30.2	300.1707
tr|A0A1S4AW82|A0A1S4AW82_TOBAC	RCA 1	Nicotiana RCA Isoform 4	WISGAGIEK	480.7611	30.2	574.3195
sp|OVAL_CHICK	NA	Ovalbumin	DEDTQAMPFR	605.2639	30.4	550.2806
sp|OVAL_CHICK	NA	Ovalbumin	DEDTQAMPFR	605.2639	30.4	621.3177
sp|OVAL_CHICK	NA	Ovalbumin	DEDTQAMPFR	605.2639	30.4	965.4509
sp|OVAL_CHICK	NA	Labelled Ovalbumin	DEDTQAMPFR[+10]	610.268	30.4	560.2889
sp|OVAL_CHICK	NA	Labelled Ovalbumin	DEDTQAMPFR[+10]	610.268	30.4	631.326
sp|OVAL_CHICK	NA	Labelled Ovalbumin	DEDTQAMPFR[+10]	610.268	30.4	975.4592
sp|OVAL_CHICK	NA	Ovalbumin	GGLEPINFQTAADQAR	844.4235	45.3	1007.4905
sp|OVAL_CHICK	NA	Ovalbumin	GGLEPINFQTAADQAR	844.4235	45.3	1121.5334
sp|OVAL_CHICK	NA	Ovalbumin	GGLEPINFQTAADQAR	844.4235	45.3	1331.6703
sp|OVAL_CHICK	NA	Labelled Ovalbumin	GGLEPINFQTAADQAR[+10]	849.4277	45.3	1017.4988
sp|OVAL_CHICK	NA	Labelled Ovalbumin	GGLEPINFQTAADQAR[+10]	849.4277	45.3	1131.5416
sp|OVAL_CHICK	NA	Labelled Ovalbumin	GGLEPINFQTAADQAR[+10]	849.4277	45.3	1341.6785
