

Metabolic Engineering Strategies to Increase Ascorbate Concentrations in Rice and Wheat

Ronan Charles Broad

orcid.org/0000-0002-2977-4718

A thesis submitted in total fulfilment for the degree of
Doctor of Philosophy

February 2020

The School of BioSciences
The University of Melbourne
Australia

Abstract

Ascorbate (ascorbic acid, vitamin C) is essential for both plants and mammals. Ascorbate is a reducing agent capable of donating electrons, enabling it to perform a range of biochemical functions, such as scavenging reactive oxygen species, assisting enzymatic activity, and reducing higher oxidative states of iron (Fe). In plants, ascorbate is the most abundant water-soluble antioxidant and plays a key role in many fundamental processes, such as photosynthesis, stress tolerance, and the transport of Fe. In humans, ascorbate is an essential micronutrient that must be obtained through diet and takes part in a range of important physiological functions, such as collagen synthesis, epigenetic programming, and Fe uptake in human digestion. Several pathways towards ascorbate biosynthesis have been proposed in plants, but there is only definitive evidence for the L-galactose pathway. The *GDP-L-galactose phosphorylase* (*GGP* or *vtc2/5*) gene encodes the first-committed and rate-limiting enzymatic step of the L-galactose pathway and represents the most promising candidate for increasing ascorbate biosynthesis in plants. In addition to transcriptional regulation, the translation of *GGP* is regulated through a highly conserved, *cis*-acting upstream open reading frame (uORF) in the 5' leader sequence of the *GGP* mRNA. Developing strategies to increase ascorbate biosynthesis in rice (*Oryza sativa* L.) and wheat (*Triticum aestivum* L.), two of the world's most important staple crops, has the potential to improve both food security and crop productivity. As part of this PhD project, two distinct metabolic engineering strategies were employed to increase ascorbate concentrations in rice: (i) constitutive overexpression of the *OsGGP* coding sequence (35S-*OsGGP* plants), and (ii) CRISPR/Cas9-targeted mutagenesis of the *OsGGP* uORF (*uorf_{OsGGP}* mutants). Ascorbate levels were negligible in both 35S-*OsGGP* and *uorf_{OsGGP}* brown rice, likely due to the decline of ascorbate levels in maturing grain reported in cereals; highlighting the challenge of increasing ascorbate levels in cereal species, such as rice. Ascorbate concentrations were significantly increased in germinated brown rice and tissues of 35S-*OsGGP* plants at the vegetative growth phase. In contrast, foliar ascorbate concentrations were significantly reduced in 35S-*OsGGP* plants at the reproductive growth phase. This was dependent on homozygosity of the 35S-*OsGGP* transgene and was associated with a significant reduction in endogenous *OsGGP* transcript levels, suggesting the occurrence of gene silencing. Foliar ascorbate concentrations were significantly increased in *uorf_{OsGGP}* mutants, without any changes to *OsGGP* transcript levels, attributed to alleviated ribosomal stalling on the *OsGGP* uORF

and enhanced translation of the *GGP* major ORF. Editing the *GGP* uORF represents an effective transgene-free strategy to increase ascorbate concentrations not only in rice, but other species. Challenging convention, automated imaging revealed that neither the 35S-*OsGGP* nor the *uorf_{OsGGP}* plants displayed increased salt tolerance at the vegetative growth phase, despite having elevated ascorbate levels. Ascorbate concentrations were positively correlated with ferritin concentrations in Caco-2 cells—an accurate predictor of Fe uptake in human digestion—exposed to *in vitro* digests of null segregant and 35S-*OsGGP* brown rice and germinated brown rice, suggesting that ascorbate-enriched crops may be able to improve Fe bioavailability in human diets. Grain Fe concentrations were not changed in the *uorf_{OsGGP}* mutants, indicating that ascorbate may not facilitate the transport of Fe into developing rice grain. Next, this PhD project identified six *TaGGP* genes in the hexaploid bread wheat genome, each with a highly conserved uORF in the 5' leader sequence. Phylogenetic analyses demonstrated that the *TaGGP* genes separate into two distinct groups and identified a duplication event of the *GGP* gene in the ancestor of the *Brachypodium*/Triticeae lineage. A microsynteny analysis revealed that the *TaGGP1* and *TaGGP2* subchromosomal regions have no shared synteny suggesting that *TaGGP2* may have been duplicated via a transposable element. A transcript analysis of the *TaGGP* genes identified that the *TaGGP1* homoeologs were broadly expressed across different tissues and developmental stages and that the *TaGGP2* homoeologs were highly expressed in anthers. Finally, transient transformation of the *TaGGP* coding sequences in *Nicotiana benthamiana* significantly increased foliar ascorbate concentrations more than five-fold, confirming their activity toward ascorbate biosynthesis *in planta*. The six *TaGGP* genes and uORFs identified in this study present an opportunity to fine-tune ascorbate biosynthesis in this important staple crop.

Declaration

I hereby declare that (i) this thesis comprises only original work towards the degree of the Doctor of Philosophy except where indicated in the preface; (ii) that due acknowledgement has been made in the text to all other material used; and (iii) this thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Ronan Charles Broad

February 2020

Preface

(i) Work towards the thesis that was carried out in collaboration with others:

Quantification of ascorbate (Chapters 2, 3, and 4) was carried out in collaboration with Sally Roden and Pawel Sadowski at the Queensland University of Technology, Brisbane, QLD, Australia.

Assessment of salt tolerance and plant growth and development using automated imaging (Chapters 2 and 3) was carried out in collaboration with Nathaniel Jewell, Chris Brien, and Bettina Berger at The Plant Accelerator, Adelaide, SA, Australia.

The assessment of grain iron bioavailability (Chapter 2) was carried out in collaboration with Jesse Beasley (University of Melbourne, Melbourne, VIC, Australia) in the laboratory of Raymond Glahn at the Robert W. Holley Center for Agriculture and Health, USDA-ARS, Ithaca, NY, United States.

Construction of the CRISPR/Cas9 vectors (Chapter 3) was carried out in collaboration with Fatima Naim and Peter Waterhouse at the Queensland University of Technology, Brisbane, QLD, Australia.

Assessment of the agro-morphological performance of the *uorf_{OsGGP}* mutants (Chapter 3) was carried out in collaboration with Sonia Sharma (University of Melbourne, Melbourne, VIC, Australia).

The syntenic analysis of the *TaGGP* genes in the bread wheat genome and analysis of the *TaGGP* promoter regions (Chapter 4) were carried out in collaboration with Ute Baumann at the University of Adelaide, Adelaide, SA, Australia.

Transient transformation of the *TaGGP* genes into *Nicotiana benthamiana* (Chapter 4) was carried out in collaboration with Sally Roden and Joshua Philips at the Queensland University of Technology, Brisbane, QLD, Australia.

(ii) Work that was carried out prior to enrolment in the degree:

Transformation of rice plants with the 35S-*OsGGP* T-DNA (Chapter 2) was carried out prior to my enrolment.

The identification of the six *TaGGP* genes in the bread wheat genome and analysis of the *TaGGP* transcripts in bread wheat tissues (Chapter 4) were carried out prior to my enrolment.

(iii) Publication status of research chapters:

Chapter 2. Effect of *OsGGP* Constitutive Overexpression in Rice on Ascorbate Biosynthesis, Stress Tolerance, and Iron Bioavailability

Submitted for publication to *Plants* on 26.02.20.

Author contributions: Conceptualization, J.P.B., R.P.H., and A.A.T.J.; data curation, N.J. and C.B.; formal analysis, R.C.B., N.J., and C.B.; funding acquisition, R.C.B. and R.P.G.; investigation, R.C.B., J.P.B., J.T.B., S.R., and P.S.; methodology, R.C.B., S.R., P.S., N.J., C.B., and B.B.; project administration, E.T., R.P.G., and B.B.; resources, R.P.G., R.P.H., and A.A.T.J.; software, N.J., and C.B.; supervision, J.P.B., R.P.H., and A.A.T.J.; visualization, R.C.B.; writing – original draft, R.C.B.; writing – review and editing, J.P.B., J.T.B., S.R., P.S., N.J., C.B., B.B., E.T., R.P.G., R.P.H., and A.A.T.J.

Chapter 3. Production and Characterization of Rice with CRISPR/Cas9-Induced Mutations in the *OsGGP* uORF

Unpublished material not submitted for publication.

Chapter 4. Genome-Wide Identification and Characterization of the *GDP-L-Galactose Phosphorylase* Gene Family in Bread Wheat

Published by *BMC Plant Biology* on 26.11.19 (Appendix 4)

Author contributions: JPB, RPH, AATJ conceived and designed the research. JPB identified the GGP genes and uORFs and RCB visualized their structures. RCB sequenced the *TaGGP* genes. RCB and JTB conducted the phylogenetic analyses. JPB and UB conducted the microsynteny analysis. JTB confirmed the specificity of the subgenome primers for the quantitative reverse transcription-PCR analysis and analysed the data. JPB analysed the supplementary gene expression data. JPB, RCB, and UB analysed the promoter regions. RCB cloned the *TaGGP* coding sequences. RCB, SR, and JGP conducted the transient transformation of *N. benthamiana* and quantified the ascorbate concentrations. JPB, RPH, and AATJ supervised the research. RCB drafted the manuscript. All authors read and approved the manuscript.

(iv) Sources of funding:

This research project was supported by grants from the following agencies:

- The Australian Plant Phenomics Facility Postgraduate Student Internship Award to Ronan Broad.
- The United States Department of Agriculture, Agricultural Research Service to Raymond Glahn.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my primary supervisor Alexander Johnson for taking me on board as part of his wonderful research team and for granting me the flexibility to pursue the research that I was most interested in. While there were many highs and lows over the course of my PhD project, I feel as though working with you over the past four years has left me well-equipped for a future in research; you have taught me many valuable skills in patience, resilience, writing, presenting, and scientific rigor. I know that we will continue to write and work together going forward, and I thoroughly look forward to it.

I would like to thank my co-supervisors, Julien Bonneau and Martin O'Brien, for your supervision and invaluable assistance in laboratory techniques—in particular, Julien's assistance in qRT-PCR and Martin's assistance with tissue culture. The two of you are some of the most hard-working and dedicated individuals I have ever met, and I thank you for all your support over the past four years.

I would also like to express my sincere gratitude to my external supervisor Roger Hellens, for both hosting my visits to the Queensland University of Technology, and for the brilliant ideas you have contributed towards my PhD project. It's been a pleasure working with you, and, again, I look forward to continuing to write and work with you in the future.

A big thank-you to all the past and present members of the Johnson lab for your wonderful company, assistance in the lab and greenhouses, and, most importantly, your party antics: Ouda Khammy, Tonatiuh Sanchez-Palacios, Laura Moreno-Moyano, Marianne Weisser, Jesse Beasley, Oscar Fung, Rucha Patil, Amy Liu, Natasha Nyangoma, Mayu Yamamoto, Sonia Sharma, Ruby Schofield, and Evangeline Kielnhofer. I look forward to our paths crossing again in the future.

Next, I need to thank all my collaborators, without whom this PhD project would not be possible: Sally Roden, Pawel Sadowski, Joshua Philips, Fatima Naim, and Peter Waterhouse from the Queensland University of Technology; Ute Baumann from The University of Adelaide; Nathaniel Jewell, Chris Brien, and Bettina Berger from The Plant Accelerator; and Elad Tako and Ray Glahn from the United States Department of Agriculture.

I would like to thank my mother, Heather McQuillan, for her constant love and support as I pursue my career in research—despite being across the Tasman. I look forward to staying with you back in Christchurch, before I embark on the next phase of my life.

Thanks to my brother, Dan Broad, and his partner Chloe Youngson, for the steady stream of Facebook messages since the arrival of my wee nephew Arthur Broad. They were truly heart-warming and motivating, and I also can't wait to spend some quality family time with you all back in Christchurch.

A very special thanks to my amazing partner, Jess Oldfield, for her love, support, encouragement, and her assistance in revising my writing—despite all the scientific jargon—over the past four years. I am so grateful to have found you and I cannot wait for the next phase of our lives together.

Finally, I dedicate this thesis to my late father Tim Broad. I know I have made him proud.

Table of Contents

Abstract	iii
Declaration	v
Preface	vi
Acknowledgements	ix
Table of Contents	xi
List of Figures	xvi
List of Tables	xviii
List of Abbreviations	xx
 Chapter 1. Metabolic Engineering Strategies to Increase Ascorbate Concentrations in Plants	 1
1.1. Introduction	2
1.2. Role of Ascorbate in Plants	4
1.3. Role of Ascorbate in Humans	7
1.4. Ascorbate Concentrations in Crops	9
1.5. Biosynthesis of Ascorbate in Plants	11
1.5.1. The L-Galactose Pathway	11
1.5.1.1. Phosphomannose Isomerase (PMI)	11
1.5.1.2. Phosphomannose Mutase (PMM)	17
1.5.1.3. GDP-D-Mannose Pyrophosphorylase (GMP)	17
1.5.1.4. GDP-D-Mannose-3',5'-Epimerase (GME)	18
1.5.1.5. GDP-L-Galactose Phosphorylase (GGP)	18
1.5.1.6. L-Galactose-1-Phosphate Phosphatase (GPP)	19
1.5.1.7. L-Galactose Dehydrogenase (L-GalDH)	20
1.5.1.8. L-Galactono-1,4-Lactone Dehydrogenase (L-GalLDH)	20
1.5.2. The L-Gulose Pathway	21
1.5.3. The Myo-Inositol Pathway	21
1.5.4. The D-Galacturonate Pathway	23
1.6. Recycling of Ascorbate in Plants	24
1.7. Regulation of Ascorbate Biosynthesis in Plants	28

1.7.1.	Ascorbic Acid Mannose Pathway Regulator 1 (AMR1)	28
1.7.2.	Basic Helix-Loop-Helix 59 Transcription Factor (bHLH59)	28
1.7.3.	Calmodulin-Like 10 (CML10)	31
1.7.4.	COP9 Signalosome Subunit 5B and 8 (CSN5B and CSN8)	31
1.7.5.	DNA-Binding with One Finger 22 (Dof22)	31
1.7.6.	Ethylene Response Factor 98 (ERF98)	32
1.7.7.	<i>GGP</i> Upstream Open Reading Frame (<i>GGP</i> uORF)	32
1.7.8.	HD-Zip I Family Transcription Factor 24 (HZ24)	33
1.7.9.	High-Pigment-1 (HP1)	33
1.7.10.	KONJAC 1 and 2 (KJC1 and KJC2)	33
1.7.11.	Myeloblastosis Transcription Factor 5 (MYB5)	34
1.7.12.	NBS-LRR 33 (NL33)	34
1.7.13.	Non-specific Lipid Transfer Protein-1 (nsLTP1)	34
1.7.14.	VITAMIN C-3 (VTC3)	35
1.7.15.	WAX1	35
1.7.16.	Zinc-finger 3 (ZF3)	35
1.8.	Rice and Wheat as Targets for Ascorbate Metabolic Engineering	37
1.9.	Thesis Aims and Objectives	39
Chapter 2.	Effect of <i>OsGGP</i> Constitutive Overexpression on Ascorbate Concentration, Salt Tolerance, and Iron Bioavailability in Rice	40
2.1.	Abstract	41
2.2.	Introduction	42
2.3.	Results	45
2.3.1	Constitutive overexpression of the <i>OsGGP</i> CDS in rice increases ascorbate concentrations in GBR and in the vegetative growth phase, but reduces levels of ascorbate at the reproductive growth phase	45
2.3.2	Constitutive overexpression of the <i>OsGGP</i> CDS in rice did not affect salt tolerance at the vegetative growth phase	48
2.3.3	Ascorbate concentrations were positively correlated with ferritin concentrations in Caco-2 cells exposed to <i>in vitro</i> digests of NS and 35S- <i>OsGGP</i> BR and GBR samples	50
2.4.	Discussion	53

2.5.	Materials and Methods	56
2.5.1.	Plant growth conditions	56
2.5.2.	Vector construction and rice transformation	56
2.5.3.	Genotyping	56
2.5.4.	Quantification of total ascorbate	57
2.5.5.	Quantitative reverse transcription-PCR analysis	57
2.5.6.	Assessment of salt tolerance with automated imaging	58
2.5.7.	Preparation of GBR	59
2.5.8.	Quantification of Fe	59
2.5.9.	Quantification of phytate	59
2.5.10.	Caco-2 cell culture assessment of Fe bioavailability	59
2.5.11.	Statistical analysis	60
2.6.	Supplementary information	61
Chapter 3.	Production and Characterization of Rice with CRISPR/Cas9-Induced Mutations in the <i>OsGGP</i> uORF	67
3.1.	Introduction	68
3.2.	Methods	71
3.2.1.	Plant growth conditions	71
3.2.2.	Vector construction	71
3.2.3.	Rice transformation	72
3.2.4.	Genotyping and sequencing	72
3.2.5.	Quantification of total ascorbate	72
3.2.6.	Preparation of germinated brown rice	74
3.2.7.	Quantitative reverse transcription-PCR analysis	74
3.2.8.	Assessment of salt tolerance and plant growth and development with automated imaging	75
3.2.9.	Elemental analysis of brown rice	76
3.2.10.	Analysis of potential off-target edits	76
3.2.11.	Statistical analysis	76
3.3.	Results	77
3.3.1.	Production of rice with CRISPR/Cas9-induced mutations in the <i>OsGGP</i> uORF	77

3.3.2.	Isolation of homozygous, transgene-free <i>uorf_{OsGGP}</i> mutants	77
3.3.3.	Ascorbate concentrations of the <i>uorf_{OsGGP}</i> mutants	80
3.3.4.	Transcript analysis of the <i>uorf_{OsGGP}</i> mutants	80
3.3.5.	Shoot growth measurements of the <i>uorf_{OsGGP}-1</i> mutant	85
3.3.6.	Agro-morphological performance of the <i>uorf_{OsGGP}-1</i> mutant	85
3.3.7.	Salt stress tolerance assessment of the <i>uorf_{OsGGP}-1</i> mutant	85
3.3.8.	Grain Fe concentrations of the <i>uorf_{OsGGP}-1</i> mutant	89
3.3.9.	Analysis of potential off-target mutations in the <i>uorf_{OsGGP}</i> mutants	89
3.4.	Discussion	95
3.4.1.	The CRISPR/Cas9 complex was active and generated new mutation as the T ₀ rice transformants developed	95
3.4.2.	CRISPR/Cas9-induced mutations increased foliar ascorbate concentrations without affecting transcript levels of the <i>OsGGP</i> gene, but did not affect ascorbate concentrations in brown rice or germinated brown rice	96
3.4.3.	Insights gained into the function of the <i>GGP</i> uORF peptide	96
3.4.4.	Increased ascorbate levels in the <i>uorf_{OsGGP}-1</i> mutant were associated with increased grain weight per plant	98
3.4.5.	Increased ascorbate levels in the <i>uorf_{OsGGP}-1</i> mutant did not affect salt tolerance	99
3.4.6.	Increased ascorbate levels in the <i>uorf_{OsGGP}-1</i> mutant did not affect grain Fe levels	99
3.4.7.	An off-target 23-bp deletion was detected in the <i>uorf_{OsGGP}-1</i> mutant	100
3.5.	Conclusion	102
Chapter 4. Genome-Wide Identification and Characterization of the <i>GDP-L-Galactose Phosphorylase</i> Gene Family in Bread Wheat		103
4.1.	Abstract	104
4.2.	Introduction	105
4.3.	Results	108
4.3.1.	Six <i>TaGGP</i> genes were identified on bread wheat chromosomal groups 4 and 5	108
4.3.2.	Each <i>TaGGP</i> gene contains a highly conserved uORF	108
4.3.3.	Graminaceous GGP proteins separate into two groups	111

4.3.4.	The <i>TaGGP1</i> and <i>TaGGP2</i> subchromosomal regions do not share synteny	111
4.3.5.	The two groups of <i>TaGGP</i> genes have distinct expression patterns in bread wheat tissues	115
4.3.6.	Transient transformation of <i>TaGGP</i> genes in <i>N. benthamiana</i> leaf increases ascorbate concentration	117
4.4.	Discussion	119
4.4.1.	The <i>TaGGP</i> genes are highly conserved, functional, and have an uORF amenable to genome editing	119
4.4.2.	A <i>Mutator</i> -like transposable element may be responsible for the duplication and specialized gene expression of <i>GGP2</i> in the <i>Brachypodium</i> /Triticeae lineage	120
4.4.3.	Applications towards plant breeding	123
4.5	Conclusions	124
4.6.	Methods	125
4.6.1.	Identification and naming of the <i>GGP</i> genes and uORFs	125
4.6.2.	Cloning and gene sequencing	125
4.6.3.	Sequence alignment and phylogenetic analysis of the <i>GGP</i> and uORF sequences	126
4.6.4.	Microsynteny analysis of the <i>TaGGP</i> , <i>BdGGP</i> , and <i>HvGGP</i> genes	126
4.6.5.	Quantitative reverse transcription-PCR analysis of the <i>TaGGP</i> genes	126
4.6.6	Analysis of <i>TaGGP</i> promoter regions	127
4.6.7.	Transient transformation of <i>N. benthamiana</i> leaves	127
4.6.8.	Ascorbate measurement	128
5.6.9.	Statistical analysis	129
4.7.	Supplementary information	130
Chapter 5.	General Discussion	147
Bibliography		155
Appendices		186

List of Figures

Chapter 1.

Figure 1.1.	The ascorbate-glutathione cycle in plants.	5
Figure 1.2.	Proposed ascorbate biosynthetic pathways and recycling in higher plants.	12

Chapter 2.

Figure 2.1.	Production and characterization of independent rice transformation events constitutively overexpressing the <i>OsGGP</i> CDS.	46
Figure 2.2.	Transcript analysis of NS and 35S- <i>OsGGP</i> plants at the reproductive growth phase	47
Figure 2.3.	Salt tolerance assessment of NS and 35S- <i>OsGGP</i> plants during the vegetative growth phase.	49
Figure 2.4.	Fe bioavailability assessment of NS and 35S- <i>OsGGP</i> BR and GBR.	51
Figure S2.1.	Foliar ascorbate concentrations of T ₃ and T ₄ NS and 35S- <i>OsGGP</i> plants at the reproductive growth phase.	63
Figure S2.2.	Foliar ascorbate concentrations in T ₁ hemizygous and homozygous NS and 35S- <i>OsGGP</i> plants at the reproductive growth phase.	64
Figure S2.3.	The sPSA and sPSA AGR of control and salt-stressed NS and 35S- <i>OsGGP</i> plants during the vegetative growth phase.	65
Figure S2.4.	Transcripts analysis of the endogenous <i>OsGGP</i> gene and 35S- <i>OsGGP</i> transgene in NS-1 and 35S- <i>OsGGP</i> -1 BR and GBR.	66

Chapter 3.

Figure 3.1.	CRISPR/Cas9-targeted mutagenesis of the <i>OsGGP</i> uORF.	78
Figure 3.2.	CRISPR/Cas9-induced mutation types and frequencies in the 31 plants with mutations detected in the <i>OsGGP</i> uORF transformed with the pUbi-Cas9-uORF T-DNA.	79
Figure 3.3.	Nucleotide sequence alignment of the parental T ₀ CRISPR/Cas9-induced <i>uorf_{OsGGP}</i> mutants used in this chapter.	81
Figure 3.4.	Nucleotide and peptide sequence alignment of the CRISPR/Cas9-induced homozygous, transgene-free <i>uorf_{OsGGP}</i> mutants used in this chapter.	82
Figure 3.5.	Ascorbate concentrations of WT and <i>uorf_{OsGGP}</i> plants.	83

Figure 3.6.	Transcript analysis of WT and <i>uorf_{OsGGP}</i> plants.	84
Figure 3.7.	Shoot growth measurements of WT and <i>uorf_{OsGGP}-1</i> plants using automated imaging over the course of a full lifecycle.	86
Figure 3.8.	Salt stress tolerance assessment of WT and <i>uorf_{OsGGP}-1</i> plants.	90
Figure 3.9.	The sPSA and sPSA AGR of control and salt-stressed WT and <i>uorf_{OsGGP}-1</i> plants.	91
 Chapter 4.		
Figure 4.1.	Gene structure of the <i>TaGGP</i> genes from bread wheat cv. Chinese Spring.	109
Figure 4.2.	Amino acid sequence alignment of GGP proteins from a range of graminaceous species.	110
Figure 4.3.	Amino acid sequence alignment of <i>GGP</i> uORFs from a range of graminaceous and non-graminaceous species.	112
Figure 4.4.	An unrooted phylogenetic tree of GGP proteins from a range of graminaceous and non-graminaceous species.	113
Figure 4.5.	Microsynteny analysis of the <i>TaGGP</i> genes from bread wheat cv. Chinese Spring.	114
Figure 4.6.	Quantitative reverse transcription-PCR analysis of the <i>TaGGP</i> genes from bread wheat cv. Chinese Spring.	116
Figure 4.7.	Transient transformation of the <i>TaGGP</i> coding sequences in <i>N. benthamiana</i> .	118
Figure S4.1.	Gene structure of the <i>AetGGP</i> , <i>BdGGP</i> , <i>HvGGP</i> , <i>OsGGP</i> , <i>SbGGP</i> , and <i>ZmGGP</i> genes.	130
Figure S4.2.	An unrooted phylogenetic tree of <i>GGP</i> coding sequences from a range of graminaceous species.	131
Figure S4.3.	An unrooted phylogenetic tree of <i>GGP</i> uORF peptides from a range of graminaceous and non-graminaceous species.	132
Figure S4.4.	Annotation of anther/pollen <i>cis</i> -acting elements within the 1-kb promoter region of the <i>TaGGP</i> genes.	133
Figure S4.5.	<i>GGP</i> gene expression data extracted from http://bar.utoronto.ca/ for wheat and barley and https://www.ebi.ac.uk for <i>Brachypodium</i> .	134
Figure S4.6.	<i>GGP</i> gene expression data extracted from https://www.ebi.ac.uk for rice and sorghum and http://bar.utoronto.ca/ for maize.	135
Figure S4.7.	Nucleotide sequence alignment of the 1-kb promoter regions of wheat, <i>Brachypodium</i> , and barley <i>GGP</i> genes.	136

List of Tables

Chapter 1.