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	NFMTLPNIK	539.2917	51.5	584.3766
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	NFMTLPNIK	539.2917	51.5	685.4243
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	NFMTLPNIK	539.2917	51.5	816.4648
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	EGPPTFEQPK	565.2798	26.9	648.3351
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	EGPPTFEQPK	565.2798	26.9	749.3828
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	EGPPTFEQPK	565.2798	26.9	846.4356
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	YLSEAALGSANEDAMATGNFFK	769.6915	60.4	915.4393
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	YLSEAALGSANEDAMATGNFFK	769.6915	60.4	986.4764
Sobic.005G231500.1.p	RCA	Themeda RCA Isoform 1	YLSEAALGSANEDAMATGNFFK	769.6915	60.4	1101.5033

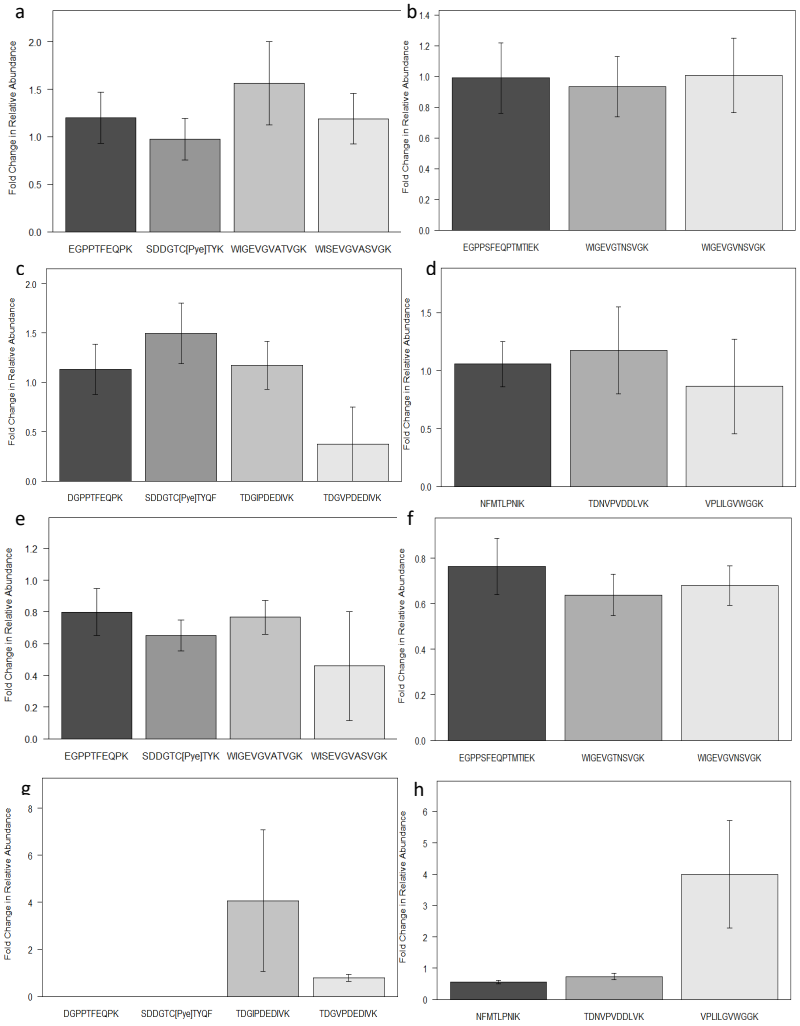
7.2 Results

7.2.1 Themeda, Gossypium, and Nicotiana Peptide Fold Changes

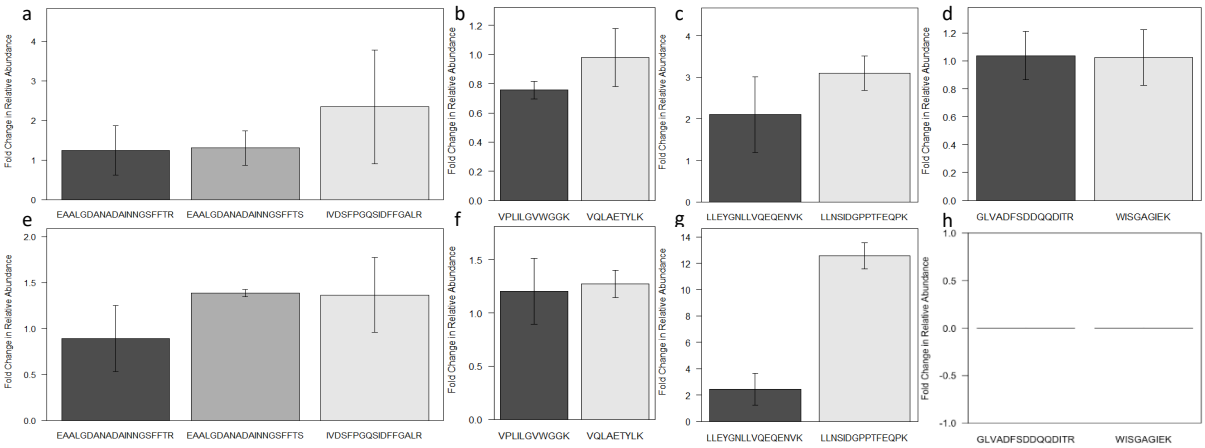
These are mean (± SE, n = 3) peptide fold changes of the relative abundance of the unique target peptides ascribed to the single Themeda RCA isoform, and the four RCA isoforms of *Gossypium* (a-d = *G. hirsutum*; e-g = *G. robinsonii*) and *Nicotiana* (a-d = *N. tabacum*; e-g = *N. megalosiphon*; as outlined in Supplementary Material Table 3), respectively, at high temperature (38^{o}C) relative to control temperature (30^{o}C).



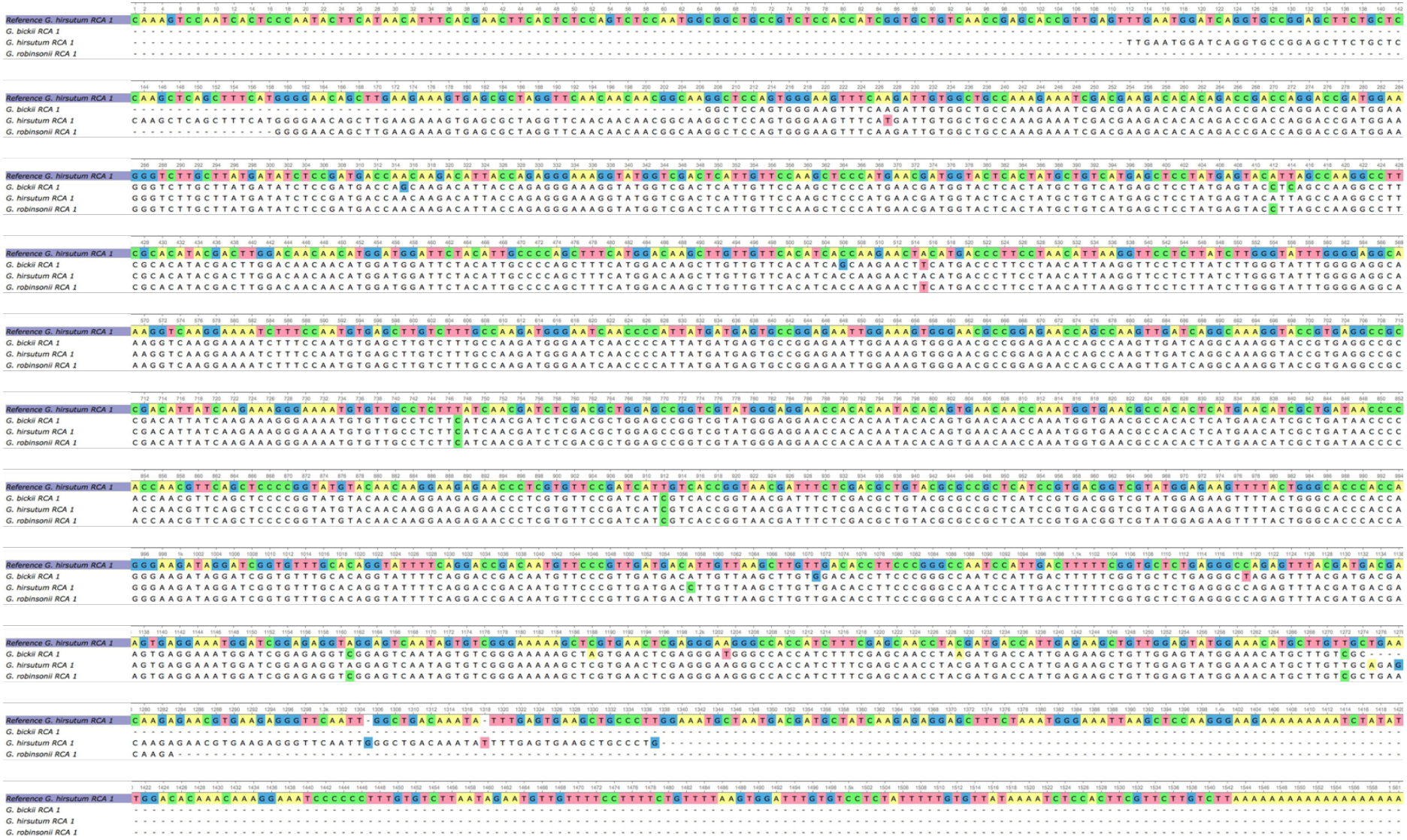
7.2.2 Gossypium Peptide Fold Change



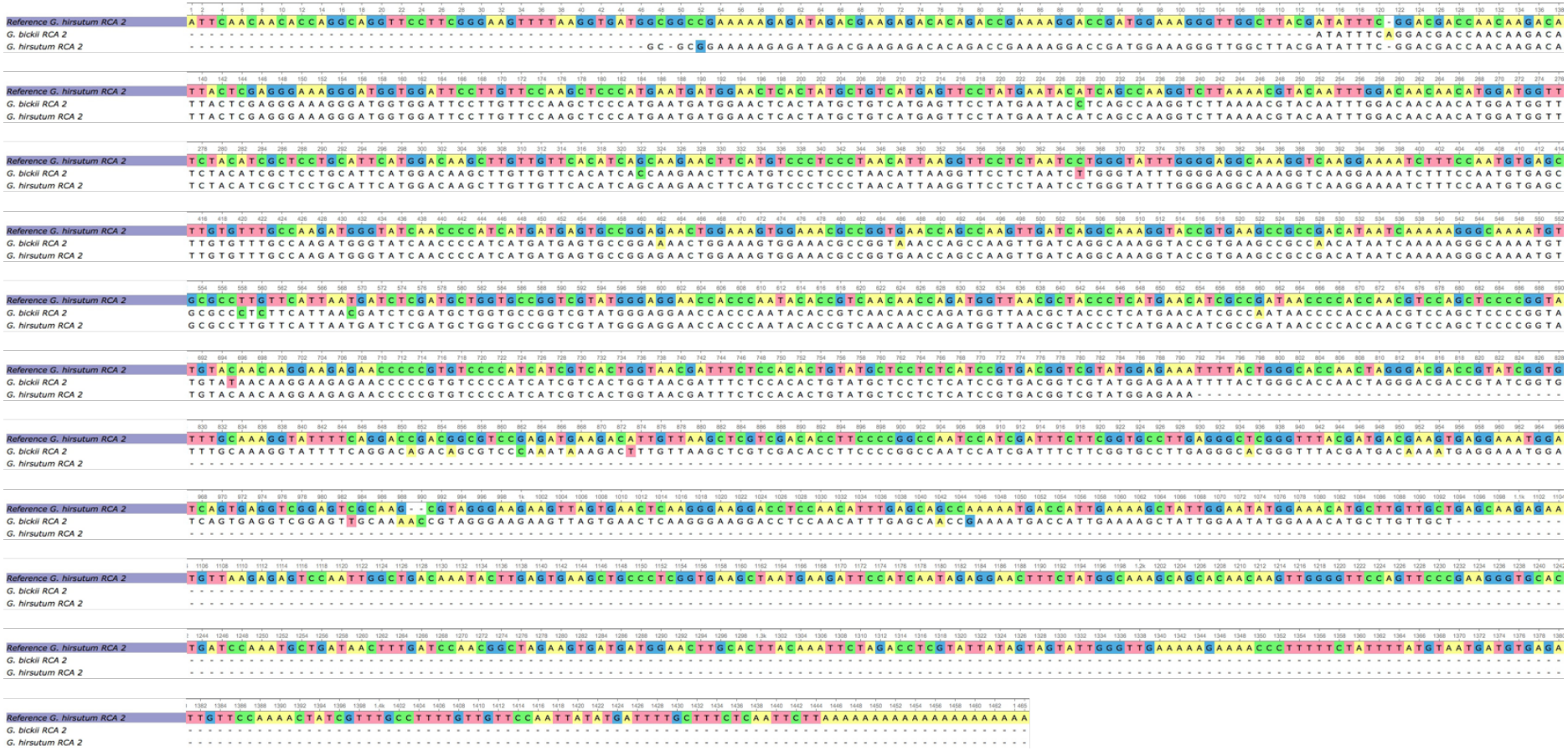
7.2.3 Nicotiana Peptide Fold Change



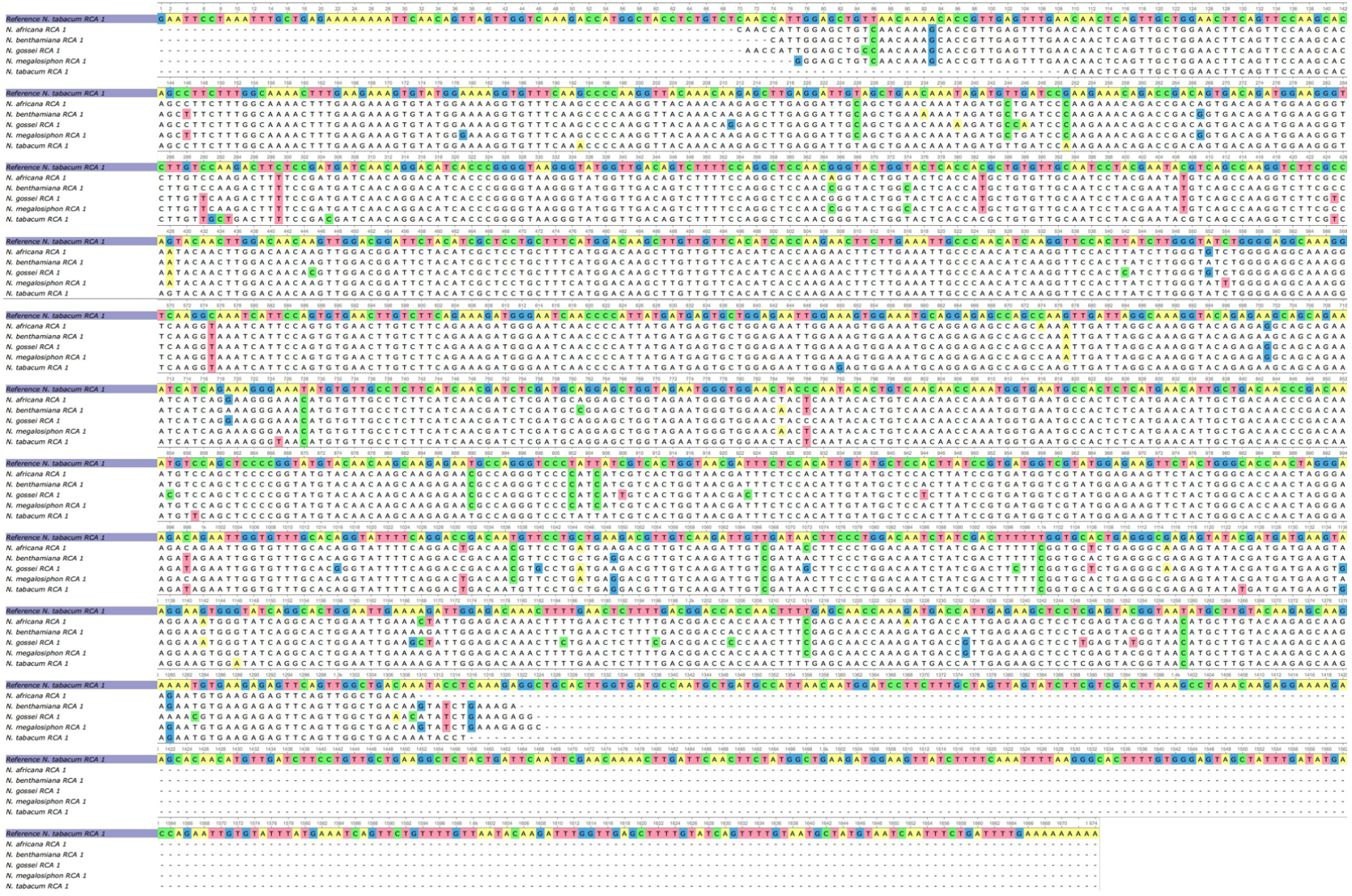
7.2.4 Gossypium RCA 1 gene sequence



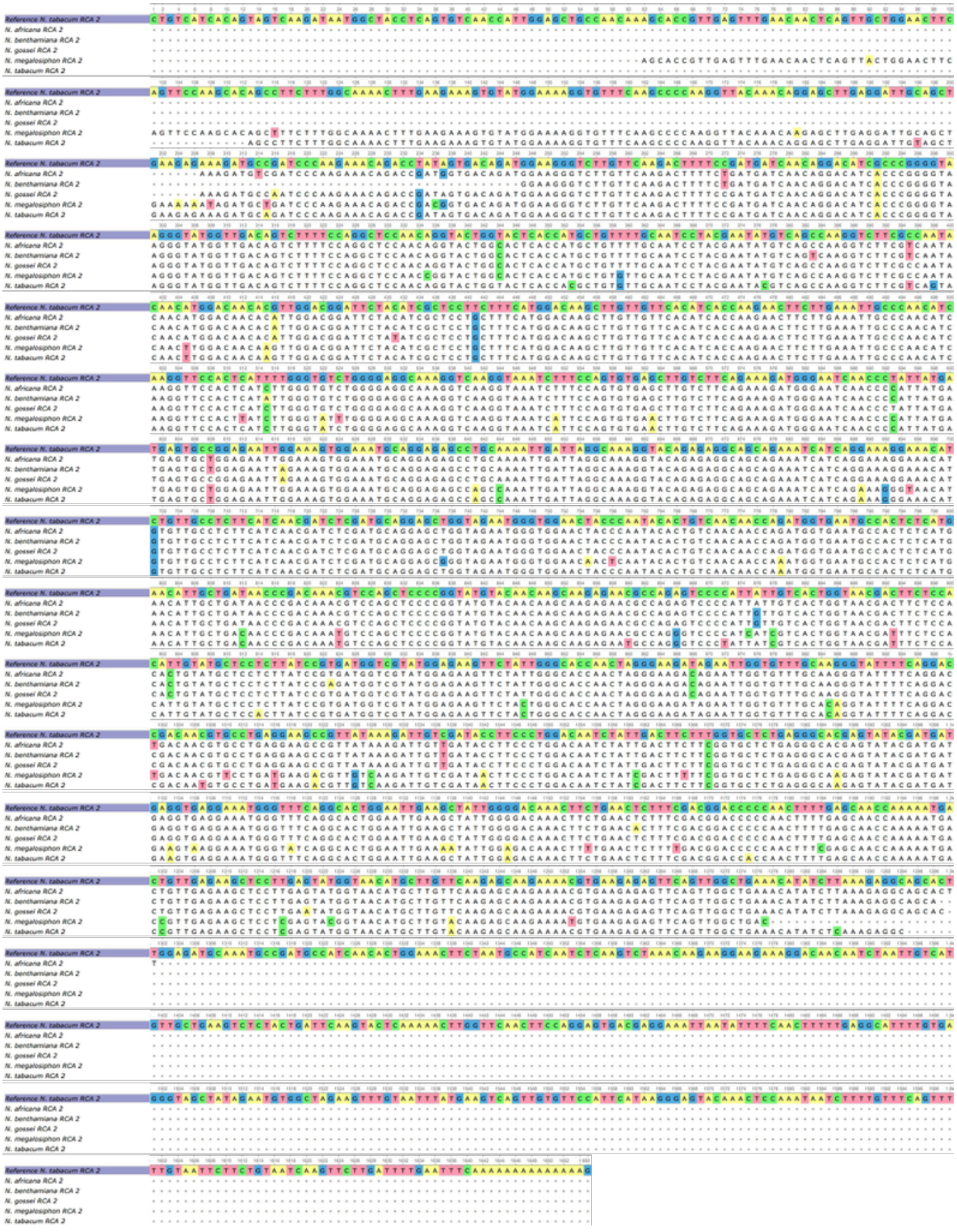
7.2.5 Gossypium RCA 2 gene sequence



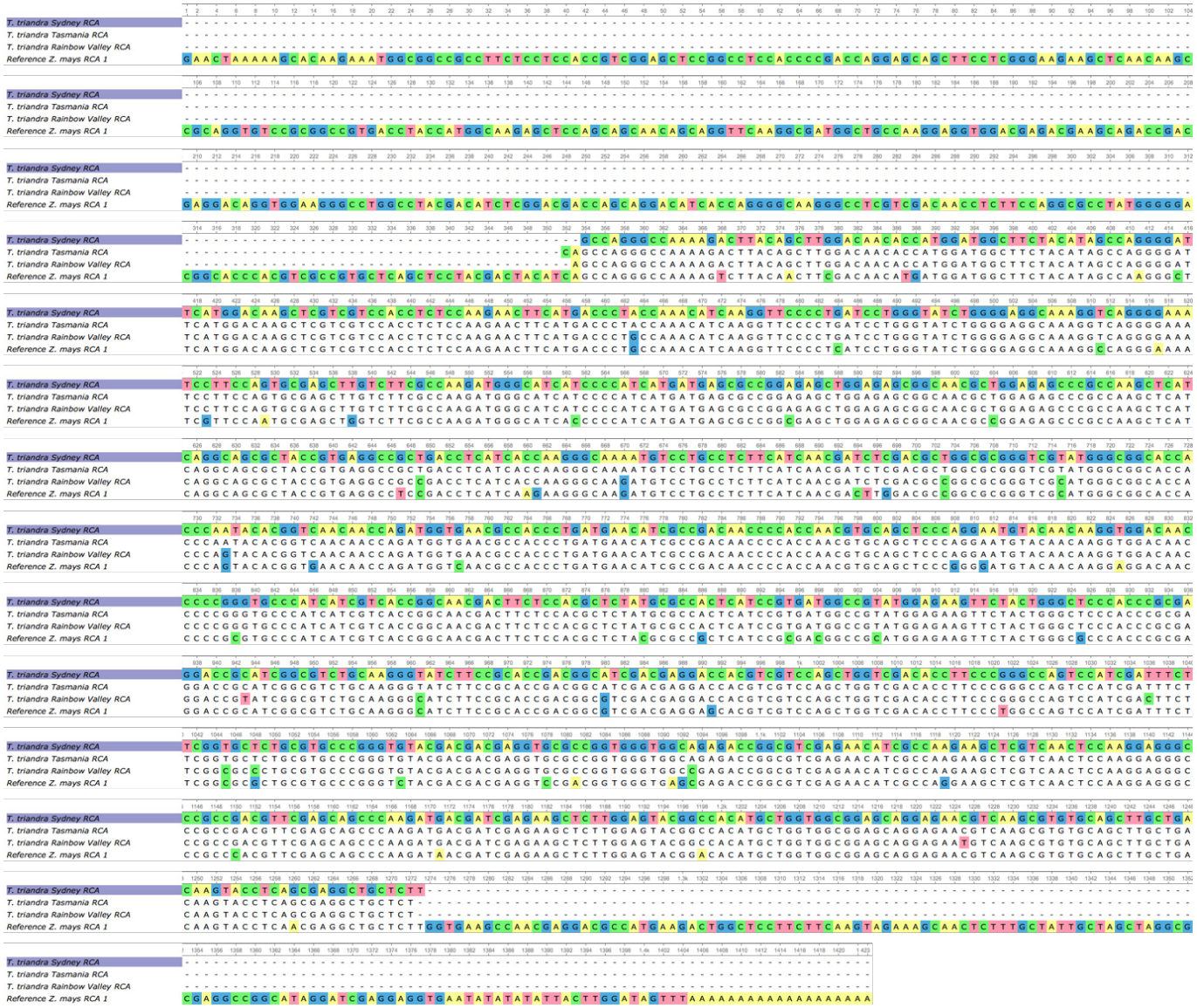
7.2.6 Nicotiana RCA 1 gene sequence



7.2.7 Nicotiana RCA 2 gene sequence



7.2.8 Themeda RCA gene sequences