Table 1.1.	Ascorbate and production values of major global crops.	10
Table 1.2.	Overview of metabolic engineering strategies utilizing ascorbate biosynthetic genes to increase ascorbate concentrations in model and crop species.	13
Table 1.3.	Overview of metabolic engineering strategies utilizing ascorbate recycling genes to increase ascorbate concentrations in model and crop species.	25
Table 1.4.	Overview of metabolic engineering strategies utilizing ascorbate regulatory genes to increase ascorbate concentrations in model and crop species.	29

Chapter 2.

Table 2.1.	Correlation analysis of ascorbate, phytate, and Fe concentrations in T ₃ homozygous NS-1 and 35S- <i>OsGGP</i> -1 BR and GBR with ferritin concentrations in Caco-2 cells exposed to <i>in vitro</i> digests of NS and 35S- <i>OsGGP</i> BR and GBR sample	52
Table S2.1.	Ascorbate concentrations of WT and T ₁ 35S- <i>OsGGP</i> GBR and number of T ₁ 35S- <i>OsGGP</i> seedlings with T-DNA present or absent.	61
Table S2.2.	List of PCR and qRT-PCR primers used in this study. The table provides primer name, forward and reverse primer sequences, PCR product length, and the application of the primers.	62

Chapter 3.

Table 3.1.	List of PCR and qRT-PCR primers used in this chapter.	73
Table 3.2.	Agro-morphological performance of T ₂ homozygous, transgene-free WT and <i>uorf_{OsGGP}</i> -1 plants.	87
Table 3.3.	Quality parameters of T ₃ homozygous, transgene-free WT and <i>uorf_{OsGGP}</i> -1 grain.	88
Table 3.4.	Elemental concentrations of T ₃ homozygous, transgene-free WT and <i>uorf_{OsGGP}</i> -1 brown rice.	92
Table 3.5.	Potential off-target sites (mismatches $\leq$ 2-bp) in genic regions of the pUbi-Cas9-uORF sgRNA.	93
Table 3.6.	Summary of off-target mutations detected in three representative T ₂ homozygous, transgene-free WT, <i>uorf_{OsGGP}</i> -1, <i>uorf_{OsGGP}</i> -2, and <i>uorf_{OsGGP}</i> -3 plants with Sanger sequencing.	94

Chapter 4.

Table S4.1.	Ten high confidence genes upstream and downstream of the six <i>TaGGP</i> genes.	137
Table S4.2.	Ten high confidence genes upstream and downstream of the two <i>BdGGP</i> genes.	141
Table S4.3.	Ten high confidence genes upstream and downstream of the two <i>HvGGP</i> genes.	143
Table S4.4.	The <i>GGP</i> genes utilized in the manuscript for sequence alignments and phylogenetic analyses.	145
Table S4.5.	Primers used for the quantitative reverse transcription-PCR analysis of the six <i>TaGGP</i> and housekeeping genes.	146

Chapter 5.

Table 5.1.	Ascorbate and Fe concentrations of major global commodity crops including the molar ratio of Fe to ascorbate and the fold-increase in ascorbate required to achieve molar ratios of 1:5, 1:10, and 1:20 of Fe to ascorbate.	151
------------	---	-----

List of Abbreviations

35S-<i>OsGGP</i>	35S promoter-driven expression of the rice <i>GDP-L-galactose phosphorylase</i> coding sequence
5' UTR	5' untranslated region
AGR	absolute growth rate
AMR	ascorbic acid mannose pathway regulator
APX	ascorbate peroxidase
Asc	ascorbate
bHLH	basic helix-loop-helix
BR	brown rice
CAT	catalase
CBF	cystathionine- β -synthase
CDS	coding sequence
CML	calmodulin-like
CSN	COP9 signalosome
DAP	days after planting
DDB	UV-DAMAGED DNA-BINDING PROTEIN
D-GalUR	D-galacturonate reductase
DHAR	dehydroascorbate reductase
Dof	DNA-binding with one finger
ERF	ethylene response factor
EST	expressed sequence tag
Fe	iron
Fe²⁺	ferrous iron
Fe³⁺	ferric iron
GBR	germinated brown rice
GGP	GDP-L-galactose phosphorylase
GlcAK	glucuronokinase
GME	GDP-D-mannose-3',5'-epimerase
GMP	GDP-D-mannose pyrophosphorylase
GPP	L-galactose-1-phosphate phosphatase
GR	glutathione reductase

HIT	histidine triad
HP	High-Pigment
hptII	hygromycin phosphotransferase II
HSP	heat shock protein
hSpCas9	human codon optimised <i>Streptococcus pyogenes</i> Cas9
HZ	HD-Zip I family transcription factor
ICP-OES	inductively coupled plasma optical emission spectrometry
IWGSC	International Wheat Genome Sequencing Consortium
KJC	KONJAC
L-GalDH	L-galactose dehydrogenase
L-GalLDH	L-galactono-1,4-lactone dehydrogenase
L-GulLO	L-gulono-1,4-lactone oxidase
MDAR	monodehydroascorbate reductase
MEM	minimum essential media
MIOX	myo-inositol oxygenase
mORF	major open reading frame
MULE	<i>mutator</i> -like transposable element
MYB	myeloblastosis transcription factor
NL	NBS-LRR
NLS	nuclear localization signal
NS	null segregant
nsLTP	non-specific lipid transfer protein
OT	off-target
PAP	purple acid phosphatase
PMI	phosphomannose isomerase
PMM	phosphomannose mutase
PSA	projected shoot area
qRT-PCR	quantitative reverse transcription-PCR
RDA	recommended daily allowance
RGR	relative growth rate
ROS	reactive oxygen species

SOD	superoxide dismutase
SOS	salt overly sensitive
sgRNA	single guide RNA
sPSA	smoothed projected shoot area
SR	subchromosomal regions
TIR	terminal inverted repeat
TSS	transcription start site
Ubi	ubiquitin
uORF	upstream open reading frame
<i>uorfosGGP</i>	rice <i>GDP-L-galactose phosphorylase</i> upstream open reading frame mutant
VTC	VITAMIN C
WT	wild-type
ZF	Zinc-finger

Chapter 1.

Metabolic Engineering Strategies to Increase Ascorbate Concentrations in Plants

1.1. Introduction

Providing adequate, safe, and nutritious food for the growing human population under changing climatic conditions represents one of the greatest challenges of the 21st century. Adverse environmental conditions such as extreme temperatures, drought, and soil salinity are the primary cause of crop losses worldwide, and are predicted to be exacerbated by climate change [1]. With the human population expected to reach 9.7 billion by the year 2050, the need for more abiotic stress tolerant crops is critical for ensuring global food security [2]. Essential vitamin and mineral deficiencies—a form of undernutrition often referred to as hidden hunger—are a major threat to the health and development of populations worldwide, affecting more than 2 billion people [3]. Hidden hunger predominantly affects women and children and is often acute and long-lasting—leading to impaired mental and physical development, poor health, reduced productivity, and death [3]. In addition to negatively impacting human health, hidden hunger restricts the socioeconomic growth of low- and middle-income countries by maintaining a cycle of malnutrition, poverty, and reduced economic productivity [3]. Improving the nutritional content of crops through agronomic practices, plant breeding, or biotechnology—a process known as biofortification—offers a low-cost and sustainable strategy to combat hidden hunger [4].

L-ascorbic acid (also known as vitamin C, and hereafter referred to as ascorbate) is an important multifunctional molecule for both plants and mammals [5]. Ascorbate is a reducing agent capable of donating electrons and primarily functions as a cellular antioxidant and enzymatic co-factor [5]. In plants, ascorbate is the most abundant water-soluble antioxidant and is well-known for its important role in photosynthetic functions and stress tolerance, but also plays a key role in regulating plant growth and developmental processes, such as the transport of Fe into developing seeds [6, 7]. In humans, ascorbate is an essential micronutrient well-known for its role in collagen formation and the prevention of the disease scurvy [8], but also plays other key roles in human physiology, such as promoting the dietary uptake of iron (Fe) in the gut [9]. Increasing ascorbate concentrations in crops has the potential to improve both the abiotic stress tolerance of crops and human health.

Four ascorbate biosynthetic pathways have been proposed in plants: the L-galactose, L-gulose, *myo*-inositol, and D-galacturonate pathways [10, 11]. The L-galactose pathway, however, is the only pathway to have all the enzymatic steps characterized. Substantial genetic evidence

now also supports the L-galactose pathway as the dominant pathway towards ascorbate biosynthesis in plants [12-16]. The *GDP-L-galactose phosphorylase* (*GGP* or *vtc2/5*) gene encodes the rate-limiting and first committed step toward ascorbate biosynthesis of the L-galactose pathway and is regulated at both the transcriptional and translational level [17]. Increasing the production of *GGP* has emerged as the most promising strategy to increase ascorbate concentrations in crops [17, 18].

In this chapter the role of ascorbate in plants and humans is first discussed, followed by an overview of ascorbate concentrations in major global crops. A review of ascorbate biosynthesis, recycling, and regulation in plants is then presented, including a summary of all metabolic engineering strategies utilized to date to increase ascorbate concentrations in model and crop species. Finally, the opportunity to metabolically engineer increased ascorbate concentrations in rice (*Oryza sativa* L.) and wheat (*Triticum aestivum* L.), two of the world's most important staple crops, is highlighted, and the aims and objectives of this thesis outlined.

1.2. Role of Ascorbate in Plants

Ascorbate is the most abundant water-soluble antioxidant in plant cells found in high concentrations in the cytosol, peroxisomes, nuclei, mitochondria, and chloroplasts [19]. Ascorbate is particularly well-known for its role in photosynthetic functions and stress tolerance in plants. This is largely due to the ability of ascorbate to counteract oxidative stress produced by normal or stressed cellular metabolism, either directly as a reactive oxygen species (ROS) scavenger, or through the ascorbate-glutathione cycle—a major antioxidant system of plants cells (Figure 1.1) [6, 20, 21]. In this cycle, H_2O_2 is detoxified to H_2O by ascorbate peroxidase (APX) using ascorbate as an electron donor. The oxidized ascorbate can then be regenerated by the ascorbate recycling enzymes monodehydroascorbate reductase (MDAR) or dehydroascorbate reductase (DHAR) using NADPH and glutathione, respectively, as electron donors (see Section 1.6) [21, 22].

Photosynthesis is a well-known source of ROS in plant cells [23]. Consequently, photoprotective mechanisms—which ascorbate plays key roles in—are critical in preventing photoinhibition and oxidative stress. For example, ascorbate serves as a co-factor for the enzyme violaxanthin de-epoxidase which catalyzes the de-epoxidation of violaxanthin and antheraxanthin to zeaxanthin—which plays a direct photoprotective role in dissipating excess energy [24, 25]. Ascorbate also plays a role in the photoprotective water-water cycle where it serves as an electron donor for APX in the reduction of H_2O_2 produced by photosynthesis to H_2O [26]. Further, ascorbate can serve as an alternative electron donor to photosystem II when the oxygen-evolving complex of photosystem II is inactivated in order to sustain electron transport activity [27, 28].

A common consequence of many abiotic stresses is the overproduction of ROS [22]. For example, stomatal closure during drought and salt stress limits carbon dioxide uptake which has both direct and indirect effects on ROS production. For instance, reduced NADP^+ regeneration through the Calvin cycle or enhanced photorespiration results in a higher leakage of photosynthetic or respiratory electrons, that in turn stimulate ROS production [29, 30]. Similarly, Na^+ and Cl^- toxicity during salt stress can disrupt the photosynthetic electron transport chain further promoting electron leakage and ROS accumulation [31]. Likewise, temperature stresses are capable of uncoupling temperature-sensitive pathways such as photosynthesis and respiration, consequently increasing ROS formation [32]. Genetic approaches elevating ascorbate concentrations in model

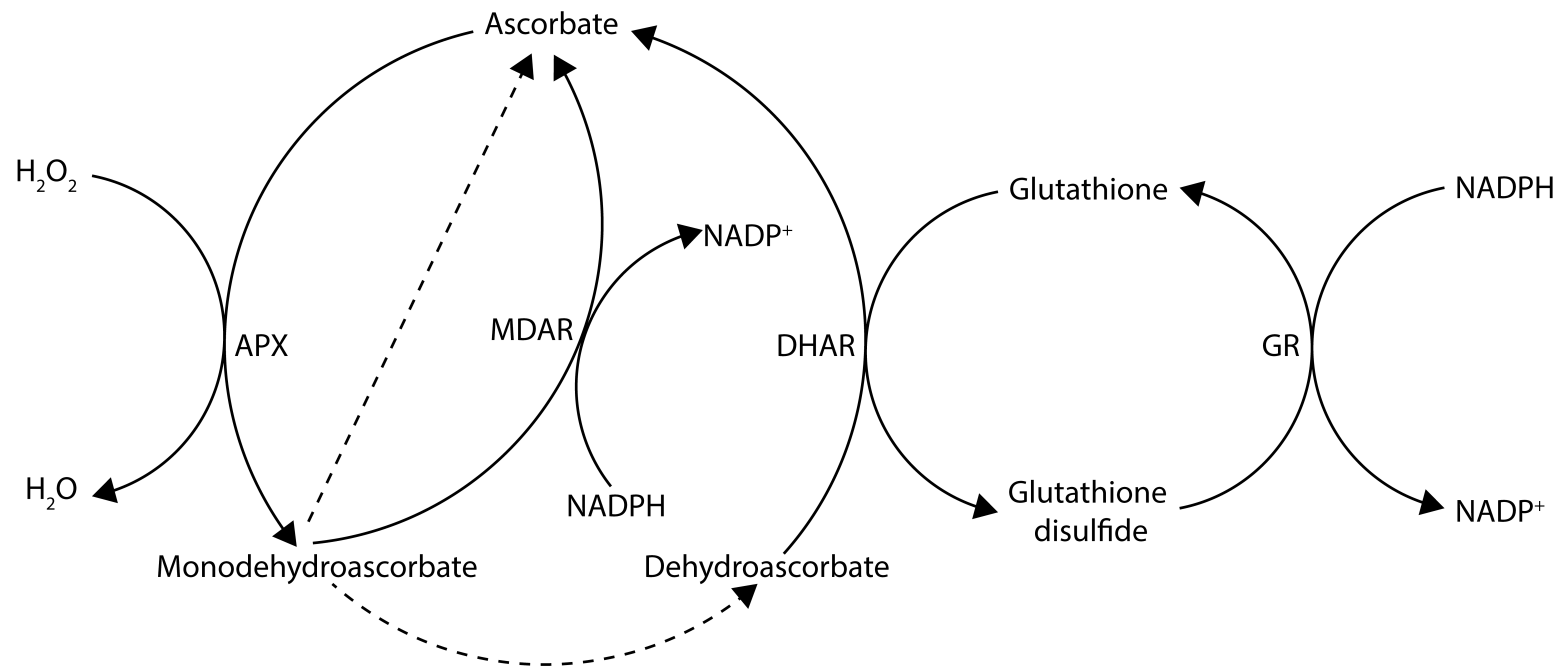


Figure 1.1. The ascorbate-glutathione cycle—a major antioxidant system of plants cells. In this cycle electrons flow from NADPH to H_2O_2 . The genes encoding the enzymatic steps are shown in italics. Dashed arrows indicate non-enzymatic disproportionation. *APX*, ascorbate peroxidase; *MDAR*, monodehydroascorbate reductase; *DHAR*, dehydroascorbate reductase; *GR*, glutathione reductase.

and crop species have demonstrated that even minor increases in ascorbate can improve plant performance under a broad range of abiotic stresses by mitigating the harmful effects of the overproduction of ROS (see Sections 1.5, 1.6, and 1.7).

Numerous other roles for ascorbate in plants include the transport of Fe into developing seeds [7], biosynthesis of ethylene and organic acids [33, 34], regenerating oxidized α -tocopherol [35], and influencing cell division [36], cell expansion [37], flowering time, and senescence [38].

1.3. Role of Ascorbate in Humans

In contrast to most other mammals, humans are not able to synthesize ascorbate due to loss-of-function mutations in the *L-gulonono-1,4-lactone oxidase* (*L-GulLO*, also known as *GULO* or *GLOase*) gene, which encodes the last enzymatic step of ascorbate biosynthesis in vertebrates [39]. These loss-of-function mutations in *L-GulLO* likely occurred in the ancestor of humans and other anthropoid primates when the ancestral diet was rich in ascorbate, and therefore would not cause a selective disadvantage [39]. This hypothesis is consistent with the diets of modern wild anthropoid primates that consume far more ascorbate than the recommended daily allowance (RDA) for adults in the United States of about 1 mg/kg/day [39]. Loss-of-function mutations in *L-GulLO* have also independently occurred in guinea pigs and fruit bats, both of which also have diets rich in ascorbate [39]. Humans must therefore obtain ascorbate through diet, principally through fresh fruits and vegetables which are rich sources of ascorbate (see Section 1.4).

Ascorbate is the most abundant water-soluble antioxidant in the plasma and tissue of humans, and plays a critical role in human physiology as a general antioxidant and as an enzymatic co-factor [40, 41]. Ascorbate is best known for its role in collagen formation and the prevention of the disease scurvy [8]; the name ascorbate is derived from ‘antiscorbutic factor’. Ascorbate is a co-factor for the dioxygenase enzymes involved in the hydroxylation of prolyl and lysyl residues of procollagen, and is responsible for maintaining the Fe in the active site of these enzymes in the reduced ferrous state (Fe^{2+}) [40]. Ascorbate has also recently been identified as a co-factor for dioxygenase enzymes involved in epigenetic programming, such as DNA demethylation and histone demethylation [42]. Another important role for ascorbate in human physiology includes the dietary uptake of Fe [9]. Ascorbate is a potent enhancer of Fe absorption and is able to overcome strong inhibitors of Fe absorption in the human digestive system, such as phytates and polyphenols [43, 44]. Ascorbate is proposed to enhance Fe absorption by reducing ferric Fe (Fe^{3+}) to bioavailable Fe^{2+} , and then forming a soluble Fe^{2+} chelate that is well absorbed in the small intestines [40, 45]. In addition to promoting Fe bioavailability, ascorbate also regulates mammalian cellular Fe uptake and metabolism [46]. For example, ascorbate stimulates the uptake of Fe-citrate complexes, modulates the transferrin Fe uptake pathway, stimulates ferritin synthesis, inhibits ferritin degradation, and decreases cellular Fe efflux [46, 47]. Beyond its critical role in human physiology, ascorbate has a wide range of positive health effects, for example by reducing

infectious diseases, arterial plaque formation, blood pressure, the risk of stroke, diabetic complications, and the formation of cancer [40].

Despite severe ascorbate deficiency, or scurvy, being regarded as a disease of historical significance, suboptimal intake of ascorbate and ascorbate deficiencies are still present in both developing and developed countries [48-50]. The re-emergence of scurvy has also been recently documented in diabetic patients in Australia due to poor dietary habits [51]. Given the importance of ascorbate in human health and physiology, suboptimal intake of ascorbate could be a significant factor in reducing healthy life years.

1.4. Ascorbate Concentrations in Crops

Ascorbate concentrations vary widely between plant species, plant cultivars, and plant tissues within a species [17]. Fresh fruits and vegetables are the richest sources of ascorbate, with some fruits such as the acerola (*Malpighia emarginata*) cherry accumulating ascorbate concentrations as high as 1677.6 mg/100 g [52]. Other fruits known for their high ascorbate concentrations include guava (*Psidium guajava* L.) (228.3 mg/100 g), capsicum (*Capsicum annuum* L.) (183.5 mg/100 g), and kiwifruit (*Actinidia chinensis*) (161.3 mg/100 g) [52]. The edible portions of many major global crops, however, have relatively low ascorbate concentrations (Table 1.1). Since ascorbate is also susceptible to oxidation, the ascorbate concentrations of these foods are further reduced by storage, processing, and cooking [53]. Achieving the RDA for adults in the United States of about 1 mg/kg/day ascorbate would be extremely challenging with diets dependent on many of these major global crops. Notably, cereal grains have negligible ascorbate levels due to the decline of ascorbate during grain maturation [54, 55], with the exception of the sweet corn variety of maize (*Zea mays* L.) which is consumed when the kernels are immature. Cornmeal, a common staple food made from ground dried maize, however, also contains negligible ascorbate levels [52].

Increasing ascorbate concentrations in these major global crops through plant breeding or biotechnology could represent a low-cost and sustainable strategy to improve the global dietary intake of ascorbate. Further, increasing ascorbate concentrations in these major global crops may help combat global Fe deficiency, the most common and widespread nutritional disorder in the world, affecting approximately two billion people worldwide [56]. Since most diets contain Fe in at least the non-heme form, it has been argued that the problem of Fe deficiency anemia could be viewed as one of inadequate Fe bioavailability. Given that ascorbate is a powerful enhancer of Fe absorption in humans, the problem of Fe deficiency anemia may be better viewed as one of inadequate ascorbate intake [40].

Table 1.1. Ascorbate^a and production^b values of major global crops^c.

Crop	Ascorbate (mg/100 g FW)	Production (tonnes)
Wheat	0.0	771,718,579
Rice	0.0	769,657,791
Barley	0.0	147,404,262
Sorghum	0.0	57,601,588
Eggplants	2.2	52,309,119
Sweet potatoes	2.4	112,835,316
Coconuts	3.3	60,773,435
Grapes	4.0	74,276,583
Apples	4.6	83,139,326
Soybeans	6.0	352,643,548
Maize	6.8 ^d	1,134,746,667
Onions	7.4	97,862,928
Watermelons	8.1	118,413,465
Bananas	8.7	113,918,763
Tomatoes	14.0	182,301,395
Yams	17.1	73,018,874
Potatoes	19.7	388,190,674
Cassava	20.6	291,992,646
Oranges	53.2	73,313,089

^aAccording to the USDA Food Composition Database [52]. ^bAccording to FAOSTAT (>50,000,000 tonnes) [57]. ^cExcluding 'sugar cane', 'oil palm fruit', 'vegetable, fresh', 'cucumbers and gherkins', 'rapeseed', 'seed cotton', 'cabbages and other brassicas', and 'mangoes, mangosteens, guavas'.

^dMaize ascorbate value is based on the sweet corn variety.

1.5. Biosynthesis of Ascorbate in Plants

Four pathways toward ascorbate biosynthesis have been proposed in plants: the L-galactose, L-gulose, *myo*-inositol, and D-galacturonate pathways [10, 11] (Figure 1.2). All four pathways share an aldono-1,4-lactone as the direct precursor to ascorbate (L-galactono-1,4-lactone for the L-galactose and D-galacturonate pathways and L-gulono-1,4-lactone for the L-gulose and *myo*-inositol pathways). While there is strong evidence for the L-galactose pathway as the dominant ascorbate biosynthetic pathway in plants, the contribution of the L-gulose, *myo*-inositol, and D-galacturonate pathways toward ascorbate biosynthesis in plants is not as well documented, and often controversial. An overview of metabolic engineering strategies utilizing ascorbate biosynthetic genes to increase ascorbate concentrations in model and crop species, including the tissue examined and any stress tolerances observed, is summarized in Table 1.2.

1.5.1. The L-Galactose Pathway

The L-galactose pathway (also known as the D-mannose/L-galactose pathway or Smirnoff-Wheeler pathway) is the only ascorbate biosynthetic pathway in plants to have all its enzymatic steps characterized. The L-galactose pathway is responsible for converting D-fructose-6-P to ascorbate via eight enzymatic steps by phosphomannose isomerase (PMI), phosphomannose mutase (PMM), GDP-D-mannose pyrophosphorylase (GMP), GDP-D-mannose-3',5'-epimerase (GME), GDP-L-galactose phosphorylase (GGP), L-galactose-1-phosphate phosphatase (GPP), L-galactose dehydrogenase (L-GalDH), and L-galactono-1,4-lactone dehydrogenase (L-GalLDH) [10, 11] (Figure 1.2). All enzymatic steps of the L-galactose pathway take place in the cytosol except for the conversion of L-galactono-1,4-lactone to ascorbate by L-GalLDH, which occurs in mitochondria [58].