7.2.9 Supplementary Table 4: Double base calls leading to amino acid substitutions in RCA sequences

Species	Gene	Position In Ref. Sequence	Base Calls	AA Substitution
G. bickii	RCA 2	133	A/C	Gln -> His
G. bickii	RCA 2	151	A/T	Lys -> Asn
G. bickii	RCA 2	157	A/G	Met -> Ile
G. bickii	RCA 2	159	T/C	Val -> Ala
G. bickii	RCA 2	174	A/G	Gln -> Arg
G. bickii	RCA 2	176	T/G	Ala -> Ser
G. bickii	RCA 2	177	C/G	Ala -> Gly
G. bickii	RCA 2	179	A/C	Pro -> Thr
G. hirsutum	RCA 1	215	T/C	Val -> Ala
G. hirsutum	RCA 1	227	A/T	Met -> Lys
G. hirsutum	RCA 1	412	A/C	Ile -> Leu
G. hirsutum	RCA 1	515	A/T	Tyr -> Phe
G. hirsutum	RCA 1	556	A/G	Ile -> Val
G. hirsutum	RCA 1	1057	A/C	Leu -> Ile
G. hirsutum	RCA 1	1171	A/G	Ser -> Gly
G. hirsutum	RCA 1	1223	A/C	Gln -> Pro
N. africana	RCA 1	85	T/C	Val -> Ala
N. africana	RCA 1	120	A/G	Ala -> Thr
N. africana	RCA 1	238	A/T	Ile -> Lys
N. africana	RCA 1	246	A/G	Asp -> Asn
N. africana	RCA 1	267	A/G	Ser -> Gly
N. africana	RCA 1	435	A/T	Leu -> Met
N. africana	RCA 1	445	A/C	Lys -> Thr
N. africana	RCA 1	446	T/G	Lys -> Asn
N. africana	RCA 1	552	A/G	Val -> Ile
N. africana	RCA 1	1045	A/C	Asp -> Ala
N. africana	RCA 1	1056	A/G	Val -> Ile
N. africana	RCA 1	1060	A/G	Lys -> Arg
N. africana	RCA 1	1167	A/G	Thr -> Ala
N. africana	RCA 1	1168	A/C	Thr -> Asn
N. africana	RCA 1	1233	A/G	Ile -> Val
N. africana	RCA 2	214	T/C	Val -> Ala
N. africana	RCA 2	237	A/G	Gly -> Ser
N. africana	RCA 2	876	A/G	Ile -> Val
N. africana	RCA 2	1030	A/G	Lys -> Arg
N. benthamiana	RCA 1	85	T/C	Val -> Ala
N. benthamiana	RCA 1	120	A/G	Ala -> Thr
N. benthamiana	RCA 1	226	T/C	Ala -> Val
N. benthamiana	RCA 1	1045	A/C	Ala -> Asp
N. benthamiana	RCA 1	1072	A/G	Asn -> Ser
N. benthamiana	RCA 1	1167	A/G	Lys -> Glu
N. benthamiana	RCA 1	1168	A/C	Lys -> Thr
N. benthamiana	RCA 1	1169	T/G	Lys -> Asn
N. benthamiana	RCA 1	1233	A/G	Ile -> Val
N. benthamiana	RCA 2	876	A/G	Val -> Ile
N. benthamiana	RCA 2	886	A/G	Gly -> Asp