1.5.1.1. Phosphomannose Isomerase (PMI)

The PMI enzyme is responsible for the first enzymatic step of the L-galactose pathway, catalyzing a reversible conversion of D-fructose-6-P to D-mannose-6-P (Figure 1.2). Knock-down of the *PMII* gene in *Arabidopsis thaliana* by RNAi reduced ascorbate to 47% of wild-type levels, supporting its role in L-galactose ascorbate biosynthesis [59]. To date, no studies have investigated the stable increased expression of the *PMI* gene in plants. However, transient expression of the

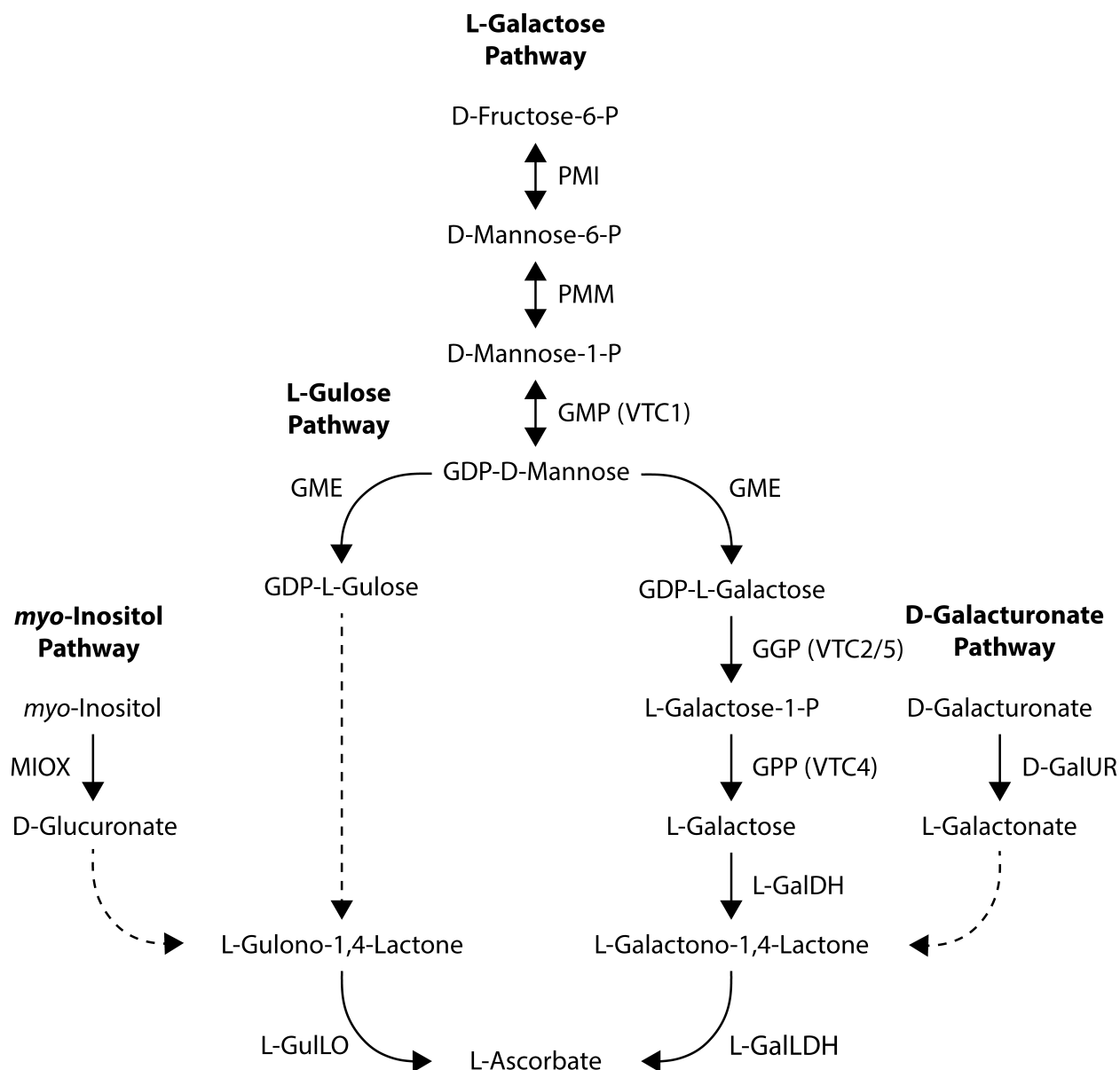


Figure 1.2. The four proposed ascorbate biosynthetic pathways in higher plants. The genes encoding the enzymatic steps are shown in *italics*. Dashed arrows indicate multiple and/or unknown biosynthetic steps in higher plants. *PMI*, *phosphomannose isomerase*; *PMM*, *phosphomannose mutase*; *GMP*, *GDP-D-mannose pyrophosphorylase*; *GME*, *GDP-D-mannose-3',5'-epimerase*; *GPP*, *GDP-L-galactose phosphorylase*; *GPP*, *L-galactose-1-P phosphatase*; *L-GalDH*, *L-galactose dehydrogenase*; *L-GalLDH*, *L-galactono-1,4-lactone dehydrogenase*; *L-GulLO*, *L-gulono-1,4-lactone oxidase*; *MIOX*, *myo-inositol oxygenase*; *D-GalUR*, *D-galacturonate reductase*.

Table 1.2. Overview of metabolic engineering strategies utilizing ascorbate biosynthetic genes to increase ascorbate concentrations in model and crop species.

Species transformed	Gene(s) transformed	Gene source(s)	Promoter	Max fold-change	Tissue examined	Stress tolerance	Reference
L-galactose pathway genes							
Tobacco	<i>PMM</i>	Acerola	<i>CaMV 35S</i>	~2.5	Leaves	-	[60]
<i>Arabidopsis</i>	<i>PMM-GFP</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.3	Leaves	Methyl viologen	[61]
<i>Arabidopsis</i>	<i>GMP</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.3	Leaves	-	[62]
<i>Arabidopsis</i>	<i>GMP</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~1.2 ^{ns}	Seedlings	-	[63]
Rice	<i>GMP</i>	<i>Arabidopsis</i>	<i>ZmUbi</i>	~1.5	Leaves	-	[64]
Tobacco	<i>GMP</i>	Peach	<i>CaMV 35S</i>	~1.3 ^{ns}	Leaves	-	[65]
Tobacco	<i>GMP</i>	Acerola	<i>MgGMP</i>	~2.5	Leaves	-	[66]
Tobacco	<i>GMP</i>	Acerola	<i>CaMV 35S</i>	~2.0	Leaves	-	[66]
Tomato	<i>GMP1</i>	Tomato	<i>CaMV 35S</i>	~1.4, ~1.2	Leaves, Red Fruit	-	[67]
Tobacco	<i>GMP + GME</i>	Peach	<i>CaMV 35S</i>	~1.4 ^{ns} , ~1.8 ^{ns} , ~1.0 ^{ns} , ~1.1 ^{ns}	Young Leaves, Old Leaves, Flower Buds, Immature Fruits	-	[65]
Tomato	<i>GMP1 + GME2</i>	Tomato	<i>CaMV 35S</i>	2.0, ~1.3	Leaves, Red Fruit	Methyl viologen	[67]
Tomato	<i>GMP1 + GME2 + GGP1 + GPP1</i>	Tomato	<i>CaMV 35S</i>	2.0, ~1.3	Leaves, Red Fruit	Methyl viologen	[67]
<i>Arabidopsis</i>	<i>GME</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.4	Leaves	-	[62]
<i>Arabidopsis</i>	<i>GME</i>	Alfalfa	<i>CaMV 35S</i>	1.8	Leaves	Low pH, drought, salt	[68]
<i>Arabidopsis</i>	<i>GME</i>	Rose	<i>CaMV 35S</i>	1.9	Leaves	-	[69]
Rice	<i>GME</i>	<i>Arabidopsis</i>	<i>ZmUbi</i>	~1.9	Leaves	Salt	[64]
Tobacco	<i>GME</i>	Peach	<i>CaMV 35S</i>	~1.3 ^{ns}	Leaves	-	[65]
Tomato	<i>GME1</i>	Tomato	<i>CaMV 35S</i>	1.4, 1.6	Leaves, Fruits	Methyl viologen, cold, salt	[70]
Tomato	<i>GME2</i>	Tomato	<i>CaMV 35S</i>	~1.8, ~1.2	Leaves, Red Fruit	-	[67]
Tomato	<i>GME2</i>	Tomato	<i>CaMV 35S</i>	1.3, 1.2	Leaves, Fruits	Methyl viologen, cold, salt	[70]
<i>Arabidopsis</i>	<i>GGP</i>	Kiwifruit	<i>CaMV 35S</i>	4.1	Leaves	-	[71]
Rice	<i>GGP</i>	Kiwifruit	<i>OsLP2</i>	2.5	Leaves	Salt, ozone	[72]

Table 1.2. (Continued).

Species transformed	Gene(s) transformed	Gene source(s)	Promoter	Max fold-change	Tissue examined	Stress tolerance	Reference
Strawberry	<i>GGP</i>	Kiwifruit	<i>CaMV 35S</i>	1.8, 2.1	Leaves, Fruit	-	[73]
Tomato	<i>GGP</i>	Kiwifruit	<i>CaMV 35S</i>	2.0, 6.2	Leaves, Fruit	-	[73]
<i>Arabidopsis</i>	<i>GGP1</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	2.9	Leaves	-	[62]
Potato	<i>GGP1</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.7 ^{ns}	Tubers	-	[73]
Potato	<i>GGP1</i>	Potato	<i>CaMV 35S</i>	1.8 ^{ns}	Tubers	-	[73]
Potato	<i>GGP1</i>	Potato	<i>StPAT</i>	3.0	Tubers	-	[73]
Potato	<i>GGP1</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.5 ^{ns}	Tubers	-	[73]
Rice	<i>GGP1</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	2.6	Leaves	Salt	[64]
Tobacco	<i>GGP1</i>	Tomato	<i>CaMV 35S</i>	1.4	Leaves	Cold	[74]
Tomato	<i>GGP1</i>	Tomato	<i>CaMV 35S</i>	2.0, ~1.1 ^{ns}	Leaves, Red Fruit	-	[67]
<i>Arabidopsis</i>	<i>GGP1</i> + <i>L-GalLDH</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	3.6	Leaves	-	[62]
<i>Arabidopsis</i>	<i>GGP1</i> + <i>GPP</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	4.1	Leaves	-	[62]
Tomato	<i>GGP1</i> + <i>GPP1</i>	Tomato	<i>CaMV 35S</i>	~1.8, ~1.0 ^{ns}	Leaves, Red Fruit	Methyl viologen	[67]
Potato	<i>GGP2</i>	Potato	<i>CaMV 35S</i>	2.4	Tubers	-	[73]
Potato	<i>GGP2</i>	Potato	<i>StPAT</i>	3.1	Tubers	-	[73]
<i>Arabidopsis</i>	<i>GPP</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.5	Leaves	-	[62]
Rice	<i>GPP</i>	<i>Arabidopsis</i>	<i>ZmUbi</i>	~1.4	Leaves	-	[64]
Tomato	<i>GPP1</i>	Tomato	<i>CaMV 35S</i>	~1.7, ~1.1 ^{ns}	Leaves, Red Fruit	-	[67]
<i>Arabidopsis</i>	<i>L-GalDH</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.2	Leaves	-	[62]
Rice	<i>L-GalDH</i>	<i>Arabidopsis</i>	<i>ZmUbi</i>	~1.7	Leaves	-	[64]
Tobacco	<i>L-GalDH</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~1.1 ^{ns}	Leaves	-	[75]
<i>Arabidopsis</i>	<i>L-GalLDH</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.8	Leaves	-	[62]
Lettuce	<i>L-GalLDH</i>	Lettuce	<i>PspetE</i>	1.3	Leaves	-	[76]
Lettuce	<i>L-GalLDH</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	3.2	Leaves	-	[77]
Lily	<i>L-GalLDH</i>	Apple	<i>CaMV 35S</i>	7.0	Leaves	-	[78]
Rice	<i>L-GalLDH</i>	Rapeseed	<i>ZmUbi</i>	~1.5	Leaves	-	[64]

Table 1.2. (Continued).

Species transformed	Gene(s) transformed	Gene source(s)	Promoter	Max fold-change	Tissue examined	Stress tolerance	Reference
Rice	<i>L-GalLDH</i>	Rice	<i>ZmUbi</i>	1.3	Leaves	-	[79]
Tobacco	<i>L-GalLDH</i>	Rose	<i>CaMV 35S</i>	2.1	Leaves	Salt, methyl viologen	[80]
Tobacco	<i>L-GalLDH</i>	Sweet potato	<i>CaMV 35S</i>	1.0 ^{ns}	Leaves	-	[81]
L-gulose and myo-inositol pathway genes							
<i>Arabidopsis</i>	<i>L-GulLO3</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~0.8 ^{ns}	Leaves	-	[82]
<i>Arabidopsis</i>	<i>L-GulLO5</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~1.1 ^{ns}	Leaves	-	[82]
Rice	<i>MIOX</i>	Rice	<i>CaMV 35S</i>	1.0 ^{ns}	Leaves	Drought	[83]
Tomato	<i>MIOX2</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	0.5, 1.4, 1.3	Leaves, Green Fruit, Red Fruit	-	[84]
<i>Arabidopsis</i>	<i>MIOX4</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.6	Leaves	Salt, cold, heat, pyrene	[85]
<i>Arabidopsis</i>	<i>MIOX4</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	3.0	Leaves	-	[86]
<i>Arabidopsis</i>	<i>MIOX4</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.7	Leaves	Heat	[27]
<i>Arabidopsis</i>	<i>MIOX4</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.0 ^{ns}	Leaves	-	[87]
D-galacturonate pathway genes							
<i>Arabidopsis</i>	<i>D-GalUR</i>	Strawberry	<i>CaMV 35S</i>	3.0	Leaves	-	[88]
Potato	<i>D-GalUR</i>	Strawberry	<i>CaMV 35S</i>	2.0	Tubers	Methyl viologen, salt, drought	[89]
Potato	<i>D-GalUR</i>	Strawberry	<i>CaMV 35S</i>	~2.1	Tubers	Methyl viologen, salt, zinc chloride	[90]
Potato	<i>D-GalUR</i>	Strawberry	<i>CaMV 35S</i>	~1.7	Tubers	Salt	[91]
Tomato	<i>D-GalUR</i>	Strawberry	<i>CaMV 35S</i>	2.5	Fruit	Methyl viologen, salt, drought	[92]
Tomato	<i>D-GalUR</i>	Strawberry	<i>CaMV 35S</i>	1.8, 2.0	Leaves, Red Fruit	Methyl viologen, salt, cold	[93]
Tomato	<i>D-GalUR</i>	Strawberry	<i>CaMV 35S</i>	1.3, ~1.0 ^{ns} , 1.4	Leaves, Green Fruit, Light Red Fruit	-	[94]
Tomato	<i>D-GalUR</i>	Strawberry	<i>SIPG</i>	~1.0 ^{ns} , ~1.0 ^{ns} , 1.4	Leaves, Green Fruit, Light Red Fruit	-	[94]

Table 1.2. (Continued).

Species transformed	Gene(s) transformed	Gene source(s)	Promoter	Max fold-change	Tissue examined	Stress tolerance	Reference
Animal and yeast genes							
Tobacco	<i>ALO</i>	Yeast	<i>CaMV 35S</i>	~2.5	Leaves	Cold	[95]
Tobacco	<i>ALO</i>	Yeast	<i>CaMV 35S</i>	~1.5, ~1.5, ~2.0	Young Leaves, Mature Leaves, Old Leaves	Methyl viologen, high light, Al toxicity	[96]
Tomato	<i>ALO</i>	Yeast	<i>CaMV 35S</i>	1.5, 1.3, 1.1 ^{ns}	Leaves, Green Fruit, Red Fruit	-	[84]
Stylo	<i>ALO + NCED</i>	Yeast/stylo	<i>CaMV 35S</i>	3.4	Leaves	Drought, cold	[95]
Tobacco	<i>ALO + NCED</i>	Yeast/stylo	<i>CaMV 35S</i>	~2.5	Leaves	Drought, cold	[95]
Tomato	<i>GMP</i>	Yeast	<i>CaMV 35S</i>	1.7, 1.5, 1.4	Leaves, Green Fruit, Red Fruit	-	[84]
<i>Arabidopsis</i>	<i>L-GulLO</i>	Rat	<i>CaMV 35S</i>	1.8	Leaves	Salt, cold, heat, pyrene	[85]
<i>Arabidopsis</i>	<i>L-GulLO</i>	Rat	<i>CaMV 35S</i>	2.0	Leaves	-	[97]
Lettuce	<i>L-GulLO</i>	Rat	<i>CaMV 35S</i>	7.0	Leaves	-	[98]
Potato	<i>L-GulLO</i>	Rat	<i>CaMV 35S</i>	2.4	Tubers	Methyl viologen, salt, drought	[99]
Tobacco	<i>L-GulLO</i>	Rat	<i>CaMV 35S</i>	7.0	Leaves	-	[98]
Tomato	<i>L-GulLO</i>	Rat	<i>CaMV 35S</i>	1.5	Red fruit	Methyl viologen, salt, drought	[92]

PMM, phosphomannose mutase; *GFP*, green fluorescent protein; *GMP*, GDP-D-mannose pyrophosphorylase; *GME*, GDP-D-mannose-3',5'-epimerase; *GGP*, GDP-L-galactose phosphorylase; *GPP*, L-galactose-1-P phosphatase; *L-GalDH*, L-galactose dehydrogenase; *L-GalLDH*, L-GalLDH L-galactono-1,4-lactone dehydrogenase; *L-GulLO*, L-gulono-1,4-lactone oxidase; *MIOX*, myo-inositol oxygenase; *D-GalUR*, D-galacturonate reductase; *ALO*, D-arabinono-1,4-lactone oxidase; *NCED*, 9-cis-epoxycarotenoid dioxygenase; *CaMV 35S*, cauliflower mosaic virus 35S constitutive promoter; *ZmUbi*, maize ubiquitin constitutive promoter; *MgGMP*, native promoter of acerola *GMP*; *OsLP2*, rice leaf panicle 2 leaf-specific promoter; *StPAT*, potato polyubiquitin constitutive promoter; *PspetE*, pea plastocyanin constitutive promoter; *SLPG*, tomato polygalacturonase fruit-specific promoter; ~, approximately; ns, non-significant.

PMII gene in *Arabidopsis* did not significantly change ascorbate concentrations, indicating that PMI is not a rate-limiting step in the L-galactose pathway in *Arabidopsis* [100].

1.5.1.2. Phosphomannose Mutase (PMM)

The PMM enzyme is responsible for the second enzymatic step of the L-galactose pathway, catalyzing a reversible conversion of D-mannose-6-P to D-mannose-1-P (Figure 1.2). Temperature-sensitive mutations in the *Arabidopsis* *PMM* gene reduced ascorbate to 20% of wild-type levels, supporting its important role in L-galactose ascorbate biosynthesis [101]. Increased expression of the *PMM* gene has increased ascorbate concentrations 1.3-fold in *Arabidopsis* [61] and 2.5-fold in tobacco (*Nicotiana tabacum* L.) [60] (Table 1.2). The small increase in ascorbate concentrations observed in the *Arabidopsis* plants with increased expression of the *PMM* gene was not proportionate to the highly upregulated PMM enzymatic activity indicating that PMM is not a rate-limiting step of the L-galactose pathway in *Arabidopsis* [61]. Increased expression of the *PMM* gene in *Arabidopsis* was associated with enhanced tolerance to methyl viologen-induced stress [61] (Table 1.2).

1.5.1.3. GDP-D-Mannose Pyrophosphorylase (GMP)

The GMP (also known as VITAMIN C-1 [VTC1]) enzyme is responsible for the third enzymatic step of the L-galactose pathway, catalyzing a reversible conversion of D-mannose-1-P to GDP-D-mannose (Figure 1.2). Missense mutations in the *Arabidopsis* *GMP* gene reduced ascorbate to 25% of wild-type levels, supporting its important role in L-galactose ascorbate biosynthesis [102-104]. Increased expression of the *GMP* gene has increased ascorbate concentrations in a range of species including 1.3-fold in *Arabidopsis* [62], 2.5-fold in tobacco [66], 1.4- to 1.7-fold in tomato (*Solanum lycopersicum* L.) [67, 84], and 1.5-fold in rice [64] (Table 1.2). Increased co-expression of the *GMP* + *GME* genes in tomato increased ascorbate concentrations 2.0-fold, a greater increase than the individual genes alone, suggesting that GMP may act synergistically with GME to increase ascorbate concentrations in tomato [67]. Increased co-expression of the *GMP* + *GME* + *GGP* + *GPP* genes in tomato also increased ascorbate concentrations 2.0-fold, however this was the same fold-change for increased expression of the *GGP* gene alone, suggesting that GMP is not a major rate-limiting step in the L-galactose pathway in tomato [67]. This conclusion is also supported by other studies increasing the expression the *GMP* gene in *Arabidopsis* and tobacco

and increasing the co-expression the *GMP* + *GME* genes in tobacco that did not report significant changes in ascorbate concentrations [63, 65] (Table 1.2).

1.5.1.4. GDP-D-Mannose-3',5'-Epimerase (GME)

The GME enzyme is responsible for the fourth enzymatic step of the L-galactose pathway, catalyzing a reversible conversion of GDP-D-mannose to produce GDP-L-galactose (Figure 1.2). The GME enzyme also catalyzes a reversible conversion of GDP-D-mannose to produce GDP-L-gulose. Consequently, an alternative L-gulose pathway has been proposed to contribute towards ascorbate biosynthesis in plants (see Section 2.2) [105]. Knock-down of the *GME* gene in tomato by RNAi reduced ascorbate to 44% of wild-type levels, supporting its important role in L-galactose ascorbate biosynthesis [106]. Increased expression of the *GME* gene has increased ascorbate concentrations 1.4- to 1.9-fold in *Arabidopsis* [62, 68, 69], 1.4- to 1.8-fold in tomato [67, 70], and 1.9-fold in rice [64] (Table 1.2). Increased expression of the *GMP* + *GME* + *GGP* + *GPP* genes in tomato only increased ascorbate concentrations 2.0-fold, the same fold-change for increased expression of the *GGP* gene alone suggesting that GME is not a major rate-limiting step in the L-galactose pathway in tomato [67]. Moreover, increased expression of the *GME* gene in tobacco did not significantly change ascorbate concentrations, suggesting that GME is not a rate-limiting step of the L-galactose pathway in tobacco [65]. Increased expression of the *GME* gene in plants has been associated with enhanced tolerance to salt, low pH, drought, cold, and methyl viologen-induced stress [64, 68, 70] (Table 1.2).

1.5.1.5. GDP-L-Galactose Phosphorylase (GGP)

The GGP (also known as VTC2/5) enzyme is responsible for the fifth enzymatic step of the L-galactose pathway, catalyzing the conversion of GDP-L-galactose to L-galactose1-P and represents the first committed step toward ascorbate biosynthesis since GDP-D-mannose and GDP-L-galactose are also used as precursors for cell wall polysaccharides and glycoproteins (Figure 1.2). Splice junction and missense mutations in the *Arabidopsis* *GGP* gene reduced ascorbate to 10% of wild-type levels, supporting its critical role in L-galactose ascorbate biosynthesis [12, 13, 107, 108]. The *ggp1(vtc2)/ggp2(vtc5)* double mutants in *Arabidopsis* also exhibit a seedling-lethal phenotype without the supplementation of L-galactose or ascorbate [12, 13]. These results support the pivotal role of the *GGP* gene in L-galactose ascorbate biosynthesis,

as well as highlight the importance of L-galactose ascorbate biosynthesis for plant growth and development. Increased expression of the *GGP* gene has increased ascorbate concentrations in a wide range of species including 2.9- to 4.1-fold in *Arabidopsis* [62, 71], 3.1-fold in potato (*Solanum tuberosum* L.) [73], 2.1-fold in strawberry (*Fragaria × ananassa*) [73], 2.0- to 6.2-fold in tomato [67, 73], 1.4-fold in tobacco [74], and 2.5- to 2.6-fold in rice [64, 72] (Table 1.2). Increased co-expression of the *GGP* + *L-GalLDH* genes or the *GGP* + *GPP* genes increased ascorbate concentrations 3.6- and 4.1-fold, respectively, in *Arabidopsis*, a greater fold-change than increased expression of the *GGP* gene alone, suggesting that GGP may act synergistically with L-GalLDH and GPP in the L-galactose pathway [62]. However, increased co-expression of the *GGP* + *GPP* genes in tomato only increased ascorbate concentrations 1.8-fold, less than the fold-change for increased expression of the *GGP* gene alone, suggesting that GGP and GPP do not act synergistically in the L-galactose pathway in tomato [67]. Relative to other genes from the L-galactose pathway, increased expression of the *GGP* gene consistently leads to the greatest increase in ascorbate concentrations, providing strong genetic evidence that GGP represents the rate-limiting enzymatic step of the L-galactose pathway in many species [62, 64, 67, 71, 100]. Consistent with being the first committed step towards ascorbate biosynthesis and rate-limiting step of the L-galactose pathway, the strongest regulation of ascorbate concentrations in plants has been proposed to occur through the *GGP* gene [17]. Increased expression of the *GGP* gene in plants has been associated with enhanced tolerance to salt, ozone, and cold stress [64, 72, 74] (Table 1.2).

1.5.1.6. L-Galactose-1-Phosphate Phosphatase (GPP)

The GPP (also known as VTC4) enzyme is responsible for the sixth enzymatic step of the L-galactose pathway, catalyzing the conversion of L-galactose1-P to L-galactose (Figure 1.2). A missense mutation in the *Arabidopsis* *GPP* gene reduced ascorbate to 50% of wild-type levels, supporting its important role in L-galactose ascorbate biosynthesis [107, 109]. Increased expression of the *GPP* gene has increased ascorbate concentrations 1.5-fold in *Arabidopsis* [62], 1.7-fold in tomato [67], and 1.4-fold in rice [64] (Table 1.2). Increased co-expression of the *GGP* + *GPP* genes or the *GMP* + *GME* + *GGP* + *GPP* genes in tomato did not increase ascorbate concentrations higher than increased expression of the *GGP* gene alone, indicating that GPP is not a major rate-limiting step in the L-galactose pathway in tomato [67].