N. benthamiana	RCA 2	1161	A/T	Thr -> Ser
N. gossei	RCA 1	85	T/C	Ala -> Val
N. gossei	RCA 1	159	A/T	Thr -> Ser
N. gossei	RCA 1	211	A/G	Arg -> Lys
N. gossei	RCA 1	234	C/G	Gln -> Glu
N. gossei	RCA 1	238	A/T	Lys -> Ile
N. gossei	RCA 1	246	A/G	Asn -> Asp
N. gossei	RCA 1	445	A/C	Thr -> Lys
N. gossei	RCA 1	552	A/G	Val -> Ile
N. gossei	RCA 1	906	A/G	Ile -> Val
N. gossei	RCA 1	1015	A/C	Thr -> Lys
N. gossei	RCA 1	1045	A/C	Asp -> Ala
N. gossei	RCA 1	1051	A/C	Asp -> Ala
N. gossei	RCA 1	1056	A/G	Val -> Ile
N. gossei	RCA 1	1167	A/G	Ala -> Thr
N. gossei	RCA 1	1168	A/C	Ala -> Asp
N. gossei	RCA 1	1307	A/C	Glu -> Asp
N. gossei	RCA 2	1161	A/T	Ser -> Thr
N. megalosiphon	RCA 1	120	A/G	Ala -> Thr
N. megalosiphon	RCA 1	211	A/G	Lys -> Arg
N. megalosiphon	RCA 1	234	C/G	Gln -> Glu
N. megalosiphon	RCA 1	235	A/G	Gln -> Arg
N. megalosiphon	RCA 1	238	A/T	Ile -> Lys
N. megalosiphon	RCA 1	246	A/G	Asp -> Asn
N. megalosiphon	RCA 1	435	A/T	Phe -> Ile