1.5.1.7. L-Galactose Dehydrogenase (L-GalDH)

The L-GalDH enzyme is responsible for the penultimate enzymatic step of the L-galactose pathway, catalyzing the conversion of L-galactose to L-galactono-1,4-lactone (Figure 1.2). Knock-down of the *L-GalDH* gene in *Arabidopsis* with RNAi had no effect on ascorbate concentrations under low light conditions, but reduced ascorbate to 45% of wild-type levels under high light conditions when a larger pool of ascorbate is required, supporting its role in L-galactose ascorbate biosynthesis [75]. Increased expression of the *L-GalDH* gene increased ascorbate concentrations 1.2-fold in *Arabidopsis* [62] and 1.7-fold in rice [64] (Table 1.2). However, increased expression of the *L-GalDH* gene in tobacco did not significantly change ascorbate concentrations, despite having a 3.5-fold increase in L-GalDH enzymatic activity, indicating that L-GalDH is not a rate-limiting step in the L-galactose pathway in tobacco [75].

1.5.1.8. L-Galactono-1,4-Lactone Dehydrogenase (L-GalLDH)

The L-GalLDH enzyme is responsible for the final enzymatic step of the L-galactose pathway, catalyzing the conversion of L-galactono-1,4-lactone to ascorbate (Figure 1.2). Similar to the *ggp1(vtc2)/ggp2(vtc5)* double mutants in *Arabidopsis*, loss-of-function *Arabidopsis l-galldh* mutants exhibited a seedling lethal phenotype in the absence of ascorbate supplementation, supporting the critical role of L-GalLDH in L-galactose ascorbate biosynthesis [110]. Increased expression of the *L-GalLDH* gene has increased ascorbate concentrations in a range of species, including 1.8-fold in *Arabidopsis* [62], 1.3- to 3.2-fold in lettuce (*Lactuca sativa* L.) [76, 77], 7.0-fold in lily (*Lilium davidii*) [78], 1.3- to 1.5-fold in rice [64, 79], and 2.1-fold in tobacco [80] (Table 1.2). One study, however, found that increased expression of the *L-GalLDH* gene in tobacco did not significantly change ascorbate concentrations, despite having 10-fold higher L-GalLDH enzymatic activity, suggesting that L-GalLDH is not a major rate-limiting step of the L-galactose pathway in tobacco [81]. Increased co-expression of the *GGP* + *L-GalLDH* genes increased ascorbate concentrations 3.6-fold in *Arabidopsis*, a greater fold-change than increased expression of each gene alone, suggesting that L-GalLDH may act synergistically with GGP in the L-galactose pathway in *Arabidopsis* [62]. Increased expression of the *L-GalLDH* gene in tobacco was associated with enhanced tolerance to salt and methyl viologen-induced stress [80] (Table 1.2).

1.5.2. The L-Gulose Pathway

The L-gulose pathway (also known as the L-gulose shunt), is a result of the epimerization of GDP-D-mannose to GDP-L-gulose by GME, in addition to the epimerization of GDP-D-mannose to GDP-L-galactose [105] (Figure 1.2). In this pathway, GDP-L-gulose is proposed to be converted to the ascorbate precursor L-gulono-1,4-lactone via L-gulose-1-P and L-gulose intermediates, however the genes and enzymes responsible for catalyzing these steps in plants have not yet been identified. The GGP and L-GalDH enzymes from the L-galactose pathway have been proposed to catalyze these reactions, however GGP does not appear to have any enzymatic activity on GDP-L-gulose, and L-GalDH has a much lower affinity for L-gulose than L-galactose [75, 111, 112]. L-gulono-1,4-lactone is then proposed to be converted to ascorbate by L-GulLO (Figure 1.2). Increased expression of the rat (*Rattus rattus*) *L-GulLO* gene has increased ascorbate concentrations in a range of species including 1.8- to 2.0-fold in *Arabidopsis* [85, 97], 7.0-fold in lettuce [98], 2.4-fold in potato [99], 7.0-fold in tobacco [98], and 1.5-fold in tomato [92] (Table 1.2). However, increased expression of the *Arabidopsis L-GulLOs* genes in *Arabidopsis* did not change ascorbate concentrations [82] (Table 1.2) and increased expression of the *Arabidopsis L-GulLOs* genes in tobacco cells only increased ascorbate concentrations when supplied with L-gulono-1,4-lactone, together suggesting that the L-gulose pathway is not a major contributor to ascorbate biosynthesis in plants [113]. Increased expression of the rat *L-GulLO* gene in plants has been associated with enhanced tolerance to salt, cold, heat, drought, and pyrene- and methyl viologen-induced stress [85, 92, 99] (Table 1.2).

1.5.3. The Myo-Inositol Pathway

The *myo*-inositol pathway (also known as the D-gluconate pathway), which operates in animals, has also been proposed to be present in plants (Figure 1.2). In this pathway, *myo*-inositol is proposed to be converted to the ascorbate precursor L-gulono-1,4-lactone via D-gluconate and D-gulonate intermediates, however the gene and enzyme responsible for catalyzing the conversion of D-gluconate to D-gulonate has not yet been identified (Figure 1.2). L-gulono-1,4-lactone is then proposed to be converted to ascorbate by L-GulLO (Figure 1.2). The contribution of the *myo*-inositol pathway towards ascorbate biosynthesis in plants has been controversial. Supporting its contribution are studies that increased the expression of the *myo-inositol oxygenase (MIOX)* gene, which encodes the enzyme responsible for catalyzing the conversion of *myo*-inositol to D-

glucuronate (Figure 1.2). For example, increased expression of the *MIOX* gene in *Arabidopsis* has increased ascorbate concentrations 1.6- to 3.0-fold, and was associated with enhanced tolerance to salt, cold, heat, and pyrene-induced stress [27, 85, 86] (Table 1.2). Varying results were obtained when the expression of the *MIOX* gene was increased in tomato, with ascorbate reduced to 50% of wild-type levels in leaves, but increased 1.4- and 1.3-fold in green and red fruit, respectively, indicating that the *myo*-inositol pathway may be tissue-specific in tomato [84]. Further support for the *myo*-inositol pathway comes from increased expression of phytases that are proposed to increase the pool of *myo*-inositol that can then serve as an entry point for ascorbate biosynthesis via the *myo*-inositol pathway. For example, increased expression of the *Arabidopsis* purple acid phosphatase 15 (*PAP15*) gene [114] and the bacterial beta-propeller phytase PHY-US417 gene [115] in *Arabidopsis* has increased ascorbate concentrations 2.0- to 2.3-fold and were associated with enhanced tolerance to drought and salt stress. Loss-of-function mutations in the *Arabidopsis* *PAP15* gene reduced ascorbate to 70% of wild-type levels, supporting the contribution of *PAP15* toward *myo*-inositol ascorbate biosynthesis [114]. The *PAP15* gene, however, has also been proposed to encode a GPP in *Arabidopsis*, which may provide an alternative explanation for these results [10]. While there are several studies supporting the contribution of the *myo*-inositol pathway towards ascorbate biosynthesis in plants, there are equally as many studies opposing its contribution. For example, increased expression of the *MIOX* gene in *Arabidopsis* [87] and rice [83] did not change ascorbate concentrations, despite the *Arabidopsis* lines with increased *MIOX* expression having higher *MIOX* enzymatic activity (Table 1.2). Further, loss-of-function mutations in the *Arabidopsis* [116] and rice [14] *MIOX* genes did not change ascorbate concentrations relative to wild-type, challenging the view that the *myo*-inositol pathway contributes towards ascorbate biosynthesis. Further opposition comes from a recent study utilizing CRISPR/Cas9-induced knockout mutations in the *Arabidopsis* *glucuronokinase1* (*GlcAK1*) gene, and radiotracer experiments with ³H-*myo*-inositol. Knockout of *GlcAK1* ought to increase the flux from *myo*-inositol to ascorbate if the *myo*-inositol pathway is operational, however no changes in ascorbate concentrations were detected in the *Arabidopsis* *glcak1* mutants. Moreover, radiolabeled ³H-*myo*-inositol in the *Arabidopsis* *glcak1* mutants accumulated as D-gluconate and ascorbate remained unlabeled, leading the authors to conclude that the *myo*-inositol pathway does not contribute toward ascorbate biosynthesis in plants [117].

1.5.4. The D-Galacturonate Pathway

Evidence for the D-galacturonate pathway was discovered in strawberry fruits [88]. In this pathway, D-galacturonate derived from pectin in the cell wall is proposed to be converted to the ascorbate precursor L-galactono-1,4-lactone via a L-galactonate intermediate, however the gene and enzyme responsible for converting L-galactonate to L-galactono-1,4-lactone has not yet been identified in plants (Figure 1.2). L-galactono-1,4-lactone is then proposed to be converted to ascorbate by L-GalLDH, the final enzymatic step of the L-galactose pathway (Figure 1.2). To date, the *D-galacturonate reductase (D-GalUR)* gene, which encodes the enzyme responsible for catalyzing the conversion of D-galacturonate to L-galactonate, has only been identified in strawberry (Figure 1.2). Knock-out or knock-down studies of the *D-GalUR* gene in strawberry or other plant species have not yet been reported, making it difficult to determine the relative contribution of the D-galacturonate pathway toward ascorbate biosynthesis in plants. Nevertheless, increased expression of the *D-GalUR* gene has increased ascorbate concentrations 3.0-fold in *Arabidopsis* [88], 1.7- to 2.1-fold in potato [89-91], and 1.0- to 2.5-fold in tomato [93, 94, 118]. Increased expression of the *D-GalUR* gene in plants has been associated with enhanced tolerance to salt, drought, cold, zinc chloride, and methyl viologen-induced stress [89-91, 93, 118] (Table 1.2).

1.6. Recycling of Ascorbate in Plants

In addition to *de novo* biosynthesis of ascorbate, oxidized ascorbate can be regenerated via the recycling enzymes MDAR and DHAR to maintain cellular ascorbate levels (Figure 1.1). Following ascorbate oxidation, a short-lived monodehydroascorbate radical is produced. Monodehydroascorbate can then be recycled back to ascorbate by MDAR using NADPH as a reductant, however if not recycled monodehydroascorbate will rapidly undergo non-enzymatic disproportionation with another monodehydroascorbate molecule to produce an ascorbate and dehydroascorbate molecule [119]. Dehydroascorbate can then be recycled back to ascorbate by DHAR using glutathione as a reductant, however if not recycled dehydroascorbate will undergo an irreversible hydrolysis to generate 2,3-diketogulonic acid [119]. An overview of metabolic engineering strategies utilizing ascorbate recycling genes to increase ascorbate concentrations in model and crop species, including the tissue examined and any stress tolerances observed, is summarized in Table 1.3.

Loss-of-function mutations in the *Arabidopsis MDAR4* gene reduced the ratio of ascorbate to dehydroascorbate, but did not affect the pool size of total ascorbate (ascorbate + dehydroascorbate), supporting the role of MDAR in ascorbate recycling [120]. Increased expression of the *MDAR* gene has increased the concentrations of reduced ascorbate 1.2- to 2.2-fold in tobacco [121-123] and 1.2-fold in tomato [124] (Table 1.3). However, other studies increasing the expression of the *MDAR* gene in tobacco [125] and tomato [126] did not report any significant changes to reduced ascorbate concentrations, indicating that MDAR may not be a rate-limiting step in the recycling of ascorbate in these species (Table 1.3). Increased expression of the *MDAR* gene in plants has been associated with enhanced tolerance to salt, ozone, drought, cold, heat, and methyl viologen-induced stress [121, 122, 124, 125] (Table 1.3).

Similar to the *mdar4* mutant in *Arabidopsis*, loss-of-function mutations in the *Arabidopsis DHAR3* gene reduced the ratio of ascorbate to dehydroascorbate, but did not affect the pool size of total ascorbate, supporting the role of DHAR in ascorbate recycling [127]. Increased expression of the *DHAR* gene has increased the concentrations of reduced ascorbate in a wide range of species including 1.1- to 3.3-fold in *Arabidopsis* [69, 128-132], 1.9- to 6.1-fold in maize [133, 134], 1.3- to 2.8-fold in potato [135-137], 1.7-fold in rice [138], 1.3- to 3.9-fold in tobacco

Table 1.3. Overview of metabolic engineering strategies utilizing ascorbate recycling genes to increase ascorbate concentrations in model and crop species.

Species transformed	Gene(s) transformed	Gene source(s)	Promoter	Max fold-change	Tissue examined	Stress tolerance	Reference
Tobacco	<i>MDAR</i>	Acerola	<i>CaMV 35S</i>	2.0	Leaves	Salt	[121]
Tobacco	<i>MDAR</i>	Mangrove	<i>CaMV 35S</i>	~1.3 ^{ns}	Leaves	Salt	[125]
Tomato	<i>MDAR</i>	Tomato	<i>CaMV 35S</i>	1.2	Leaves	Cold, heat, methyl viologen	[124]
Tomato	<i>MDAR</i>	Tomato	<i>FMV 34S</i>	1.2 ^{ns} , 1.0 ^{ns}	Leaves, red fruit	-	[126]
Tobacco	<i>MDAR1</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	2.2	Leaves	Ozone, salt, drought	[122]
Tobacco	<i>MDAR1</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~1.2	Roots	-	[123]
<i>Arabidopsis</i>	<i>DHAR</i>	Chinese tulip tree	<i>CaMV 35S</i>	~1.4	Leaves	Salt, drought	[128]
<i>Arabidopsis</i>	<i>DHAR</i>	Rose	<i>CaMV 35S</i>	3.0	Leaves	-	[69]
Maize	<i>DHAR</i>	Wheat	<i>ZmUbi</i>	1.8, 1.9	Leaves, kernels	-	[133]
Potato	<i>DHAR</i>	Sesame	<i>CaMV 35S</i>	1.5, 1.6	Leaves, tubers	-	[135]
Potato	<i>DHAR</i>	Sesame	<i>PtPal</i>	1.3	Tubers	-	[135]
Tobacco	<i>DHAR</i>	Wheat	<i>CaMV 35S</i>	2.1	Leaves	Ozone	[139]
Tobacco	<i>DHAR</i>	Human	<i>CaMV 35S</i>	1.6	Leaves	Methyl viologen, H ₂ O ₂ , cold, salt	[140]
Tobacco	<i>DHAR</i>	Human	<i>CaMV 35S</i>	1.6, 2.0	Young leaves, mature leaves	Methyl viologen	[141]
Tobacco	<i>DHAR</i>	Wheat	<i>CaMV 35S</i>	2.4, 3.9, 2.2	Expanding leaves, mature leaves, presenescent leaves	-	[133]
Tomato	<i>DHAR</i>	Tomato	<i>FMV 34S</i>	1.1 ^{ns} , 1.6, 1.6	Leaves, green fruit, red fruit	-	[126]
Tobacco	<i>DHAR</i> + <i>CuZnSOD</i> + <i>APX</i>	Human + pea + pea	<i>CaMV 35S</i>	1.5	Leaves	Methyl viologen, salt	[142]
<i>Arabidopsis</i>	<i>DHAR1</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	3.3	Leaves	High light, heat, paraquat	[129]
<i>Arabidopsis</i>	<i>DHAR1</i>	Kiwifruit	<i>CaMV 35S</i>	~1.5	Leaves	Salt	[130]
<i>Arabidopsis</i>	<i>DHAR1</i>	Rice	<i>CaMV 35S</i>	1.2	Leaves	Salt	[131]
Maize	<i>DHAR1</i>	Rice	<i>HvHor</i>	6.1	Kernels	-	[134]
Potato	<i>DHAR1</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	2.8	Leaves	Methyl viologen, drought, salt	[136]
Potato	<i>DHAR1</i>	Potato	<i>CaMV 35S</i>	1.7, 1.3	Leaves, tubers	-	[137]
Rice	<i>DHAR1</i>	Rice	<i>ZmUbi</i>	~1.7	Leaves	-	[138]
Tobacco	<i>DHAR1</i>	Rice	<i>PsPrn</i>	1.6	Leaves	Salt, cold	[143]

Table 1.3. (Continued).

Species transformed	Gene(s) transformed	Gene source(s)	Promoter	Max fold-change	Tissue examined	Stress tolerance	Reference
Tomato	<i>DHAR1</i>	Potato	<i>CaMV 35S</i>	1.9, 1.4	Leaves, fruit	Methyl viologen, salt	[144]
Tobacco	<i>DHAR1 + GR</i>	Rice + <i>E. coli</i>	<i>PsPrn</i>	~2.5	Leaves	Salt, cold, methyl viologen	[143]
<i>Arabidopsis</i>	<i>DHAR2</i>	Kiwifruit	<i>CaMV 35S</i>	~1.4	Leaves	Salt	[145]
Potato	<i>DHAR2</i>	Potato	<i>CaMV 35S</i>	1.5, ~1.1 ^{ns}	Leaves, tubers	-	[137]
Tobacco	<i>DHAR2</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	2.1	Leaves	Ozone, drought, salt	[146]
Tobacco	<i>DHAR2</i>	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~1.3	Roots	Aluminium	[123]
Tomato	<i>DHAR2</i>	Potato	<i>CaMV 35S</i>	1.8, ~1.1 ^{ns}	Leaves, fruit	Methyl viologen, salt	[144]
Tomato	<i>DHAR2</i>	Pear	<i>CaMV 35S</i>	1.5	Leaves	Salt, cold	[147]
<i>Arabidopsis</i>	<i>DHAR3</i>	Sweet potato	<i>CaMV 35S</i>	~1.1 ^{ns}	Leaves	Salt, drought	[132]

MDAR, monodehydroascorbate reductase; *DHAR*, dehydroascorbate reductase; *CuZnSOD*, copper zinc superoxide dismutase; *APX*, ascorbate peroxidase; *GR*, glutathione reductase; *E. coli*, *Escherichia coli*; *CaMV 35S*, cauliflower mosaic virus 35S constitutive promoter; *FMV 34S*, figwort mosaic virus 34S constitutive promoter; *ZmUbi*, maize *ubiquitin* constitutive promoter; *PtPal*, potato *patatin* tuber-specific promoter; *HvHor*, barley *D-hordein* endosperm-specific promoter; *PsPrn*, pea plastid rRNA operon constitutive promoter; ~, approximately; ns, non-significant.

[123, 133, 139-143, 146], and 1.5-fold to 1.9-fold in tomato [126, 144, 147] (Table 1.3). These studies suggest that DHAR is the rate-limiting step in ascorbate recycling in plants relative to MDAR. Increased expression of the *DHAR* gene in plants has been associated with enhanced tolerance to salt, drought, high light, heat, ozone, cold, aluminum, hydrogen peroxide, and methyl viologen-induced stress [123, 128-132, 136, 139-144, 146, 147] (Table 1.3).

1.7. Regulation of Ascorbate Biosynthesis in Plants

The genes and enzymes responsible for the major biosynthetic route towards ascorbate biosynthesis—the L-galactose pathway—and the recycling of ascorbate in plants have been well characterized for over a decade now [108, 112]. The regulation of ascorbate biosynthesis and recycling, on the other hand, has remained relatively unexplored until more recently. Several key regulators have now been identified, and it is becoming clear that ascorbate biosynthesis and recycling is tightly regulated at the transcriptional, translational, and post-translational level [17]. An overview of metabolic engineering strategies utilizing ascorbate regulatory factors to increase ascorbate concentrations in model and crop species, including the tissue examined and any stress tolerances observed, is summarized in Table 1.4.

1.7.1. Ascorbic Acid Mannose Pathway Regulator 1 (AMR1)

The *ascorbic acid mannose pathway regulator 1 (AMR1)* gene was identified in an activation-tagged *Arabidopsis* mutant that contained 40% of wild-type ascorbate levels [148]. The *AMR1* gene encodes an F-box protein that negatively regulates the transcription of genes of the L-galactose pathway, including *GMP*, *GME*, *GGP1*, *GPP*, *L-GalDH*, and *L-GalLDH* [148]. The expression of the *AMR1* gene was developmentally and environmentally controlled, with increased expression as the plants aged and decreased expression under high light [148]. Loss-of-function mutations in the *AMR1* gene in *Arabidopsis* increased ascorbate concentrations 2.0-fold and was associated with enhanced ozone stress [148] (Table 1.4).

1.7.2. Basic Helix-Loop-Helix 59 Transcription Factor (bHLH59)

The tomato *basic helix-loop-helix 59 (bHLH59)* gene was recently identified as the likely candidate underlying the ascorbate quantitative trait locus TFA9 in tomato [149]. The bHLH59 transcription factor positively regulates the transcription of genes of the L-galactose pathway, including *PMI*, *PMM*, *GMP1*, *GMP2*, *GMP3*, *GMP4*, and *GME1* [149]. Knock-down of the *bHLH59* gene with RNAi in tomato reduced ascorbate to 65% of wild-type levels [149]. Increased expression of the *bHLH59* gene in tomato increased ascorbate concentrations 1.5-fold and was associated with enhanced tolerance to methyl viologen-induced stress [149] (Table 1.4).

Table 1.4. Overview of metabolic engineering strategies utilizing ascorbate regulatory factors to increase ascorbate concentrations in model and crop species.

Species	Regulatory Factor	Strategy	Gene source	Promoter	Max fold-change	Tissue examined	Stress tolerance	Reference
<i>Arabidopsis</i>	<i>AMR1</i>	T-DNA insertion	-	-	2.0	Leaves	Ozone	[148]
Tomato	<i>bHLH59</i>	Increased expression	Tomato	<i>CaMV 35S</i>	~1.5	Fruit	Methyl viologen	[149]
<i>Arabidopsis</i>	<i>CSN5B</i>	T-DNA insertion	-	-	~1.4	Seedlings	Salt	[150]
<i>Arabidopsis</i>	<i>CSN8</i>	T-DNA insertion	-	-	~1.8	Seedlings	-	[150]
Tomato	<i>Dof22</i>	RNAi	-	-	1.3, 1.6	Leaves, red fruit	-	[151]
<i>Arabidopsis</i>	<i>ERF98</i>	Increased expression	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.7	Leaves	Salt	[152]
<i>Arabidopsis</i>	<i>GGP1</i> uORF	Genome editing	-	-	1.7	Leaves	-	[153]
Lettuce	<i>GGP1</i> uORF	Genome editing	-	-	1.4	Leaves	Methyl viologen	[153]
Lettuce	<i>GGP2</i> uORF	Genome editing	-	-	2.6	Leaves	Methyl viologen	[153]
Tomato	<i>GGP2</i> uORF	Genome editing	-	-	~1.4	Leaves	-	[154]
Tomato	<i>HZ24</i>	Increased expression	Tomato	<i>CaMV 35S</i>	1.5, ~1.2	Leaves, breaker fruit	Methyl viologen	[155]
<i>Arabidopsis</i>	<i>KJC1</i>	Increased expression	<i>Arabidopsis</i>	<i>CaMV 35S</i>	1.4	Seedlings	-	[63]
<i>Arabidopsis</i>	<i>KJC1</i> + <i>GMP</i>	Increased expression	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~1.0 ^{ns}	Seedlings	-	[63]
<i>Arabidopsis</i>	<i>KJC2</i>	Increased expression	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~1.0 ^{ns}	Seedlings	-	[63]
Tobacco	<i>MYB5</i>	Increased expression	Pear	<i>CaMV 35S</i>	~1.3	Leaves	Cold	[156]
Tomato	<i>NL33</i>	RNAi	-	-	2.7, 1.3	Leaves, red fruit	Methyl viologen, <i>Botrytis cinerea</i>	[157]
Potato	<i>nsLTP1</i>	Increased expression	Potato	<i>CaMV 35S</i>	2.3	Leaves	Heat, drought, salt	[158]
<i>Arabidopsis</i>	<i>VTC3</i>	Increased expression	<i>Arabidopsis</i>	<i>CaMV 35S</i>	~0.8 ^{ns}	Seedlings	-	[159]
<i>Arabidopsis</i>	<i>WAX1</i>	Increased expression	Saltwater cress	<i>SP</i>	1.3	Leaves	-	[160]
<i>Arabidopsis</i>	<i>WAX1</i>	Increased expression	Saltwater cress	<i>AtRD29A</i>	~1.4*	Leaves	Drought	[160]
<i>Arabidopsis</i>	<i>ZF3</i>	Increased expression	Tomato	<i>CaMV 35S</i>	1.8	Leaves	Salt	[161]
Tomato	<i>ZF3</i>	Increased expression	Tomato	<i>CaMV 35S</i>	~2.1	Leaves	Salt	[161]

AMR1, ascorbic acid mannose pathway regulator 1; *CSN5B*, COP9 signalosome subunit 5B; *Dof22*, DNA-binding with one finger 22; *ERF98*, ethylene response factor 98; *GGP* uORF, GDP-L-galactose phosphorylase upstream open reading frame; *HZ24*, HD-Zip I transcription factor 24; *KJC*, KONJAC; *MYB5*, myeloblastosis transcription factor 5; *NL33*, NBS-LRR 33; *nsLTP1*, non-specific lipid

transfer protein-1; *VTC3*, *VITAMIN C-3*; *CaMV 35S*, cauliflower mosaic virus 35S constitutive promoter; *SP*, super-promoter constitutive promoter; *AtRD29A*, *Arabidopsis* stress-inducible promoter; ~, approximately; ns, non-significant; *under drought stress.

1.7.3. Calmodulin-Like 10 (CML10)

A calcium sensor, calmodulin-like 10 (CML10), was recently identified to interact with *Arabidopsis* PMM in a yeast two-hybrid screen [162]. The expression of the *CML10* gene was upregulated under oxidative stress, and the CML10 protein increased PMM enzymatic activity in a calcium-dependent manner [162]. Knock-down of the *CML10* gene in *Arabidopsis* by RNAi reduced ascorbate to 70% of wild-type levels and was associated with decreased drought and H₂O₂-induced stress [162]. To date, no studies have investigated increased expression of the *CML10* gene in plants.

1.7.4. COP9 Signalosome Subunit 5B and 8 (CSN5B and CSN8)

The *Arabidopsis* COP9 signalosome subunit 5B (CSN5B), a component of the photomorphogenic COP9 signalosome, was identified to interact with GMP in a yeast two-hybrid screen [150]. Under darkness, CSN5B promotes the ubiquitination and degradation of GMP through the 26S proteasome pathway [150]. Loss-of-function mutations in the *CSN5B* gene in *Arabidopsis* increased ascorbate concentrations 1.4-fold and was associated with enhanced salt stress [150] (Table 1.4). Loss-of-function mutations in the *CSN8* gene in *Arabidopsis* also increased ascorbate concentrations 1.8-fold (Table 1.4), however CSN8 did not interact with GMP in the yeast-two hybrid screen, and its mechanism of action has yet to be determined [150].

1.7.5. DNA-Binding with One Finger 22 (Dof22)

The DNA-binding with one finger 22 (Dof22), a member of the Dof family of plant specific transcription factors, has recently been proposed as a negative transcriptional regulator of L-galactose pathway, ascorbate recycling, and ascorbate-glutathione cycle genes in tomato, such as *GGP1*, *GGP2*, *L-GalDH*, *L-GalLDH*, *MDAR*, cytosolic *APX*, and *glutathione reductase (GR)* [151]. Knock-down of the *Dof22* gene with RNAi in tomato increased ascorbate concentrations up to 1.6-fold (Table 1.4), but was associated with reduced tolerance to salt stress due to the transcriptional downregulation of the *salt overly sensitive 1 (SOS1)* gene, which encodes a plasma membrane Na⁺/H⁺ antiporter [151].

1.7.6. Ethylene Response Factor 98 (ERF98)

In a reverse genetic screen of a pool of *Arabidopsis ethylene response factor (ERF)* T-DNA insertion mutants, loss-of-function mutations in the *ERF98* gene reduced ascorbate to 65% of wild-type levels [152]. The *ERF98* gene encodes a transcription factor belonging to the AP2/ERF superfamily that positively regulates the transcription of ascorbate biosynthetic, ascorbate recycling, and ascorbate-glutathione cycle genes, including *GMP*, *GGP1*, *L-GalDH*, *L-GalLDH*, *MIOX4*, *MDAR3*, chloroplastic *DHAR*, cytosolic *DHAR*, and *GRI* [152]. Increased expression of the *ERF98* gene in *Arabidopsis* increased ascorbate concentrations 1.7-fold and was associated with enhanced tolerance to salt stress [152] (Table 1.4).

1.7.7. *GGP* Upstream Open Reading Frame (*GGP* uORF)

The translation of *GGP* has been discovered to be post-transcriptionally regulated through a highly conserved, *cis*-acting upstream open reading frame (uORF) in the long 5' leader sequence of the *GGP* mRNA [163, 164]. The *GGP* uORF is proposed to initiate from a non-canonical AUC or ACG start-codon, and encode a 60- to 65-residue long peptide [163]. Disruption of the *GGP* uORF increased ascorbate concentrations when a *GGP* promoter-*uORF*-*GGP* construct was transiently transformed in *Nicotiana benthamiana* [163]. It is proposed that under high ascorbate concentrations the *GGP* uORF is translated and causes ribosomal stalling, thereby preventing translation of the *GGP* major ORF (mORF), while under low ascorbate concentrations, the uORF is skipped and the *GGP* mORF is translated [163]. The precise mechanism of how ascorbate influences the translation of the *GGP* uORF or mORF, however, has yet to be determined. Recently, disruption of the *GGP* uORF with the CRISPR/Cas9 genome editing system increased ascorbate concentrations 1.7-fold in *Arabidopsis* [153], 1.4- to 2.6-fold in lettuce [153], and 1.4-fold in tomato [154] (Table 1.4). Moreover, disruption of the lettuce *GGP* uORFs was associated with enhanced tolerance to methyl viologen-induced stress [153] (Table 1.4). Disrupting the *GGP* uORF with the CRISPR/Cas9 genome editing system has established itself as a very promising transgene-free strategy to engineer plants with enhanced production of *GGP* for increased ascorbate biosynthesis in crops.

1.7.8. HD-Zip I Family Transcription Factor 24 (HZ24)

The HD-Zip I family transcription factor 24 (HZ24) was recently identified based on a yeast-one hybrid assay to the promoter region of tomato *GMP3* [155]. Knock-down mutants of the *HZ24* gene with RNAi in tomato reduced ascorbate to 60% of wild-type levels [155]. The HZ24 transcription factor positively regulates the transcription of a wide range of genes from the L-galactose pathway in tomato, including *PMM*, *GMP4*, *GME1*, *GME2*, *GGP*, *GPP1*, *GPP2*, and *L-GalDH* [155]. Increased expression of the *HZ24* gene in tomato increased ascorbate concentrations 1.5-fold and was associated with enhanced tolerance to methyl viologen-induced stress [155] (Table 1.4).

1.7.9. High-Pigment-1 (HP1)

The *High-Pigment-1* (*HP1*) tomato mutants—previously characterized as having increased carotenoid and flavonoid concentrations—were found to have reduced ascorbate concentrations at various stages of fruit development and ripening relative to wild-type [165]. The *HP1* gene is orthologous to the *Arabidopsis* *UV-DAMAGED DNA-BINDING PROTEIN 1* (*DDB1*) gene which encodes a component of the CUL4-based E3 ligase complex [166]. Fourteen genes involved in ascorbate biosynthesis and recycling were differentially expressed in the developing fruit of the *hp1* mutant, and it was proposed that HP1 positively regulates the transcription of the *GMP* and *GME1* genes, but negatively regulates the transcription of the *L-GalDH* gene in tomato, ultimately resulting in the reduced ascorbate concentrations observed in the developing and ripening fruit [165]. To date, no studies have investigated the effect of increasing the expression of the *HP1* gene on ascorbate concentrations in plants.

1.7.10. KONJAC 1 and 2 (KJC1 and KJC2)

Two nucleotide sugar phosphorylase-like proteins KONJAC 1 and 2 (KJC1 and KJC2) were identified based on their phylogenetic relationship to GMP [63]. Loss-of-function mutations in the *KJC1* and *KJC2* genes in *Arabidopsis* reduced ascorbate to 40% and 70% of wild-type levels, respectively, due to reducing GMP enzymatic activity to 10% and 70% of wild-type levels, respectively [63]. The KJC1 and KJC2 proteins are proposed to post-translationally regulate GMP enzymatic activity, and may function opposite that of the COP9 signalosome in the regulation of GDP-D-mannose biosynthesis [63]. Increased expression of the *KJC1* gene increased ascorbate

concentrations 1.4-fold in *Arabidopsis*, but increased co-expression of the *KJC1* and *GMP* genes or increased expression of the *KJC2* gene did not change ascorbate concentrations [63] (Table 1.4).

1.7.11. Myeloblastosis Transcription Factor 5 (MYB5)

Pear (*Pyrus betulaefolia*) myeloblastosis transcription factor 5 (MYB5) was recently identified in a yeast one-hybrid screen to interact with the promoter region of the *DHAR2* gene [156]. The expression of the *MYB5* gene was induced by a wide range of stresses including cold, drought, and salt stress [156]. The MYB5 protein positively regulates the transcription of ascorbate recycling, ascorbate-glutathione cycle, and stress-associated genes, including *MDAR*, *DHAR2*, *APX*, *cystathionine- β -synthase 1 (CBF1)*, *CBF2*, and *CBF3* [156]. Increased expression of the *MYB5* gene in tobacco increased ascorbate concentrations 1.3-fold and was associated with enhanced tolerance to cold stress [156] (Table 1.4).

1.7.12. NBS-LRR 33 (NL33)

The tomato NBS-LRR 33 (NL33) resistance protein was recently identified based on a candidate gene-based association study with ascorbate concentrations in tomato [157]. The NL33 protein negatively regulates the transcription of ascorbate biosynthetic, ascorbate recycling, and ascorbate-glutathione cycle genes, including *PMM*, *GMP1*, *GMP4*, *GME1*, *GGP1*, *GPP1*, *GPP2*, *L-GalDH*, *L-GalLDH*, *MIOX*, *MDAR*, *DHAR*, and cytosolic *APX* [157]. Knock-down of the *NL33* gene with RNAi in tomato increased ascorbate concentrations 2.7-fold and was associated with enhanced tolerance to methyl viologen-induced stress and gray mold infection (*Botrytis cinerea*) [157] (Table 1.4).

1.7.13. Non-specific Lipid Transfer Protein-1 (nsLTP1)

The potato thermo-tolerance gene *non-specific lipid transfer protein-1 (nsLTP1)* was first identified in a functional screen of potato high-temperature stress-responsive genes to impart heat tolerance in yeast [167]. The nsLTP1 protein in potato has now recently been identified to positively regulate the transcription of ascorbate-glutathione cycle and stress-related genes, including *APX*, *catalase (CAT)*, *superoxide dismutase (SOD)*, *heat shock protein 20 (HSP20)*, *HSP70*, and *HSP90* [158]. Increased expression of the *nsLTP1* gene increased ascorbate

concentrations 2.3-fold in potato and was associated with enhanced tolerance to heat, drought, and salt stress [158] (Table 1.4).

1.7.14. VITAMIN C-3 (VTC3)

The *Arabidopsis vtc3* mutants were first identified in a forward genetic screen for ozone-sensitive mutants, and contained 40% of wild-type ascorbate levels [107]. Over a decade later the *VTC3* gene was mapped to a unique dual-function polypeptide containing an N-terminal protein kinase domain and a C-terminal protein phosphatase domain [159]. The *Arabidopsis vtc3* mutants were reported to be defective in their ability to increase ascorbate concentrations in response to light and heat [159]. Little is known about the mechanism of action for VTC3 but it is proposed to act as a signal transduction protein that is capable of perceiving a change in the environment to influence cellular ascorbate levels [159]. It has been suggested that VTC3 may play a role in the *GGP* uORF sensing mechanism of ascorbate levels [17]. Increased expression of the *VTC3* gene in the *Arabidopsis vtc3* mutants recovered ascorbate concentrations back to wild-type levels, however, increased expression of the *VTC3* gene in wild-type *Arabidopsis* did not significantly change ascorbate concentrations [159] (Table 1.4).

1.7.15. WAX1

The saltwater cress (*Eutrema salsugineum*) *WAX1* gene was identified based on being rapidly induced in abiotic stress-treated saltwater cress plants [160]. The *WAX1* gene encodes a MYB transcription factor that positively regulates wax related genes, as well as ascorbate biosynthetic genes, including *GGPI*, *L-GalDH*, and *MIOX4* [160]. Increased expression of the *WAX1* gene with the constitutive super-promoter in *Arabidopsis* increased ascorbate concentrations 1.3-fold but severely disrupted plant growth and development [160] (Table 1.4). Increased expression of the *WAX1* gene with the *Arabidopsis RD29A* stress-inducible promoter, instead of the constitutive super-promoter, increased ascorbate concentrations 1.4-fold under drought stress and was associated with enhanced tolerance to drought stress [160] (Table 1.4).

1.7.16. Zinc-finger 3 (ZF3)

The tomato C2H2-type *zinc finger 3* (*ZF3*) gene was first identified based on being induced in salt-stressed tomato plants [168]. The ZF3 protein in tomato has now recently been identified to interact

with CSN5B and inhibit the interaction between CSN5B and GMP [161]. By doing so, ZF3 prevents the ubiquitination and degradation of GMP through the 26S proteasome pathway, promoting the accumulation of GMP [161]. Knock-down of the *ZF3* gene with RNAi in tomato reduced ascorbate to 80% of wild-type levels [161]. Increased expression of the tomato *ZF3* gene increased ascorbate concentrations 1.8-fold and 2.1-fold in *Arabidopsis* and tomato, respectively, and was associated with enhanced tolerance to salt stress [161] (Table 1.4).

1.8. Rice and Wheat as Targets for Ascorbate Metabolic Engineering

Rice is the most important source of human calories—providing 3.5 billion humans with more than 20% of their daily caloric intake—and has a global production of more than 700 million tonnes (Table 1.1) [169]. Rice plays an important role as a staple crop in many parts of the world, particularly in Eastern Asia, Southern Asia, and South-East Asia, where it provides a major source of calories for those populations [169]. Not only is rice one of the most important staple crops in the world, but rice also serves as an important model organism for plant biology [170]. Relative to other major cereal crops, rice has a small genome of 430 Mb—approximately 3.5 times the size of the *Arabidopsis* genome—and was the first crop to have its genome sequenced in 2005 [171]. Rice has since served as a model for other cereal crops with much larger genomes, such as maize (2.4 Gb) and wheat (17 Gb). Efficient protocols for the genetic modification of rice via *Agrobacterium tumefaciens*-mediated transformation have also been developed over the past two decades, and adopted by many laboratories for basic and applied research [172]. Further, rice plants display high editing efficiencies with the CRISPR/Cas9 genome editing system with mutation frequencies often observed around 80-85% [173]. Due to the high efficiency of transformation and genome editing in rice, genetic modification and genome editing strategies are often first tested in rice before being applied to other cereals [172].

Bread wheat is the second most important source of human calories after rice, accounting for 20% of calories consumed by humans, is cultivated on more land than any other crop, and has a global production of more than 700 million tonnes (Table 1.1) [169, 174]. Wheat plays an important role as a staple crop in many parts of the world, particularly in Northern Africa, Central Asia, and Western Asia, where it provides a major source of calories for those populations [169]. Wheat has a large and complex hexaploid genome of 17 Gb—approximately 40 times the size of the rice genome—which has created many challenges in the assembly of a high-quality reference sequence. However, in 2018, the International Wheat Genome Sequencing Consortium released the first high-quality annotated reference sequence for the bread wheat genome, presenting an opportunity for researchers to further understand wheat biology and engage in genomics-assisted molecular breeding [174].

Abiotic stresses such as drought, flooding, salinity, and extreme temperatures—which are predicted to be exacerbated by climate change—are major limiting factors in global rice and wheat

productivity [175, 176]. Developing strategies to increase ascorbate concentrations in rice and wheat may provide a means of enhancing the stress tolerance of these plants and help to mitigate the yield losses associated with these abiotic stresses, improving global food security. Moreover, since rice and wheat grains contain negligible levels of ascorbate, developing strategies to increase ascorbate concentrations in rice and wheat grains may provide a novel means of improving human dietary intakes of ascorbate and bioavailable Fe.

1.9. Thesis Aims and Objectives

This PhD project aimed to develop strategies to increase ascorbate concentrations in rice and wheat plants and investigate the effect that increased ascorbate concentrations may have on grain nutrition, grain Fe bioavailability, abiotic stress tolerance, and plant growth and development. This aim was accomplished through the completion of the following three major research objectives:

- 1) Production and characterization of rice with constitutive overexpression of the *OsGGP* coding sequence.
- 2) Production and characterization of rice with CRISPR/Cas9-induced mutations in the *OsGGP* uORF.
- 3) Genome-wide identification and characterization of the *TaGGP* gene family in bread wheat.

Chapter 2.

Effect of *OsGGP* Constitutive Overexpression on Ascorbate Concentration, Salt Tolerance, and Iron Bioavailability in Rice

2.1 Abstract

Ascorbate (vitamin C) is an essential multifunctional molecule for both plants and mammals. In plants, ascorbate is the most abundant water-soluble antioxidant that supports stress tolerance. In humans, ascorbate is an essential micronutrient and promotes iron (Fe) absorption in the gut. Engineering crops with increased ascorbate levels has the potential to improve both crop stress tolerance and human health. Here, rice (*Oryza sativa* L.) plants were engineered to constitutively overexpress the rice *GDP-L-galactose phosphorylase* coding sequence (35S-*OsGGP*) which encodes the rate-limiting enzymatic step of the L-galactose pathway. Ascorbate concentrations were negligible in both null segregant (NS) and 35S-*OsGGP* brown rice (BR) but significantly increased in 35S-*OsGGP* germinated brown rice (GBR) relative to NS. Foliar ascorbate concentrations were significantly increased in 35S-*OsGGP* plants in the vegetative growth phase relative to NS, but significantly reduced at the reproductive growth phase and were associated with reduced *OsGGP* transcript levels. The 35S-*OsGGP* plants did not display altered salt tolerance at the vegetative growth phase despite having elevated ascorbate concentrations. Ascorbate concentrations were positively correlated with ferritin concentrations in Caco-2 cells—an accurate predictor of Fe bioavailability in human digestion—exposed to *in vitro* digests of NS and 35S-*OsGGP* BR and GBR samples.

2.2 Introduction

Producing sufficient and nutritious food to feed the growing world population represents one of the great challenges of the 21st century. Abiotic stresses such as drought, salinity, and extreme temperatures are a major limiting factor of global crop productivity, and are predicted to be exacerbated and co-occur more frequently due to the effects of climate change [177, 178]. Common to many abiotic stresses is the excess production of reactive oxygen species (ROS) due to the disruption of cellular homeostasis [22]. The development of crops with an enhanced capacity to mitigate the damaging effects caused by the excess production of ROS represents one solution to help offset the forecasted yield losses associated with climate change [179]. Micronutrient deficiencies—a form of undernutrition commonly referred to as hidden hunger—are a major threat to the health and development of populations worldwide, affecting more than 2 billion people [180]. Hidden hunger is most prominent in the developing world, often resulting from diets dependent micronutrient-poor staple crops [180]. The production of crops containing increased micronutrient density and/or bioavailability—a process known as biofortification—offers a cost-effective and sustainable intervention to reduce hidden hunger [181].

Ascorbate (also known as vitamin C) is a vital multifunctional molecule for both plants and mammals [5]. Ascorbate is a reducing agent capable of donating electrons, and primarily serves a role as a cellular antioxidant and enzymatic co-factor [5]. In plants, ascorbate performs crucial roles in photosynthetic functions and stress tolerance, largely due to its ability to combat excess ROS produced by normal or stressed cellular metabolism as a ROS scavenger or through the ascorbate-glutathione cycle—an indispensable antioxidant system for controlling cellular redox homeostasis [21]. Beyond its functions in photosynthesis and stress tolerance, ascorbate plays many other roles in plants, including the biosynthesis of ethylene and organic acids [33, 34], regenerating oxidized α -tocopherol [35], transporting iron (Fe) [7], and influencing cell division [36], cell expansion [37], and the onset of flowering and senescence [38]. In humans, ascorbate is an essential micronutrient well-known for its essential role in collagen biosynthesis and prevention of the disease scurvy [8]. Ascorbate also plays many other key roles in human physiology, including the cellular uptake of dietary Fe in the gut. Ascorbate is a powerful enhancer of Fe bioavailability—the proportion of Fe that is absorbed and is available for use or storage—and can overcome Fe absorption inhibitors such as phytate and polyphenolic compounds [44, 182-186]. In

addition to enhancing Fe absorption in the gut, ascorbate has also been reported to regulate cellular Fe uptake and metabolism [9, 47]. Ascorbate has many further positive effects on human health, for instance by reducing the risk of infections, cardiovascular disease, stroke, and cancer [40].

Several routes toward ascorbate biosynthesis have been proposed in plants, however, substantial genetic evidence supports the L-galactose pathway—which converts D-fructose-6-P to ascorbate via eight enzymatic steps—as the predominant pathway toward ascorbate biosynthesis in plants [12-16]. The *GPD-L-galactose phosphorylase* (*GGP*, also known as *VTC2/VTC5*) gene encodes the fifth enzymatic step of the L-galactose pathway catalyzing the conversion of GDP-L-galactose to L-galactose 1-P and represents the first committed step toward ascorbate biosynthesis. Relative to other genes from the L-galactose pathway, increased expression of the *GGP* gene consistently results in the largest increases in ascorbate concentrations in a wide range of species, providing strong genetic evidence for *GGP* as the rate-limiting enzymatic step of the L-galactose pathway. For example, increased expression of the *GGP* gene has increased ascorbate concentrations 2.9- to 4.1-fold in *Arabidopsis thaliana* [62, 71], 3.1-fold in potato (*Solanum tuberosum* L.) [73], 2.1-fold in strawberry (*Fragaria × ananassa*) [73], 2.0- to 6.2-fold in tomato (*Solanum lycopersicum* L.) [67, 73], 1.4-fold in tobacco (*Nicotiana tabacum* L.) [74], and 2.5- to 2.6-fold in rice (*Oryza sativa* L.) [64, 187]. Increased expression of the *GGP* gene has further been associated with enhanced tolerance to multiple abiotic stresses including salt, ozone, and cold stress [64, 74, 187]. Although many studies have increased ascorbate concentrations in model and crop species, none have investigated whether the elevated levels of ascorbate might improve Fe bioavailability in the human diet.

Rice is the most important source of calories for humans, providing nearly half the world's population with more with more than 20% of their daily caloric intake [169]. It is an important staple crop for many parts of the world, particularly in Eastern Asia, Southern Asia, and South-East Asia [169]. Abiotic stresses such as drought, flooding, and salinity are major limiting factors in global rice productivity [175], therefore developing strategies to increase ascorbate concentrations in rice may help to mitigate the yield losses associated with these abiotic stresses. Moreover, rice contains negligible concentrations of ascorbate and is a poor source of bioavailable Fe [52, 188], thus strategies to increase ascorbate levels in rice could improve dietary intakes of both ascorbate and bioavailable Fe in humans. To date, several studies have increased the

expression of ascorbate biosynthetic or recycling genes to increase foliar ascorbate concentrations in rice [64, 79, 138, 187]. Whether these strategies alter ascorbate concentrations in the rice grain and/or other tissues is unclear.

In this study, we aimed to increase ascorbate concentrations in rice via constitutive overexpression of the *OsGGP* coding sequence (CDS) and to determine the effect that elevated levels of ascorbate have on salt stress tolerance and Fe bioavailability. We found that constitutive overexpression of the *OsGGP* CDS in rice did not affect ascorbate concentrations in brown rice (BR), but significantly increased ascorbate concentrations in germinated brown rice (GBR) and tissues at the vegetative growth phase. At the reproductive growth phase, ascorbate concentrations were significantly reduced in homozygous 35S-*OsGGP* plants and were associated with significantly reduced transcript levels of the endogenous *OsGGP* gene. Automated imaging revealed that the 35S-*OsGGP* plants did not display increased salt tolerance at the vegetative growth phase, despite having elevated ascorbate concentrations. Finally, we found that ascorbate concentrations were positively correlated with ferritin concentrations in Caco-2 cells—an accurate predictor of Fe bioavailability in human digestion—exposed to *in vitro* digests of NS and 35S-*OsGGP* BR and GBR samples.

2.3. Results

2.3.1. Constitutive overexpression of the *OsGGP* CDS in rice increases ascorbate concentrations in GBR and in the vegetative growth phase, but reduces levels of ascorbate at the reproductive growth phase

Rice cv. Nipponbare transformants constitutively overexpressing the *OsGGP* CDS were generated through *Agrobacterium tumefaciens*-mediated transformation of a T-DNA containing the *OsGGP* CDS under transcriptional control of the dual CaMV 35S promoter (Figure 2.1a). A total of 27 independent, hemizygous 35S-*OsGGP* transformation events were regenerated from tissue culture and ascorbate concentrations were measured in the T₁ BR. Ascorbate concentrations were negligible in both wild-type (WT) and 35S-*OsGGP* BR but significantly increased up to 5.0-fold in 35S-*OsGGP* GBR relative to WT (Table S2.1). Two events (hereafter referred to as 35S-*OsGGP*-1 and 35S-*OsGGP*-2) with increased ascorbate concentrations in T₁ GBR and Mendelian segregation ratios of 3:1 for presence/absence of the T-DNA in T₁ seedling were advanced and homozygous and null segregant (NS) progeny were identified (Table S2.1). Ascorbate concentrations were negligible in homozygous NS and 35S-*OsGGP* BR and did not differ significantly (Figure 2.1b). Ascorbate concentrations were, however, significantly increased 8.7- and 5.1-fold in homozygous 35S-*OsGGP*-1 and 35S-*OsGGP*-2 GBR relative to NS-1 and NS-2, respectively (Figure 2.1b). At the vegetative growth phase, foliar ascorbate concentrations were significantly increased 1.8-fold in both homozygous 35S-*OsGGP*-1 and 35S-*OsGGP*-2 plants relative to NS-1 and NS-2, respectively (Figure 2.1c). Similarly, root ascorbate concentrations were significantly increased 4.2- and 4.5-fold in homozygous 35S-*OsGGP*-1 and 35S-*OsGGP*-2 plants relative to NS-1 and NS-2, respectively (Figure 2.1d). In contrast to the vegetative growth phase, ascorbate concentrations at the reproductive growth phase foliar were significantly reduced 2.0-fold in both homozygous 35S-*OsGGP*-1 and 35S-*OsGGP*-2 plants relative to NS-1 and NS-2, respectively (Figure 2.1e and Figure S2.1). The reduction in ascorbate concentrations in 35S-*OsGGP* plants relative to NS at the reproductive growth phase was only observed in homozygous 35S-*OsGGP* plants; hemizygous 35S-*OsGGP* plants had significantly increased ascorbate concentrations relative to NS at the reproductive growth phase (Figure S2.2). Transcript levels of the endogenous *OsGGP* gene were significantly reduced 2.0-fold in homozygous 35S-*OsGGP*-1 plants relative to NS-1 at the reproductive phase.