N. megalosiphon	RCA 1	445	A/C	Lys -> Thr
N. megalosiphon	RCA 1	552	A/G	Ile -> Val
N. megalosiphon	RCA 1	1045	A/C	Asp -> Ala
N. megalosiphon	RCA 1	1060	A/G	Lys -> Arg
N. megalosiphon	RCA 1	1072	A/G	Asn -> Ser
N. megalosiphon	RCA 1	1167	A/G	Lys -> Glu

Species	Gene	Position In Ref. Sequence	Base Calls	AA Substitution
N. megalosiphon	RCA 1	1168	A/C	Lys -> Thr
N. megalosiphon	RCA 1	1169	T/G	Lys -> Asn
N. megalosiphon	RCA 1	1195	A/T	Phe -> Tyr
N. megalosiphon	RCA 1	1233	A/G	Val -> Ile
N. megalosiphon	RCA 2	90	A/G	Thr -> Ala
N. megalosiphon	RCA 2	105	T/C	Pro -> Ser
N. megalosiphon	RCA 2	522	A/G	Ile -> Val
N. megalosiphon	RCA 2	1021	A/C	Asp -> Ala
N. megalosiphon	RCA 2	1026	A/G	Val -> Ile
N. megalosiphon	RCA 2	1042	A/C	Asn -> Thr
N. megalosiphon	RCA 2	1137	A/G	Asn -> Asp
N. megalosiphon	RCA 2	1138	A/C	Asn -> Thr
N. megalosiphon	RCA 2	1139	T/G	Asn -> Lys
N. megalosiphon	RCA 2	1203	A/G	Val -> Ile
N. tabacum	RCA 1	193	A/G	Asn -> Ser
N. tabacum	RCA 1	234	C/G	Gln -> Glu