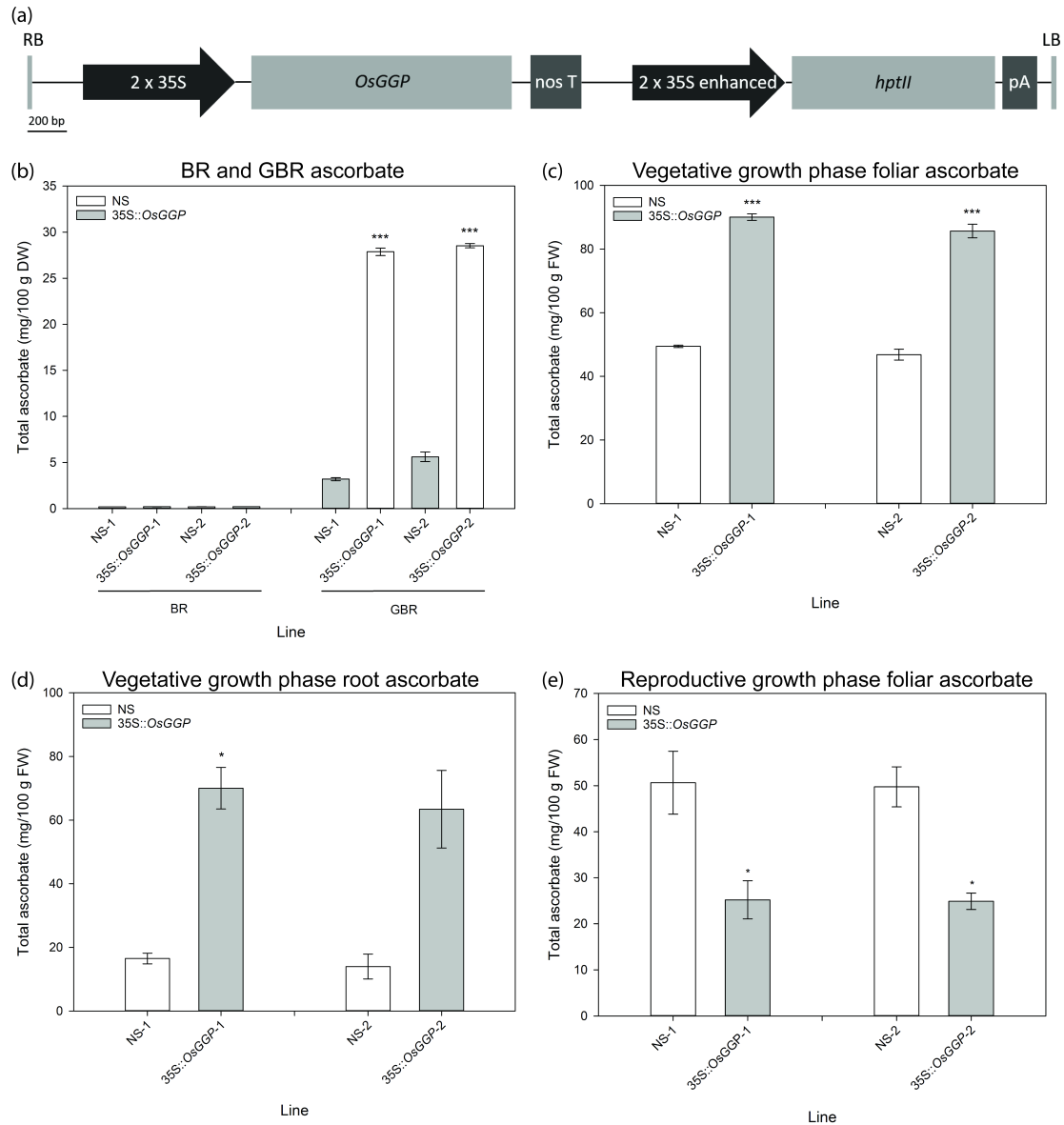


Figure 2.1. Production and characterization of independent rice transformation events constitutively overexpressing the *OsGGP* CDS. (a) Schematic representation of the T-DNA used for constitutive overexpression of the *OsGGP* CDS. RB, right border; 2 x 35S, dual CaMV 35S promoter; *OsGGP*, *OsGGP* CDS; nos T, nopaline synthase terminator; 2 x 35S enhanced, dual CaMV 35S promoter enhanced; *hptII*, hygromycin phosphotransferase II; pA, CaMV poly(A) signal; LB, left border. (b) Ascorbate concentrations of T₃ homozygous NS and 35S-*OsGGP* BR and GBR. Bars represent mean ± SEM of three independent replicates of approximately fifty grain. (c) Foliar and (d) root ascorbate concentrations of T₃ homozygous NS and 35S-*OsGGP* plants at the vegetative growth phase. (e) Foliar ascorbate concentrations of T₂ homozygous NS and 35S-*OsGGP* plants at the reproductive growth phase. Bars represent mean ± SEM of three biological replicates. Asterisks indicate statistically significant differences between NS and 35S-*OsGGP* plants (two-sample t-test; * p-value ≤ 0.05; *** p-value ≤ 0.001).

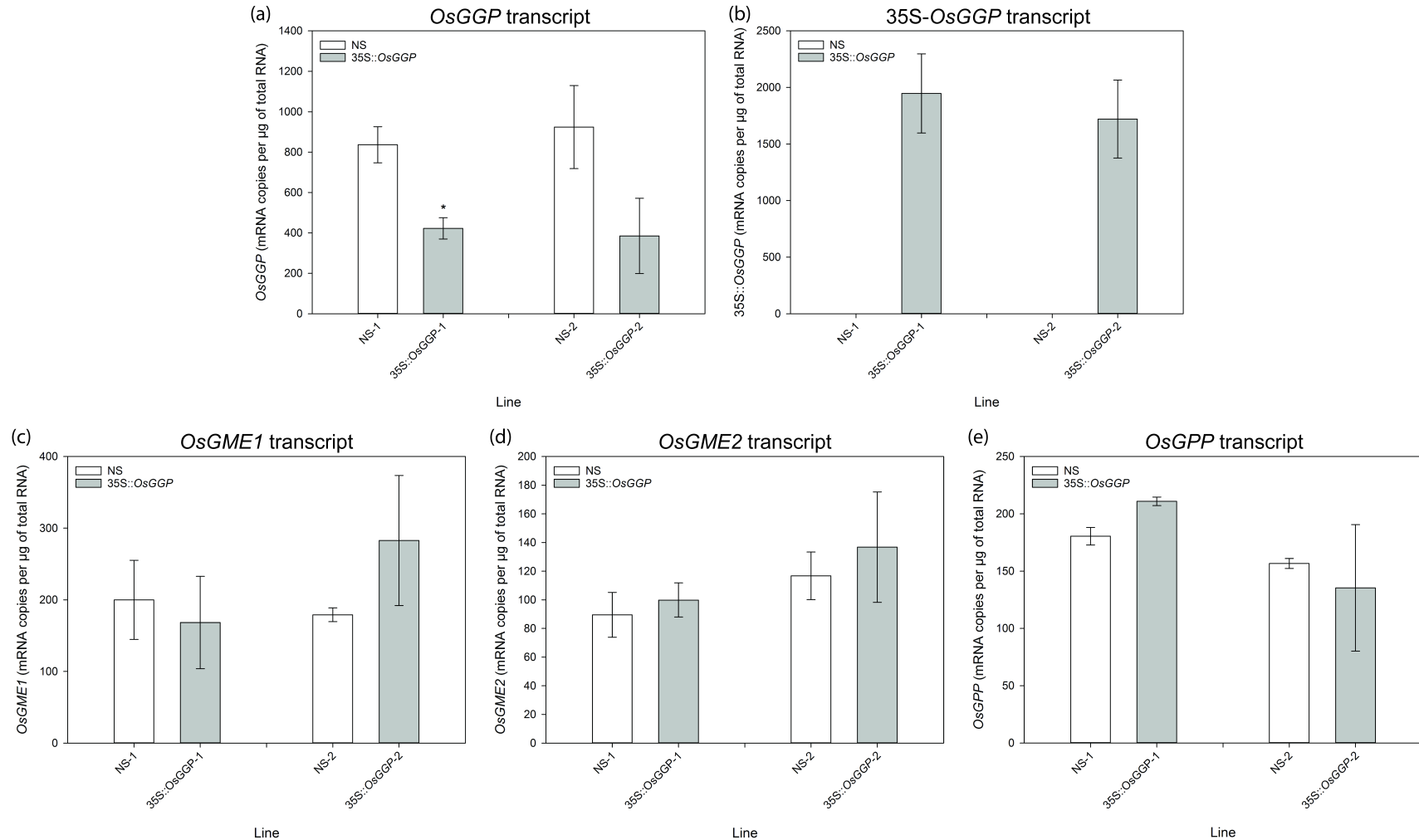


Figure 2.2. Transcript analysis of NS and 35S-*OsGGP* plants at the reproductive growth phase. Transcript levels of the (a) endogenous *OsGGP* gene, (b) 35S-*OsGGP* transgene, (c) *OsGME1* gene, (d) *OsGME2* gene, and (e) *OsGPP* gene of T₂ homozygous NS and 35S-*OsGGP* plants at the reproductive growth phase. Bars represent mean $\pm$ SEM of three biological replicates. Asterisks indicate statistically significant differences between NS and 35S-*OsGGP* plants (two-sample t-test; * p-value ≤ 0.05).

Similarly, transcript levels of the endogenous *OsGGP* gene were reduced 2.4-fold in homozygous 35S-*OsGGP*-2 plants relative to NS-2 but did not differ significantly (Figure 2.2a). Even though foliar ascorbate concentrations and endogenous *OsGGP* transcript levels were significantly reduced, high transcript levels of the 35S-*OsGGP* transgene were detected in the 35S-*OsGGP* plants (Figure 2.2b). Transcript levels of the *GDP-D-mannose-3',5'-epimerase 1* and 2 (*OsGME1* and *OsGME2*) genes, which encode the enzyme responsible for the fourth enzymatic step of the L-galactose pathway, and the *L-galactose-1-phosphate phosphatase* (*OsGPP*) gene, which encodes the sixth enzymatic step of the L-galactose pathway, did not differ significantly between the NS and 35S-*OsGGP* plants (Figure 2.2c-e).

2.3.2. Constitutive overexpression of the *OsGGP* CDS in rice did not affect salt tolerance at the vegetative growth phase in terms of shoot biomass accumulation

Shoot growth measurements of control and salt-stressed 35S-*OsGGP* plants during the vegetative growth phase were carried out using automated imaging. Salt stress was applied in two steps to the plants at days after planting (DAP) 20 and 23 and the plants were imaged daily from DAP 14 to 40 inclusive. Under control conditions, foliar ascorbate concentrations were significantly increased 1.2-fold in 35S-*OsGGP*-1 plants relative to NS-1 at DAP 40, however foliar ascorbate concentrations did not differ significantly between NS-2 and 35S-*OsGGP*-2 plants (Figure 2.3a). Under salt conditions, foliar ascorbate concentrations were significantly increased 1.3-fold in both 35S-*OsGGP*-1 and 35S-*OsGGP*-2 plants relative to NS-1 and NS-2, respectively, at DAP 40 (Figure 2.3b). From smoothed projected shoot area (sPSA) values—a strong predictor of shoot biomass—the sPSA absolute growth rate (AGR) and sPSA relative growth rate (RGR) were calculated at the following DAP intervals: 16-20, 20-24, 24-30, 30-35, and 35-40 (Figure 2.3c-f and Figure S2.3). As expected, the sPSA RGR declined more rapidly in salt-stressed plants than control plants (Figure 2.3c-f). However, a similar trend in the sPSA RGR was observed for both control and salt-stressed 35S-*OsGGP*-1 plants relative to NS-1, with the 35S-*OsGGP*-1 plants not differing significantly relative to NS-1 at any of the calculated DAP intervals (Figure 2.3c,d). Likewise, a similar trend in the sPSA RGR was observed for both control and salt-stressed 35S-*OsGGP*-2 plants relative to NS-2, with the 35S-*OsGGP*-2 plants having a significantly lower sPSA RGR at the DAP interval 30-35 relative to NS-2, but

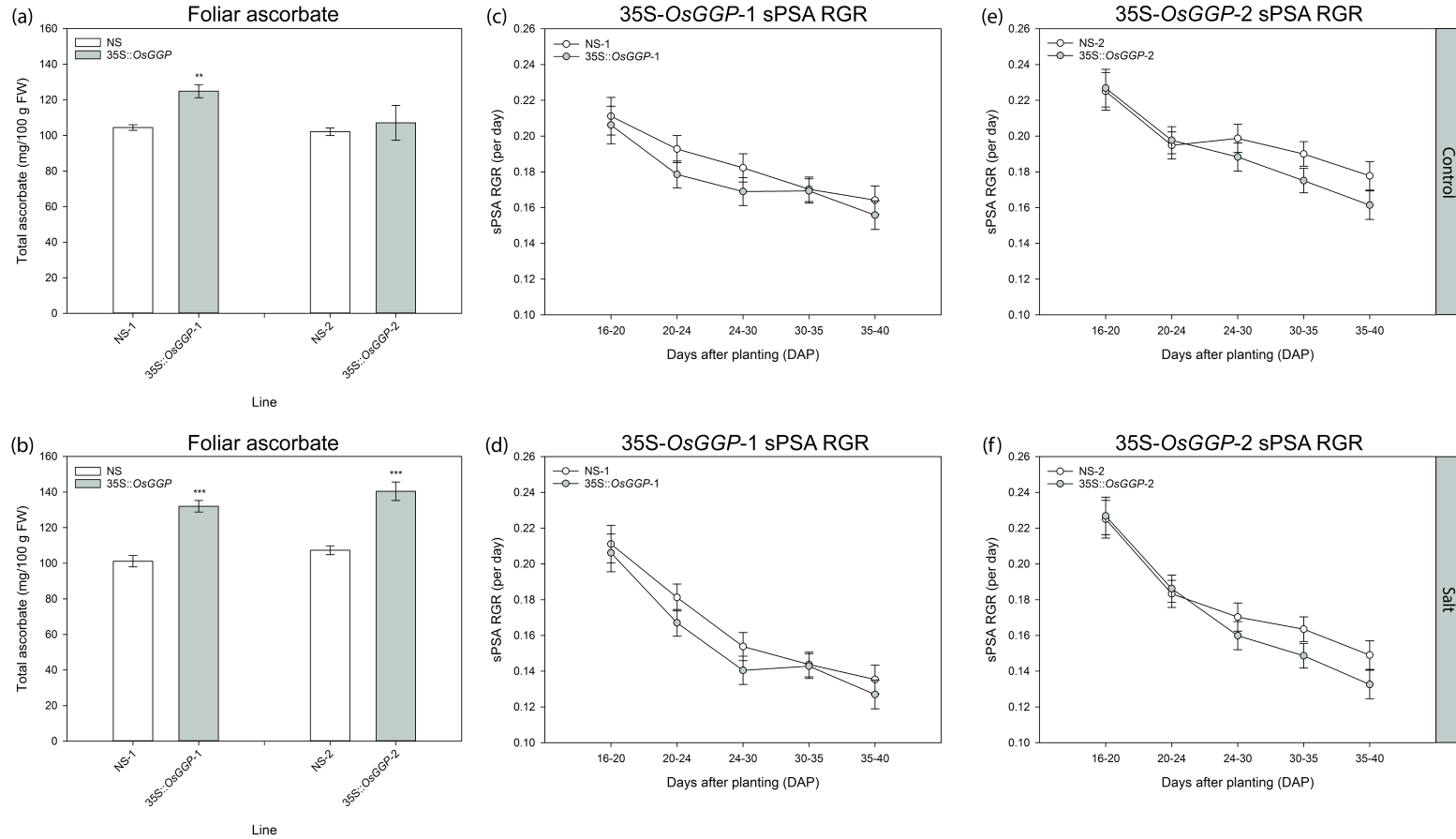


Figure 2.3. Salt tolerance assessment of NS and 35S-*OsGGP* plants during the vegetative growth phase. Foliar ascorbate concentrations of (a) control and (b) salt-stressed T₄ homozygous NS-1 and 35S-*OsGGP*-1 plants and T₃ homozygous NS-2 and 35S-*OsGGP*-2 plants at DAP 40. Bars represent mean $\pm$ SEM of six biological replicates. Asterisks indicate statistically significant differences between NS and 35S-*OsGGP* plants (two-sample t-test; ** p-value ≤ 0.01 ; *** p-value ≤ 0.001). The sPSA RGR of (c) control and (d) salt-stressed T₄ homozygous NS-1 and 35S-*OsGGP*-1 plants and of (e) control and (f) salt-stressed T₃ homozygous NS-2 and 35S-*OsGGP*-2 plants. Salt was applied at DAP 20 and 23. Values represent mean $\pm$ half least significant (5%) pairwise difference of six biological replicates. Non-overlapping error bars indicate significant differences at $\alpha = 0.05$.

otherwise did not differ significantly for any of the other calculated DAP intervals (Figure 2.3e,f).

2.2.3. Ascorbate concentrations were positively correlated with ferritin concentrations in Caco-2 cells exposed to *in vitro* digests of NS and 35S-*OsGGP* BR and GBR samples

The bioavailability of Fe in NS-1 and 35S-*OsGGP*-1 BR and GBR was determined using the coupled *in vitro* digestion/Caco-2 cell culture assay. Ascorbate concentrations were negligible in NS-1 and 35S-*OsGGP*-1 BR and did not differ significantly but were significantly increased 6.6-fold in 35S-*OsGGP*-1 GBR relative to NS-1 (Figure 2.4a). The concentrations of ascorbate in the NS-1 and 35S-*OsGGP* BR and GBR samples were strongly associated with transcript levels of the endogenous *OsGGP* gene and 35S-*OsGGP* transgene (Figure S2.4). Phytate and Fe concentrations did not differ significantly between NS-1 and 35S-*OsGGP*-1 BR or GBR (Figure 2.4b,c). The NS-1 and 35S-*OsGGP*-1 BR had molar ratios of 1:31.5:0.0 and 1:38.0:0.0, respectively, of Fe to phytate to ascorbate. The NS-1 and 35S-*OsGGP*-1 GBR had molar ratios of 1:25.3:0.4 and 1:30.3:2.5, respectively, of Fe to phytate to ascorbate. Ferritin concentrations of Caco-2 cells exposed to *in vitro digests* of NS-1 and 35S-*OsGGP*-1 BR were similar and did not differ significantly (Figure 2.4d). Ferritin concentrations of Caco-2 cells exposed to *in vitro digests* of 35S-*OsGGP*-1 GBR were increased 1.8-fold relative to NS-1 but did not differ significantly (Figure 2.4d). Ascorbate concentrations were positively correlated with ferritin concentrations in Caco-2 cells ($r = 0.620$), whereas Fe and phytate concentrations were not ($r = 0.210$ and $r = 0.042$, respectively) (Table 2.1).

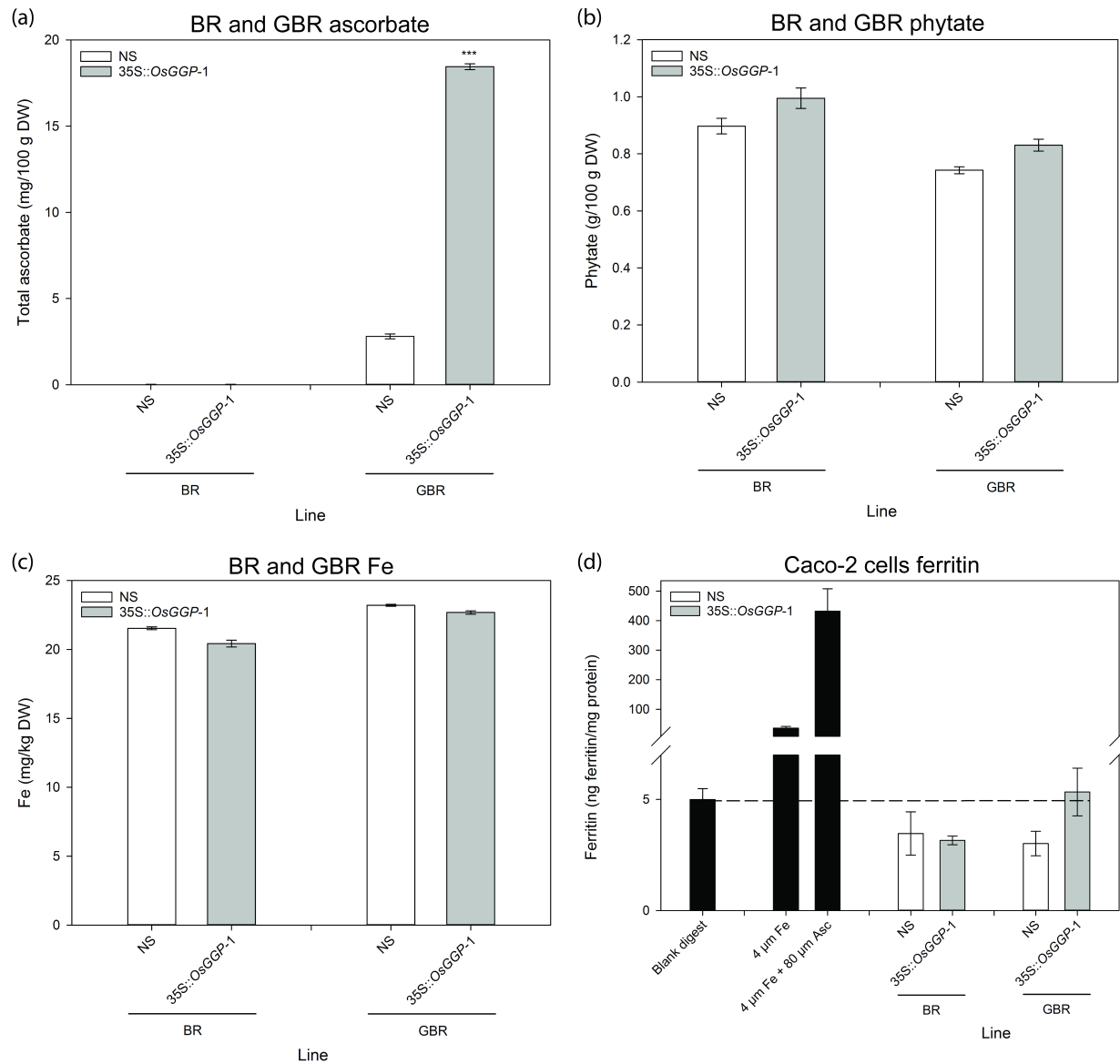


Figure 2.4. Fe bioavailability assessment of NS and 35S-*OsGGP* BR and GBR. Concentrations of (a) ascorbate, (b) phytate, and (c) Fe in T₃ homozygous NS-1 and 35S-*OsGGP*-1 BR and GBR. Bars represent mean $\pm$ SEM of three replicates from a bulked pool of lyophilized grain. (d) Ferritin concentrations of Caco-2 cells exposed to *in vitro* digests of T₃ homozygous NS-1 and 35S-*OsGGP*-1 BR and GBR. A solution of 4 μ m Fe and a solution of 4 μ m Fe + 80 μ m ascorbate (Asc) were used as positive controls to verify the responsiveness of the Caco-2 cells to Fe uptake. Bars represent $\pm$ SEM of three biological replicates. Asterisks indicate significant differences between the between the NS-1 and 35S-*OsGGP*-1 plants (two-sample t-test, *** p-value ≤ 0.001).

Table 2.1. Correlation analysis of ascorbate, phytate, and Fe concentrations in T₃ homozygous NS-1 and 35S-*OsGGP*-1 BR and GBR with ferritin concentrations in Caco-2 cells exposed to *in vitro* digests of NS-1 and 35S-*OsGGP*-1 BR and GBR samples.

	Fe (µg/g)	Phytate (g/100 g)	Ferritin (ng/mg protein)
Ascorbate (mg/100 g)	0.494	-0.320	0.620*
Fe (µg/g)		-0.935***	0.210
Phytate (g/100 g)			0.042

Values represent correlation coefficients of twelve replicates. Asterisks indicate significant correlations (Pearson correlation method, * p -value ≤ 0.05 ; *** p -value ≤ 0.001).

2.4. Discussion

Constitutive overexpression of the *OsGGP* CDS did not affect ascorbate concentrations in BR, a result that is consistent with the reported decline of ascorbate to negligible levels during grain maturation in cereals [54, 55]. Overcoming the decline in ascorbate concentrations during grain maturation represents a significant challenge to ascorbate biofortification of cereal species such as rice. Further understanding this decline in ascorbate concentrations in the maturing grains of cereals may help identify strategies to reduce this decline, and therefore present avenues to biofortify cereals with ascorbate, and ought to be the subject of future research. Alternatively, future research in the field of ascorbate biofortification may benefit from focusing on crops that contain meaningful concentrations of ascorbate at the time of harvest, such as fruits and vegetables. Constitutive overexpression of the *OsGGP* CDS did, however, significantly increase ascorbate concentrations in GBR [64]. To our knowledge, the 8.7-fold increase in ascorbate concentrations in 35S-*OsGGP*-1 GBR, relative to NS-1, represents the largest fold increase in ascorbate reported in any plant tissue to date. It is possible that GBR, an increasingly popular functional food [189, 190], could serve as an ascorbate biofortified cereal product, but further studies are required to determine the effect of germination conditions, processing, and cooking on the final concentrations of ascorbate. Constitutive overexpression of the *OsGGP* CDS also significantly increased ascorbate concentrations in leaves and roots at the vegetative growth phase. The 1.8-fold change in foliar ascorbate concentrations is similar to that reported in previous studies with increased expression of the *Arabidopsis AtGGP* and kiwifruit (*Actinidia chinensis* L.) *AcGGP* gene in rice [64, 187]. To our knowledge, the 4.5-fold increase in ascorbate concentrations in the 35S-*OsGGP*-1 roots, relative to NS-1, represents the largest fold increase reported in plant roots to date and indicates that GGP is likely also a significant rate-limiting enzymatic step toward ascorbate biosynthesis in roots. Contrary to the results in GBR and tissues at the vegetative growth phase, constitutive overexpression of the *OsGGP* CDS significantly reduced foliar ascorbate concentrations at the reproductive growth phase. This reduction in ascorbate concentrations was dependent on homozygosity of the 35S-*OsGGP* transgene and was associated with reduced transcript levels of the endogenous *OsGGP* gene. These results suggest that post-transcriptional silencing of the endogenous *OsGGP* gene and 35S-*OsGGP* transgene may be occurring in response to elevated transcript levels of *OsGGP* in a homozygous-dependent manner [191].

Detection of high transcript levels of the 35S-*OsGGP* transgene does not necessarily fit the traditional model of post-transcriptional gene silencing, however similar observations were made when the kiwifruit *AcGGP* gene was expressed in *Arabidopsis*, and was interpreted as incomplete processing of the silenced gene transcripts thereby leaving mRNA template for detection by quantitative reverse transcription-PCR [71]. The identification of *OsGGP* derived small interfering RNAs, for example, may indicate the occurrence of gene silencing in the 35S-*OsGGP* plants and warrants further investigation. Future efforts to increase ascorbate concentrations in plants could avoid the use of transgene overexpression by utilizing genome editing tools such as the CRISPR/Cas9 system to disrupt transcriptional or translational repressors of ascorbate biosynthetic genes, such as the highly conserved *cis*-acting upstream open reading frame that controls translation of the *GGP* gene [17, 153, 154, 163, 164, 192, 193].

We also examined whether the 35S-*OsGGP* plants displayed altered salt tolerance at the vegetative growth phase using automated imaging. Previous studies have demonstrated that non-destructive measurements such as sPSA is a strong predictor of shoot biomass, and that the sPSA RGR can be used effectively to screen for salt tolerance in rice [194-197]. However, a similar trend in the sPSA RGR was observed for both control and salt-stressed 35S-*OsGGP* plants relative to NS at all calculated DAP intervals, suggesting that the elevated ascorbate concentrations from constitutive overexpression of the *OsGGP* CDS in rice does not affect salt tolerance in terms of shoot biomass accumulation during this early growth phase in response to a final concentration of 90 mM NaCl. This result is incongruent with many previous studies that have reported enhanced salt tolerance in model and crop species with increased expression of ascorbate biosynthetic genes [64, 68, 70, 80, 89-93, 99, 118, 187, 198]. Particularly, our results differ with a previous study that found increased expression of the *Arabidopsis AtGGP* gene in rice to enhance seedling salt tolerance in terms of biomass and RGRs [64]. Our results may be more consistent with a recent study that reported that increased expression of the kiwifruit *AcGGP* gene in rice enhanced salt tolerance in terms of visible symptoms and lipid peroxidation, but not shoot biomass [187]. Investigating oxidative stress markers such as lipid peroxidation in the salt-stressed 35S-*OsGGP* plants may determine whether constitutive overexpression of the *OsGGP* CDS in rice ameliorates salt-induced oxidative stress and will be the subject of future research. One limitation of this study was the possibility that gene silencing was being induced in the 35S-*OsGGP* plants and could therefore have confounded the interpretation of the results. Nevertheless, the 35S-*OsGGP* plants

had a 1.3-fold increase in foliar ascorbate concentrations at DAP 40 under salt stress conditions and thus an enhanced antioxidant capacity to combat any salt-induced oxidative stress.

The ascorbate-enriched 35S-*OsGGP* GBR that we generated in this study was used to test whether engineering ascorbate-enriched crops could improve Fe bioavailability in human diets utilizing the Caco-2 cell line model—an accurate predictor of human Fe absorption [199]. Previous studies have reported that phytate significantly reduces Fe bioavailability, beginning at a 1:1 molar ratio of Fe to phytate, with maximal inhibition occurring at a ratio around 1:10 [44, 184, 200]. Ferritin concentrations below or equivalent to the blank digest for the Caco-2 cells exposed to *in vitro* digests of NS-1 and 35S-*OsGGP* BR and GBR in this study can therefore likely be explained by the high levels of phytate [44, 200]. Ferritin concentrations of Caco-2 cells exposed to *in vitro* digests of the ascorbate-enriched 35S-*OsGGP* GBR were increased 1.8-fold relative to NS-1, albeit non-significantly. This result indicates that the increased ascorbate levels in the 35S-*OsGGP* GBR may have had a slightly promotive effect on Fe bioavailability, but an even greater increase in ascorbate levels to counteract the high levels of phytate would likely be required to reveal a significant effect. Ascorbate concentrations—but not Fe or phytate concentrations—positively correlated with Caco-2 cell ferritin formation in our study. This is consistent with an *in vitro* bioavailability study of Andean potato clones with varying Fe, ascorbate, and phenolic concentrations, that reported that ascorbate concentrations were the strongest predictor of ferritin formation [201]. One limitation of our study is that we did not investigate the concentrations of phenolic compounds—that are known to inhibit Fe absorption—in the BR and GBR [44, 184, 200]. Total phenolic compounds have been reported to increase in germinating BR, and thus it is possible that increased phenolic compounds may have reduced the promoting effects of ascorbate on Fe bioavailability in the 35S-*OsGGP* GBR [202]. Taken together, our Caco-2 Fe bioavailability results of the ascorbate-enriched 35S-*OsGGP* indicate that increasing ascorbate concentrations in crops may be able to improve Fe bioavailability in human diets, but further *in vivo* assessments of ascorbate-enriched crops on Fe bioavailability are needed to support this.

2.5. Materials and Methods

2.5.1. Plant growth conditions

Oryza sativa cv. Nipponbare was used for all experiments. Rice grain were surface sterilized with 70% (v/v) ethanol (Chem-Supply, SA, Australia) for 1 min and 30% (v/v) bleach (White King, NSW, Australia) with a few drops of Tween-20 (Sigma-Aldrich, MO, USA) for 30 min and washed three times with sterile dH₂O. Surface sterilized rice grain were germinated in a petri dish with moist sterile filter paper (Whatman, UK) for 7 to 9 days prior to transplanting to 1 L pots filled with potting mix in a glasshouse maintained at 26°C and 70% relative humidity at The University of Melbourne (Melbourne, VIC, Australia). The potting mix was prepared by mixing one part washed fine sand (Col Smith, VIC, Australia), one part propagating sand (Brunnings, VIC, Australia), two parts premium vermiculate (Exfoliators, VIC, Australia), and one part General Mix potting media (Australian Growing Solutions, VIC, Australia) fertilized with Osmocote Exact Standard 8-9 M (ICL, NSW, Australia) at a rate of 6 g/L.

2.5.2. Vector construction and rice transformation

The *OsGGP* (LOC_Os12g08810) CDS was PCR-amplified from rice cv. Nipponbare cDNA and recombined into the Gateway-compatible pMDC32 vector [203], which placed the *OsGGP* CDS under transcriptional control of a dual CaMV 35S promoter and also contained the *hygromycin phosphotransferase II* (*hptII*) plant-selectable marker gene (Table S2.2). *Agrobacterium*-mediated transformation of rice callus was carried out using established protocols [204].

2.5.3. Genotyping

Genomic DNA was extracted using the Extract-N-Amp™ Plant PCR Kit (Sigma-Aldrich) according to manufacturer's instructions. Presence of the 35S-*OsGGP* T-DNA was determined via multiplex PCR amplification of *hptII* with *OsACT1*. The PCR amplification cycles consisted of 1 cycle = 1 min 95°C; 40 cycles = 20 s 95°C, 20 s 60°C, 90 s 72°C. The PCRs were performed in a final volume of 20 µL for MyTaq™ HS Red DNA Polymerase (Bioline, UK) according to manufacturer's instructions.

2.5.4. Quantification of total ascorbate

Total ascorbate (ascorbate + dehydroascorbate) was extracted and measured as previously described with modifications [205]. Briefly, total ascorbate was extracted from ground, lyophilized tissue homogenised in extraction fluid containing 8% metaphosphoric acid, 2 mM EDTA, and 2 mM TCEP (Sigma-Aldrich) at 40°C for 2 h. The extract was centrifuged and 5 µL of supernatant was injected onto a C18 3 µm 33 x 7 mm Alltima Rocket column (Hichrom Limited, UK) maintained at 40°C with a flow rate of 1 mL/min. The concentration was determined by reverse phase chromatography on a Shimadzu Nexera UHPLC system (Shimadzu, Japan) with simultaneous UV and MS detection. Mobile phase A consisted of MS grade water with 0.1% formic acid (ThermoFisher Scientific, MA, USA). The elution procedure utilized a 1 min gradient from 0% to 90% of mobile phase B: MS grade acetonitrile with 0.1% formic acid (ThermoFisher Scientific). The UV absorption signal was acquired using a Shimadzu SPD-20A detector at 245 nm wavelength. The MS data was acquired on a Shimadzu LCMS-8050 triple quadrupole mass spectrometer equipped with a DUIS source. The MS instrument was operated in Multiple Reaction Monitoring mode monitoring ions in both positive (177>95, 177>85) and negative (175 >115, 175>71, 175>59) ionisation modes. The product ion 115 *m/z* was used as a quantifier. A calibration curve in the range of 1.95 to 250 ppm of L-ascorbic acid analytical standard (Sigma-Aldrich) was used to determine absolute concentration of analyte in the extracts.

2.5.5. Quantitative reverse transcription-PCR analysis

Total RNA from foliar tissue was isolated using the Direct-zol™ RNA MiniPrep Kit (Zymo Research, CA, USA) according to manufacturer's instructions. To isolate total RNA from BR and GBR, sub-samples of snap-frozen BR and GBR were ground using a Tube Mill with a chilled 40 mL grinding chamber (IKA, Germany) and suspended in an extraction buffer containing 0.1M glycine-NaOH (Chem-Supply), 100 mM NaCl (Chem-Supply), 10 mM EDTA (Chem-Supply), 2% SDS (Sigma-Aldrich), and 1% sodium lauryl-sarcosine (Sigma-Aldrich) at pH 9 for 7 min at 1,400 RPM [206]. The solution was then mixed with an equal volume of phenol to chloroform to isoamyl alcohol at a ratio of 25:24:1 (Sigma-Aldrich) for 1 min at 1,400 RPM followed by incubation on ice for 5 min. After centrifugation for 20 min at 4°C, total RNA was then purified from the upper aqueous phase using the Direct-zol™ RNA MiniPrep Kit (Zymo Research) according to manufacturer's instructions. Quantification, DNase treatment, and reverse

transcription of the RNA was performed using the QuantiFluor® RNA System (Promega, WI, USA), RQ1 RNase-Free DNase (Promega), and the Tetro cDNA Synthesis Kit (Bioline), respectively, according to manufacturer's instructions. The quantitative reverse transcription-PCR (qRT-PCR) primers were designed to amplify the *OsGPP*, *35S-OsGPP*, *OsGME1*, *OsGME2*, and *OsGPP* genes using Primer 3 software [207, 208] (Table S2.2). Primer efficiency was $\geq 93\%$ for all primer pairs. The qRT-PCR analysis was carried out on a CFX Connect™ Real-Time System (Bio-Rad, CA, USA) with Hard-Shell® 96-Well PCR Plates (Bio-Rad). The qRT-PCR amplification cycles consisted of 1 cycle = 3 min 95°C; 40 cycles = 20 s 95°C, 20 s 60°C, 20 s 72°C and were followed by a melt curve from 65°C to 95°C in 0.5°C steps for 5 s/step. The qRT-PCRs were performed in a final volume of 10 µL for KAPA SYBR® FAST (Kapa Biosystems, MA, USA) according to manufacturer's instructions, each with four technical replicates. The absolute quantification of transcript copy number was determined using a 10-fold serial dilution of 10^8 – 10^2 copies for each PCR product and a standard curve generated with the Bio-Rad CFX Manager 3.1 software (Bio-Rad). The PCR products were purified and quantified with a DNA Clean & Concentrator™-5 Kit (Zymo Research) and the QuantiFluor® ONE dsDNA System (Promega), respectively, prior to serial dilution. The geometric mean expression of three optimal housekeeping genes selected from *OsGAPDH*, *OsELF1*, *OsACT1* and *OsUBQ5* using the NormFinder Excel add-in [209] was used to normalize the expression of the *35S-OsGPP*, *OsGPP*, *OsGME1*, *OsGME2*, and *OsGPP* genes.

2.5.6 Assessment of salt tolerance with automated imaging

Surface sterilized rice grain were germinated in plastic boxes on moist paper towels (January 9, 2019, denoted DAP 0). At DAP 6, three uniformly germinated grain per line were transplanted to 2.8 L pots filled with 2.6 kg of UC Davis Mix, placed into square containers for water to collect, and grown on benches in the NW Smarthouse maintained at 26°C and 70% relative humidity at The Plant Accelerator, Adelaide, Australia under manual watering (not to weight). At DAP 14, plants were thinned to one seedling per pot and loaded onto the NW Smarthouse for automated imaging. A split-unit design was used to randomize pots (i.e. line and treatment combinations) to cart positions within the Smarthouse. The plants were automatically watered daily so that approximately 500 mL of water was maintained in the soil. The salt plants were treated with saline solution on DAP 20 and 23 (100 mL of 225mM NaCl) to reach a final concentration of 90 mM

NaCl in the soil solution, while the control plants received equivalent water applications as previously described [196]. Imaging was carried out daily from DAP 14 to 40 inclusive. From these images the PSA of the plant, as viewed using an RGB camera, was obtained as previously described [210]. For analysis purposes, PSA was defined as the sum of the areas as measured (in kilopixels) from six side/oblique camera views plus twice the shoot area (top view). The data was prepared for analysis using growthPheno [211], a package for the R statistical computing environment [212].

2.5.7. Preparation of GBR

Following surface sterilization, BR samples were immediately snap-frozen in liquid nitrogen, lyophilized, and ground using a Tube Mill with a 40 mL grinding chamber (IKA), whereas GBR samples were imbibed in sterile dH₂O for 12 h in the dark at 28°C, washed three times with sterile dH₂O, and spread in a petri dish with moist sterile filter paper (Whatman) for 72 h in the dark at 28°C. The GBR was then washed three times with sterile dH₂O, snap-frozen in liquid nitrogen, lyophilized, and ground using a Tube Mill with a 40 mL grinding chamber (IKA).

2.5.8. Quantification of Fe

Inductively coupled plasma optical emission spectrometry (ICP-OES) of the ground, lyophilized BR and GBR samples was conducted at the Robert W. Holley Centre for Agriculture and Health (USDA-ARS, Ithaca, NY) to determine Fe concentrations as previously described [213].

2.5.9. Quantification of phytate

Phytate concentrations of the ground, lyophilized BR and GBR samples were measured as the proportion of phosphorus released by phytase and alkaline phosphatase using the Phytic Acid/Total Phosphorus Assay Kit (Megazyme International, Ireland). Phosphorus concentrations in free and total phosphorus solutions were determined by absorbance at 655 nm.

2.5.10 Caco-2 cell culture assessment of Fe bioavailability

Ground, lyophilized BR and GBR samples were digested for Caco-2 cell Fe bioavailability analysis as previously described [199, 214, 215]. Briefly, the Caco-2 cells were maintained at 37°C in supplemented Dulbecco's modified Eagle medium (ThermoFisher Scientific) for 11 d post-

seeding and replaced with supplemented minimum essential media (MEM) solution (ThermoFisher Scientific) 48 h prior to the experiment day. On the experiment day, gastric-digested samples were added to cylindrical Transwell inserts (Corning Life Sciences, NY, USA) fitted with a Spectra/Por 2.1 15,000 Da molecular weight cut-off dialysis membrane (Spectrum Medical, CA, USA). The inserts were placed within wells containing Caco-2 cell monolayers and incubated for 2 h at 37°C, after which the inserts were removed, and additional MEM was added to the cells before incubation for 22 h at 37°C. After incubation, the growth medium was removed by aspiration and the Caco-2 cells were washed twice with a rinse solution containing 140 mmol/L NaCl, 5 mmol/L KCl, and 10 mmol/L PIPES (Sigma-Aldrich) at pH 7 and harvested with the addition of dH₂O followed by brief sonication. From an aliquot of the Caco-2 cell solution, ferritin and protein content was determined using a FER-IRON II Ferritin Assay (Ramco Laboratories, TX, USA) and DCTM Protein Assay (Bio-Rad), respectively, according to manufacturer's instructions. As Caco-2 cells synthesize ferritin in response to intracellular Fe, the ratio of ferritin/total protein (expressed as ng ferritin/mg protein) was used as an index of cellular Fe uptake. A solution of 4 µm Fe (High Purity Standards, SC, USA) and a solution of 4 µm Fe (High Purity Standards) and 80 µm ascorbic acid (Sigma-Aldrich) were used as positive controls to verify the responsiveness of the Caco-2 cells to Fe uptake.

2.5.11. Statistical analysis

Statistically significant differences between the NS and 35S-*OsGGP* plants were detected with the two-sample t-test using the 5% level of significance and calculated with Minitab 17.1.0 (<http://www.minitab.com/en-us/>). The significance of the correlations between ascorbate, phytate, and Fe concentrations in NS and 35S-*OsGGP* BR and GBR with ferritin concentrations in Caco-2 cells were examined with the Pearson correlation method at the 5% significance level using Minitab 17.1.0. A mixed model analysis was performed on the data for each trait from the imaged plants using the R packages ASReml-R [216] and asremlPlus [217]. Estimated means were obtained from the analysis and the significances of their differences were investigated using the Least Significant Difference at the 5% significance level.

2.6. Supplementary Information

Table S2.1. Ascorbate concentrations of WT and T₁ 35S-*OsGGP* GBR and the number of T₁ 35S-*OsGGP* seedlings with the 35S-*OsGGP* T-DNA present or absent.

Genotype	Total ascorbate (mg/100 g DW) ^a	No. of seedlings with T-DNA	No. of seedlings without T-DNA	Chi-square p-value ^b
WT	1.66	-	-	-
Event 1	4.35	8	2	0.715
Event 2	2.88	10	2	0.505
Event 4	2.37	10	1	0.223
Event 5	2.94	11	1	0.182
Event 6	3.07	9	2	0.602
Event 7	5.64	10	2	0.505
Event 8	4.44	8	4	0.505
Event 10	3.82	11	1	0.182
Event 13	4.10	11	1	0.182
Event 16	2.31	11	1	0.182
Event 19	2.62	11	1	0.182
Event 20 (35S- <i>OsGGP</i> -2)	5.19	8	2	0.715
Event 22	2.17	9	1	0.273
Event 24	5.36	9	1	0.273
Event 25	6.29	10	1	0.223
Event 26	5.18	8	2	0.715
Event 27	6.26	10	1	0.223
Event 30	1.84	10	1	0.223
Event 33	0.95	8	0	0.103
Event 41	2.26	11	1	0.182
Event 42	4.07	7	1	0.414
Event 43	3.15	8	3	0.862
Event 45	2.52	10	1	0.223
Event 46	6.11	12	0	0.046*
Event 49	4.22	10	2	0.505
Event 50	3.74	11	1	0.182
Event 51 (35S- <i>OsGGP</i> -1)	8.30	10	2	0.505

^aTotal ascorbate values represent one replicate of four to eleven grain. ^bAsterisks indicate statistically significant differences between observed and expected segregation ratio of 3:1 for the presence to absence of the 35S-*OsGGP* T-DNA (chi-square test; * p-value ≤0.05).

Table S2.2. List of PCR and qRT-PCR primers used in this study. The table provides primer name, forward and reverse primer sequences, PCR product length, and the application of the primers.

Primer name	Forward primer sequence (5' - 3')	Reverse primer sequence (5' - 3')	Product length (bp)	Application
<i>OsGGP</i>	ATGGAGATGAAGCTGACCATCAAG	CTGCCTCGTCCTTCAGTGA	1317	Cloning the <i>OsGGP</i> CDS
<i>OsACT1</i>	TCAGCAACTGGGATGATATGGAG	GCCGTTGTGGTGAATGAGTAAC	634	Detecting the 35S- <i>OsGGP</i> T-DNA
<i>hptII</i>	CTTGTATGGAGCAGCAGACGC	CTATTTCTTTGCCCTCGGACG	308	Detecting the 35S- <i>OsGGP</i> T-DNA
qRT-PCR-35S- <i>OsGGP</i>	AGGAAGTCAGCGAAGACGAG	AAGACCGGCAACAGGATTC	249	Transcript analysis of the 35S- <i>OsGGP</i> transgene
qRT-PCR- <i>OsGGP</i>	AGATCTTCCTCTTCCCTCAGT	CCGACAACCTGTTCTACCGT	368	Transcript analysis of the endogenous <i>OsGGP</i> gene
qRT-PCR- <i>OsGME1</i>	CGTATCAATGGTGTGAAGAGGTT	GGCCAGGCATCAGATTCCTT	107	Transcript analysis of the <i>OsGME1</i> gene
qRT-PCR- <i>OsGME2</i>	GATATGGGAGGGATGGGATT	GCCAAGCATCTGACTCCTTC	196	Transcript analysis of the <i>OsGME2</i> gene
qRT-PCR- <i>OsGPP</i>	TGCCTTCCTCAATGGTTCTC	CACACATGTTCAAGGCCAAC	188	Transcript analysis of the <i>OsGPP</i> gene
qRT-PCR- <i>OsGAPDH</i>	GGGCTGCTAGCTTCAACATC	TTGATTGCAGCCTTGATCTG	190	qRT-PCR housekeeping gene
qRT-PCR- <i>OsELF1</i>	GAAAGCAGCAGAAGAACGGG	TCCTCAAGCTTTGCCATGTCT	126	qRT-PCR housekeeping gene
qRT-PCR- <i>OsACT1</i>	ATCACTGCCTTGGCTCCTAG	TCGTACTCAGCCTTGCAA	140	qRT-PCR housekeeping gene
qRT-PCR- <i>OsUBQ5</i>	TACACCAAGCCCAAGAAGCAG	GTGAGGCCGCACTTACCG	191	qRT-PCR housekeeping gene

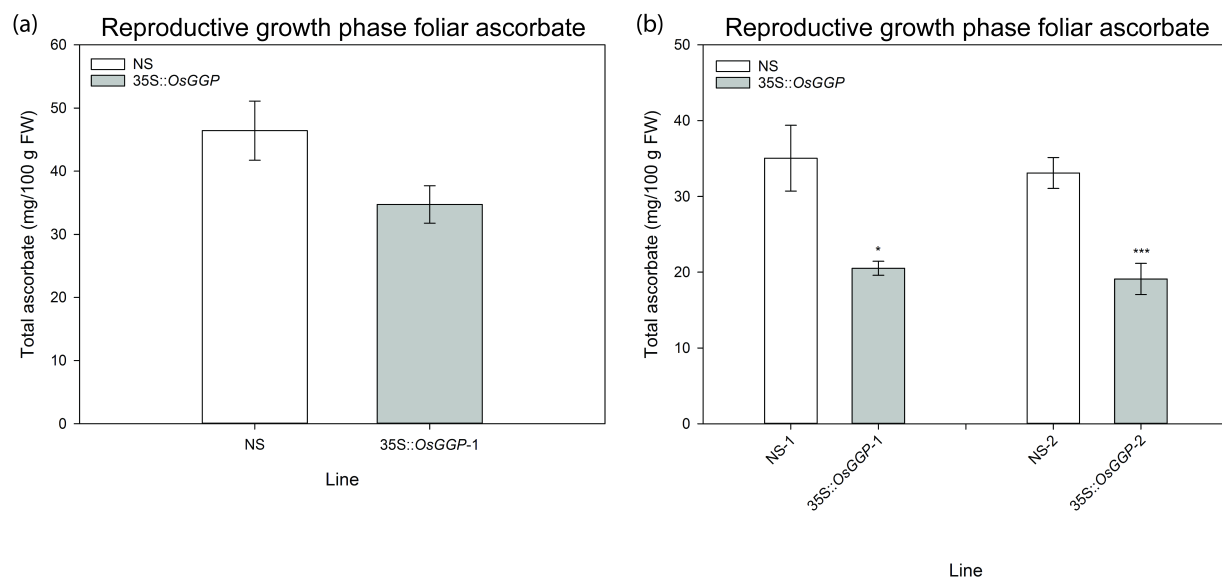


Figure S2.1. Foliar ascorbate concentrations of T₃ and T₄ NS and 35S-*OsGGP* plants at the reproductive growth phase. **(a)** Foliar ascorbate concentrations in T₃ homozygous 35S-*OsGGP*-1 plants at the reproductive growth phase. Bars represent mean $\pm$ SEM of three biological replicates. **(b)** Foliar ascorbate concentrations of T₄ homozygous NS-1 and 35S-*OsGGP*-1 and T₃ NS-2 and 35S-*OsGGP*-2 plants at the reproductive growth phase. Bars represent mean $\pm$ SEM of five to six biological replicates. Asterisks indicate statistically significant differences between NS and 35S-*OsGGP* plants (two-sample t-test; * p-value ≤ 0.05 ; *** p-value ≤ 0.001).