N. tabacum	RCA 1	244	T/C	Val -> Ala
N. tabacum	RCA 1	292	A/C	Ala -> Asp
N. tabacum	RCA 1	1146	A/G	Ile -> Val
N. tabacum	RCA 1	1155	A/G	Thr -> Ala
N. tabacum	RCA 2	163	A/G	Ser -> Asn
N. tabacum	RCA 2	181	A/G	Arg -> Lys
N. tabacum	RCA 2	196	T/C	Val -> Ala
N. tabacum	RCA 2	263	A/T	Gln -> His
N. tabacum	RCA 2	291	A/G	Thr -> Ala
N. tabacum	RCA 2	405	A/T	Leu -> Met
N. tabacum	RCA 2	415	A/C	Lys -> Thr
N. tabacum	RCA 2	441	T/G	Ala -> Ser
N. tabacum	RCA 2	522	A/G	Ile -> Val
N. tabacum	RCA 2	985	A/C	Thr -> Lys
N. tabacum	RCA 2	1026	A/G	Val -> Ile
N. tabacum	RCA 2	1042	A/C	Asn -> Thr
N. tabacum	RCA 2	1116	A/G	Val -> Ile
N. tabacum	RCA 2	1125	A/G	Thr -> Ala
N. tabacum	RCA 2	1137	A/G	Ala -> Thr
N. tabacum	RCA 2	1138	A/C	Ala -> Asp
N. tabacum	RCA 2	1203	A/G	Val -> Ile
N. tabacum	RCA 2	1277	A/C	Glu -> Asp
N. tabacum	RCA 2	1279	A/C	Thr -> Lys
T. triandra (Rainbow Valley)	RCA	443	T/G	Ser -> Ala
T. triandra (Rainbow Valley)	RCA	950	A/G	Val -> Ile