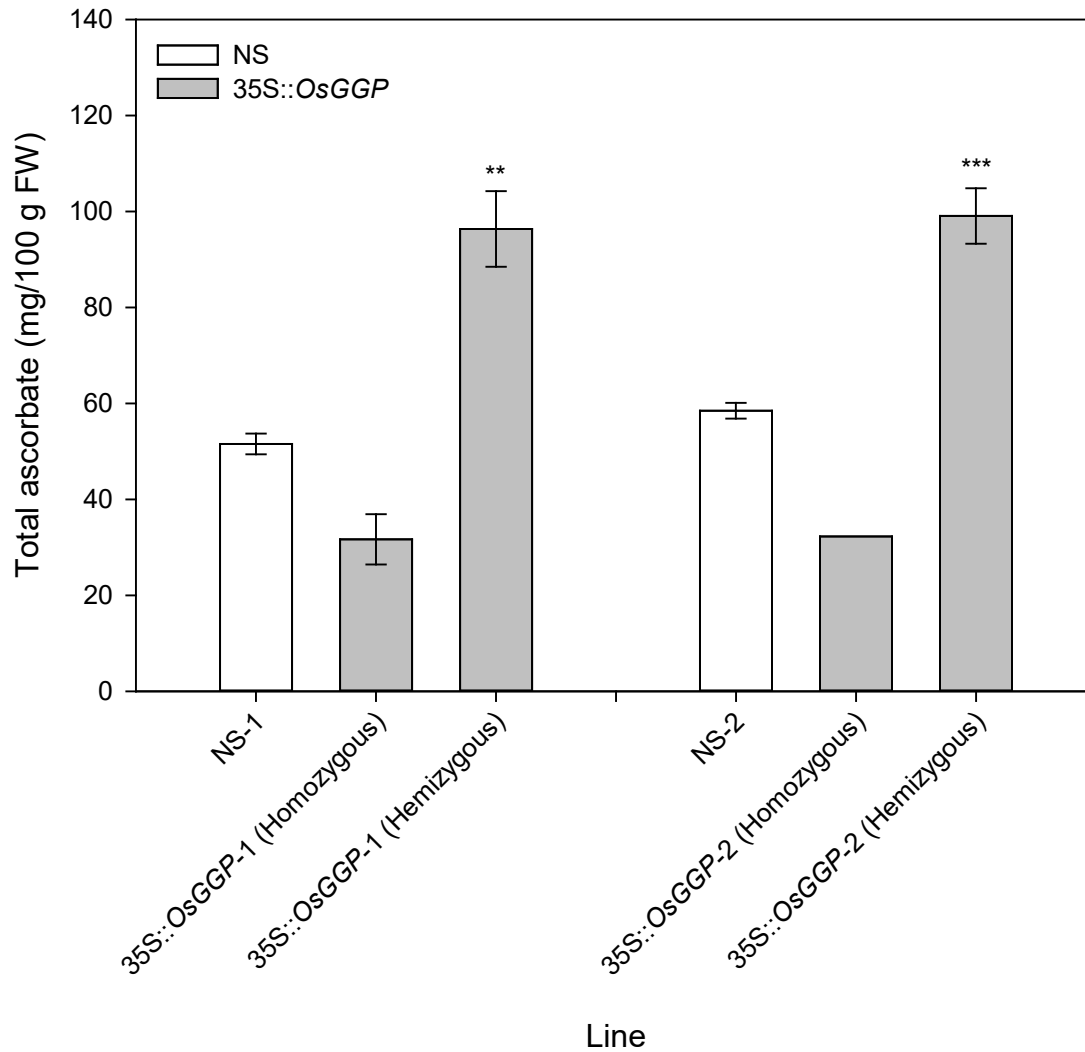


Figure S2.2. Foliar ascorbate concentrations in T₁ hemizygous and homozygous NS and 35S-*OsGGP* plants at the reproductive growth phase. Bars represent mean $\pm$ SEM of one to seven biological replicates. Asterisks indicate statistically significant differences between NS and 35S-*OsGGP* plants (two-sample t-test; ** p-value ≤ 0.01 ; *** p-value ≤ 0.001).

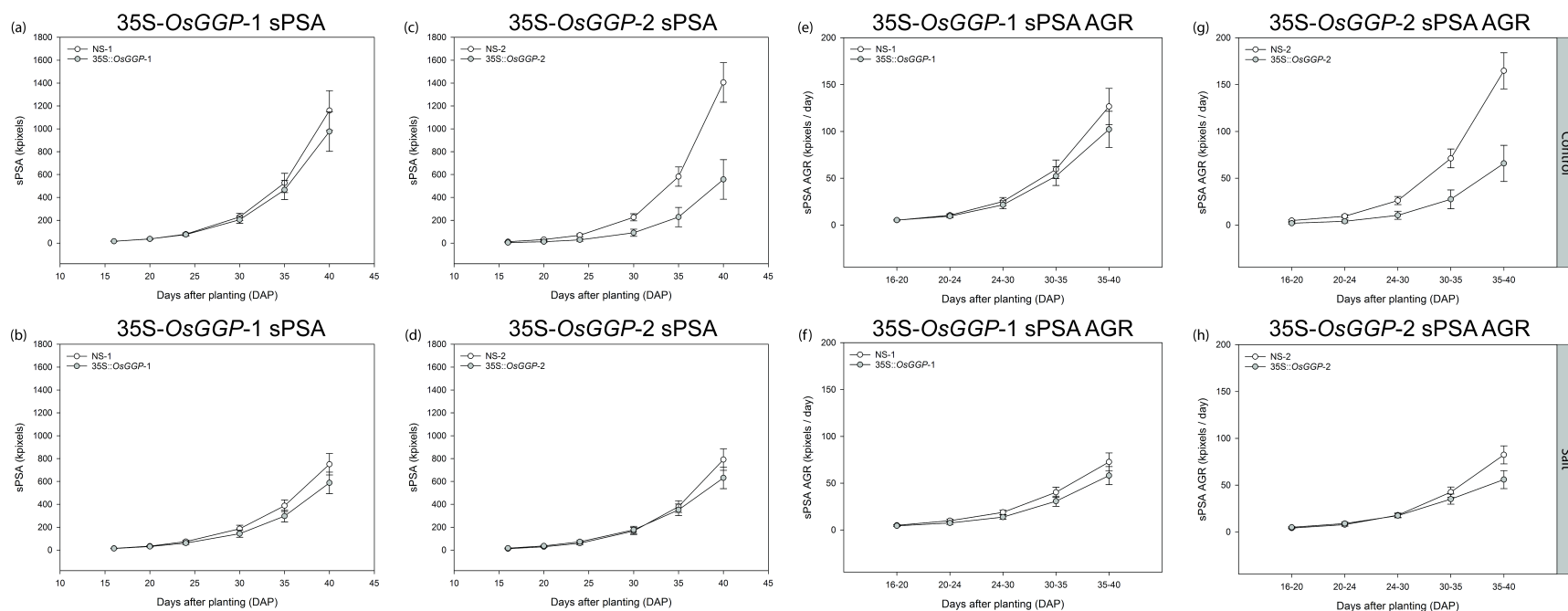


Figure S2.3. The sPSA and sPSA AGR of control and salt-stressed NS and 35S-*OsGGP* plants during the vegetative growth phase. The sPSA of (a) control and (b) salt-stressed T₄ homozygous NS-1 and 35S-*OsGGP*-1 plants and of (c) control and (d) salt-stressed T₃ homozygous NS-2 and 35S-*OsGGP*-2 plants. The sPSA AGR of (e) control and (f) salt-stressed T₄ homozygous NS-1 and 35S-*OsGGP*-1 plants and of (g) control and (h) salt-stressed T₃ homozygous NS-2 and 35S-*OsGGP*-2 plants. Salt was applied at DAP 20 and 23. Values represent mean $\pm$ half least significant pairwise (5%) difference of six biological replicates. Non-overlapping error bars indicate significant differences at $\alpha = 0.05$.

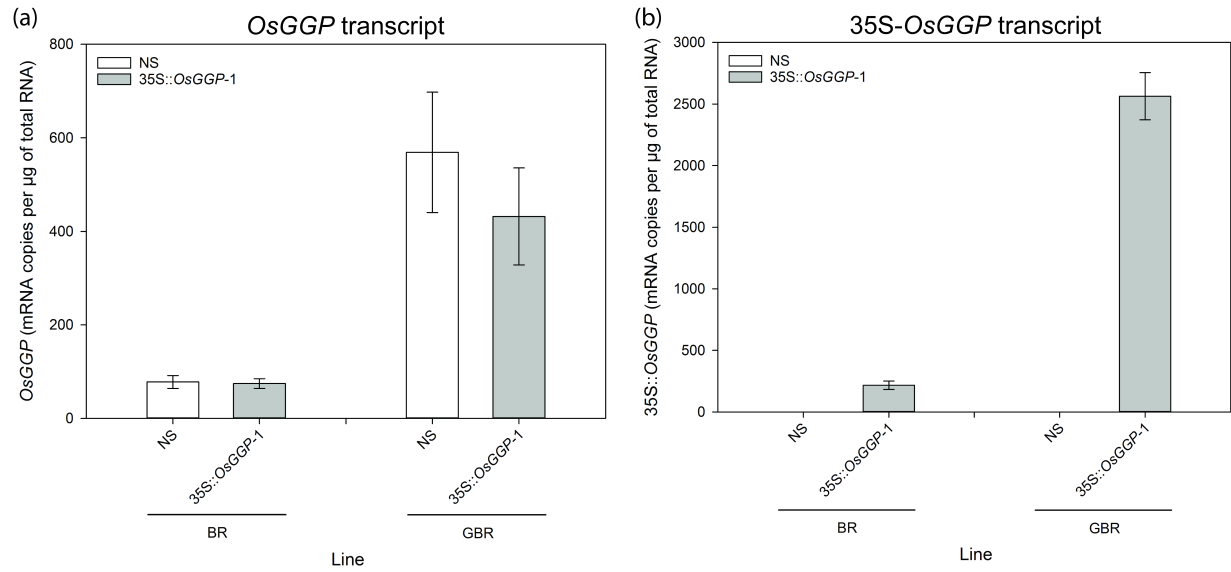


Figure S2.4. Transcript analysis of the endogenous *OsGGP* gene and 35S-*OsGGP* transgene in NS and 35S-*OsGGP* BR and GBR. Transcript levels of the (a) endogenous *OsGGP* gene and (b) 35S-*OsGGP* transgene in T₃ homozygous NS-1 and 35S-*OsGGP*-1 BR and GBR. Bars represent mean $\pm$ SEM of three independent replicates of subsampled grain, each with four technical replicates. No statistically significant differences were detected between NS and 35S-*OsGGP* plants (two-sample t-test).

Chapter 3.

Production and Characterization of Rice with CRISPR/Cas9-Induced Mutations in the *OsGGP* uORF

3.1. Introduction

Ascorbate, also known as vitamin C, is a vital multifunctional compound for both plants and mammals [5]. In plants, ascorbate has a well-established role in stress tolerance and photosynthetic functions. This is in large part due to ascorbate's ability to mitigate the damaging effects of reactive oxygen species (ROS) produced during normal or stressed cellular metabolism as a ROS scavenger or through the ascorbate-glutathione cycle [21]. Recently, ascorbate has also been reported to play a role in the reduction and transport of iron (Fe) into developing seeds [7, 218]. Therefore, developing strategies to increase ascorbate concentrations in crops could not only improve stress tolerance and photosynthetic performance, but also increase the density of Fe in seeds. In humans, ascorbate is an essential micronutrient well-known for its role in collagen formation and the prevention of the disease scurvy, but it also plays many other important roles in human physiology, such as promoting the dietary uptake of Fe in the gut [46] and epigenetic programming [42]. Ascorbate has many further beneficial effects on human health and is associated with reduced risk of infection, cardiovascular disease, stroke, diabetic complications, and cancer [40]. Developing strategies to increase ascorbate concentrations in crops could therefore also improve the dietary intake of this essential micronutrient in humans.

Several pathways towards ascorbate biosynthesis have been proposed in plants, but substantial genetic evidence now supports the L-galactose pathway as the dominant ascorbate biosynthetic pathway in plants [12-16]. The *GDP-L-galactose phosphorylase* (*GGP* or *VTC2/5*) gene—which encodes the enzyme responsible for the fifth enzymatic step of the L-galactose pathway, catalysing the conversion of GDP-L-galactose to L-galactose1-P—has emerged as the most promising candidate for increasing ascorbate concentrations in crops. Relative to other ascorbate biosynthetic genes, increased expression of the *GGP* gene consistently leads to the greatest fold increases in ascorbate concentrations in a wide range of model and crop species [62, 64, 67, 71, 73, 74, 187]. Moreover, as the first committed step toward ascorbate biosynthesis in the L-galactose pathway, the catalysis of GDP-L-galactose to L-galactose1-P by *GGP* is the most strongly regulated step in the L-galactose pathway, with the production of *GGP* being post-transcriptionally regulated through a highly conserved *cis*-acting upstream open reading frame (uORF) in the *GGP* mRNA, in addition to transcriptional regulation [17, 163].

The regulation of translation by uORFs—a class of small ORFs found in the 5' leader sequence upstream of the major ORF (mORF) of mRNAs—is a dynamic regulatory mechanism for the control of protein translation [219]. Generally, uORFs repress translation of the mORF by allowing fewer ribosomes to reach the mORF, for example by causing ribosome stalling in response to metabolite levels [220]. The metabolic control of translation by uORFs in plants has been reported for several metabolites in addition to ascorbate, including polyamines [221-230], sucrose [231-235], phosphocholine [236, 237], and galactinol [238].

The highly conserved *GGP* uORF is predicted to initiate from a non-canonical ATC or ACG start codon in the long 5' leading sequence of the *GGP* mRNA and encode a peptide of approximately 60 to 65-residues in length [163]. It has been proposed that under high ascorbate concentrations, the ribosome initiates at the *GGP* uORF and translation of the *GGP* uORF peptide stalls the ribosome, thereby preventing translation of the *GGP* mORF. Whereas under low ascorbate concentrations either the ribosome does not initiate at the *GGP* uORF, thereby skipping it, or the translation of the *GGP* uORF peptide does not cause the ribosome to stall, thereby allowing the *GGP* mORF to be translated. However, the mechanism of how ascorbate influences ribosomal initiation or ribosomal stalling of the *GGP* uORF has yet to be determined [163]. Disruption of the *GGP* uORF was reported to stop the ascorbate-mediated regulation of translation and increase foliar ascorbate concentrations when a *GGP* promoter-*uORF*-*GGP* construct was transiently transformed in *Nicotiana benthamiana* [163]. Recently, two studies reported that disrupting the *GGP* uORF with the CRISPR/Cas9 genome editing system increased ascorbate concentrations 1.7-fold in *Arabidopsis thaliana* [153], 1.4- to 2.6-fold in lettuce (*Lactuca sativa* L.) [153], and 1.4-fold in tomato (*Solanum lycopersicum* L.) [154]. Furthermore, the increased ascorbate concentrations in lettuce were associated with enhanced tolerance to methyl viologen-induced oxidative stress [153].

The editing of uORFs with the CRISPR/Cas9 genome editing system is emerging as a promising transgene-free strategy to fine tune gene translation for the improvement of crop traits [239]. The editing of uORFs offers many advantages over transgenic plants, for example by avoiding the occurrence of transgene silencing and the random insertion of T-DNAs [239]. Moreover, since transgenic plants require the introduction of foreign DNA they are considered a biosecurity risk and are thus subject to regulatory oversight [239]. Transgenic plants have also suffered from poor public acceptance due to the perceived risks of genetically modified organisms and values held by consumers [240]. Since plants with genome edited-induced indels (small insertions and deletions) due to errors in non-homologous end joining do not rely

on the presence of foreign DNA, they are more likely to be accepted by the public and subject to less regulatory oversight.

Chapter 2 described the production and characterization of rice (*Oryza sativa* L.) plants with constitutive overexpression of the *OsGGP* coding sequence. Characterization of homozygous 35S-*OsGGP* plants revealed that foliar ascorbate concentrations were reduced at the reproductive growth phase and was associated with reduced transcript levels of the endogenous *OsGGP* gene, suggesting the occurrence of gene silencing. This chapter therefore aimed to produce transgene-free, ascorbate-enriched rice plants with CRISPR/Cas9-induced mutations in the *OsGGP* uORF. The specific objectives of this chapter were to: (i) investigate the effect that CRISPR/Cas9-induced mutations in the *OsGGP* uORF have on ascorbate concentrations and *OsGGP* transcript levels, and (ii) investigate the effects that increased ascorbate concentrations in rice have on plant growth and development, agro-morphological performance, salt tolerance, and grain Fe levels.

3.2. Methods

3.2.1. Plant growth conditions

Oryza sativa cv. Nipponbare was used for all experiments. Rice grain were surface sterilized with 70% (v/v) ethanol (Chem-Supply, SA, Australia) for 1 min and 30% (v/v) bleach (White King, NSW, Australia) with a few drops of Tween-20 (Sigma-Aldrich, MO, USA) for 30 min and washed three times with sterile dH₂O. Surface sterilized rice grain were germinated in a petri dish with moist sterile filter paper (Whatman, UK) for 7 to 9 days prior to transplanting to 1 L pots filled with potting mix in a glasshouse maintained at 26°C and 70% relative humidity at The University of Melbourne (Melbourne, VIC, Australia). The potting mix was prepared by mixing one part washed fine sand (Col Smith, VIC, Australia), one part propagating sand (Brunnings, VIC, Australia), two parts premium vermiculate (Exfoliators, VIC, Australia), and one part General Mix potting media (Australian Growing Solutions, VIC, Australia) fertilized with Osmocote Exact Standard 8-9 M (ICL, NSW, Australia) at a rate of 6 g/L.

3.2.2. Vector construction

Two CRISPR/Cas9 vectors, p35S-Cas9 and pUbi-Cas9, containing the human codon optimised *Streptococcus pyogenes* Cas9 (*hSpCas9*) gene under the transcriptional control of the dual CaMV 35S promoter and the maize ubiquitin (Ubi) promoter, respectively, a single guide RNA (sgRNA) targeting the most conserved region of the *OsGGP* uORF under the transcriptional control of the rice OsU3 promoter, and the *hygromycin phosphotransferase II* (*hptII*) plant-selectable marker gene were assembled as previously described [241]. The sgRNAs were inserted downstream of the OsU3 promoter using the tRNA-gRNA system and BsaI cloning strategy as previously described [242]. The p35S-Cas9 sgRNA was assembled in two fragments with primers S1 (5'-GGTCTCATGGCAACAAAGCACCAGTGGTCTAG-3') and AS1 (5'-TAGGTCTCAGGCTGGGCTGGATGCACCAGCCGGGAATC-3') and S2 (5'-TAGGTCTCCAGCCCCGCACGGGTTTTAGAGCTAGAAATAG-3') and AS2 (5'-TAGGCATGACTATGGTCTCTGTAGTAAAAAAAAAAGCACCGACTCGGTGCCAC-3'), respectively. Similarly, the pUbi-Cas9 sgRNA was assembled in two fragments with primers S3 (5'-GGTCTCATGGCAACAAAGCACCAGTGGTCTAG-3') and AS3 (5'-TAGGTCTCAGAAGAGGAGGTCTGCACCAGCCGGGAATC-3') and S4 (5'-TAGGTCTCCCTTCCTCGCCGGGTTTTAGAGCTAGAAATAG-3') and AS4 (5'-

TAGGCATGACTATGGTCTCTGTAGTAAAAAAAAAAGCACCGACTCGGTGCCAC-3'), respectively. The resulting vectors were named p35S-Cas9-uORF and pUbi-Cas9-uORF, respectively, and were mobilised into *Agrobacterium tumefaciens* (strain AGL1) by electroporation prior to rice transformation.

3.2.3. Rice transformation

Transformation of rice callus was performed using established protocols [204]. Briefly, seed callus was induced on NB medium for 4 weeks (28°C, dark) followed by multiplication of embryonic units on fresh NB medium for a further 10-14 d (28°C, dark). Calli were then co-cultured with *A. tumefaciens* and placed on R2-CS medium for 3 d (25°C, dark). Transformed calli were then selected on R2-S medium for 14 d (28°C, dark), proliferated on NBS medium for 2-3 weeks (28°C, dark), and matured on PR-AG medium for 7-10 d (28°C, dark) prior to regeneration on RN medium for 2 d (28°C, dark) and then a further 4-6 weeks (28°C, 12 h light/12 h dark). Plantlets were then rooted on P medium for 3-4 weeks (28°C, 12 h light/12 h dark) prior to greenhouse acclimation in Jiffy peat pellets for 15 d, followed by transfer to soil pots.

3.2.4. Genotyping and sequencing

All primers used in this chapter were designed using Primer 3 software [207, 208]. Genomic DNA extractions were carried out using an Extract-N-Amp™ Plant PCR Kit (Sigma-Aldrich) according to manufacturer's direction. The PCR-amplification of target DNA was carried out in a final volume of 20 µL using MyTaq™ HS Red DNA Polymerase (Bioline, UK) according to manufacturer's instructions. The PCR amplification cycles consisted of 1 cycle = 1 min 95°C; 35 cycles = 20 s 95°C, 20 s 60°C, 30 to 90 s (depending on PCR product length) 72°C. Sanger sequencing was carried out by the Australian Genome Research Facility Ltd (<http://www.agrf.org.au/>). The primers used to multiplex-PCR *hptII* with *OsACT1* to detect the CRISPR/Cas9 T-DNAs, to sequence the *OsGGP* uORF, and detect the *uorf_{OsGGP-1}*, *uorf_{OsGGP-2}*, and *uorf_{OsGGP-3}* mutations are listed in Table 3.1.

3.2.5. Quantification of total ascorbate

Total ascorbate (ascorbate + dehydroascorbate) was extracted and measured as previously described with modifications [205]. Briefly, total ascorbate was extracted from ground, lyophilized tissue homogenised in extraction fluid containing 8% metaphosphoric acid, 2 mM EDTA, and 2 mM TCEP (Sigma-Aldrich) at 40°C for 2 h. The extract was centrifuged and 5

Table 3.1. List of PCR and qRT-PCR primers used in this chapter. The table provides primer name, forward and reverse primer sequences, PCR product length, and the application of the primers.

Primer Name	Forward primer sequence (5' - 3')	Reverse primer sequence (5' - 3')	Product length (bp)	Application
<i>OsACT1</i>	TCAGCAACTGGGATGATATGGAG	GCCGTTGTGGTGAATGAGTAAC	634	Detecting the CRISPR/Cas9 T-DNA
<i>hptII</i>	CTTGATGGAGCAGCAGACGC	CTATTTCTTTGCCCTCGGACG	308	Detecting the CRISPR/Cas9 T-DNA
<i>uorf_{OsGGP}</i>	GCCGACTGAGAAGAGGAAGAG	GAAGACGGCTTTTTCGGACT	1,127	Sequencing the <i>OsGGP</i> uORF
<i>uorf_{OsGGP-1}</i>	TGACCTCCTCTTCCTCGCA	TTCGGAATCCAAGGTGTAATAAAGA	356	Detecting the <i>uorf_{OsGGP-1}</i> mutation
<i>uorf_{OsGGP-2}</i>	CCCTCTGACCTCCTCTTCCTC	GATTTCGGAATCCAAGGTGTAATAAT	363	Detecting the <i>uorf_{OsGGP-2}</i> mutation
<i>uorf_{OsGGP-3}</i>	AAAGACGGCTAGCGTTGAAG	GATTGGATGTTAGAGAAGGTGGAT	228	Detecting the <i>uorf_{OsGGP-3}</i> mutation
qRT-PCR- <i>OsGGP</i>	AGATCTTCCTCTTCCTCAGT	CCGACAACCTGTTCTACCGT	368	Transcript analysis of the <i>OsGGP</i> gene
qRT-PCR- <i>OsGME1</i>	CGTATCAATGGTGTGAAGAGGTT	GGCCAGGCATCAGATTCCTT	107	Transcript analysis of the <i>OsGME1</i> gene
qRT-PCR- <i>OsGME2</i>	GATATGGGAGGGATGGGATT	GCCAAGCATCTGACTCCTTC	196	Transcript analysis of the <i>OsGME2</i> gene
qRT-PCR- <i>OsGPP</i>	TGCCTTCCTCAATGGTTCTC	CACACATGTTCAAGGCCAAC	188	Transcript analysis of the <i>OsGPP</i> gene
qRT-PCR- <i>OsGAPDH</i>	GGGCTGCTAGCTTCAACATC	TTGATTGCAGCCTTGATCTG	190	qRT-PCR housekeeping gene
qRT-PCR- <i>OsELF1</i>	GAAAGCAGCAGAAGAACGGG	TCCTCAAGCTTTGCCATGTCT	126	qRT-PCR housekeeping gene
qRT-PCR- <i>OsACT1</i>	ATCACTGCCTTGGCTCCTAG	TCGTA CTAGCCTTGGCAA	140	qRT-PCR housekeeping gene
qRT-PCR- <i>OsUBQ5</i>	TACACCAAGCCAAGAAGCAG	GTGAGGCCGCACTTACCG	191	qRT-PCR housekeeping gene
OT-1	TGTGGAGGTGATTGTGGAGA	ATCCGAGAACAGTCCGAAGA	602	Sequencing the OT-1 site
OT-2	TGAAGCGAGTTGAACGAGTG	ATTGTGAGGAAGGGGATCG	509	Sequencing the OT-2 site
OT-3	TCCTTCACTGCATCAGCAAC	GCAAAAACGCAAAACAAGCTC	455	Sequencing the OT-3 site
OT-4	GCCACCTCTCTCAGTGCTTC	CTCCTCCA ACTCCAATCCAA	785	Sequencing the OT-4 site
OT-5	TGTGCTTGCTCAATGTTGT	AAGCAGATCGACAGTGTAACG	528	Sequencing the OT-5 site
OT-6	GAGGGGTGAGGAAGAGGAAG	CAGACAGATGGGTCCCAGAG	532	Sequencing the OT-6 site

μL of supernatant was injected onto a C18 3 μm 33 x 7 mm Alltima Rocket column (Hichrom Limited, UK) maintained at 40°C with a flow rate of 1 mL/min. The concentration was determined by reverse phase chromatography on a Shimadzu Nexera UHPLC system (Shimadzu, Japan) with simultaneous UV and MS detection. Mobile phase A consisted of MS grade water with 0.1% formic acid (ThermoFisher Scientific, MA, USA). The elution procedure utilized a 1 min gradient from 0% to 90% of mobile phase B: MS grade acetonitrile with 0.1% formic acid (ThermoFisher Scientific). The UV absorption signal was acquired using a Shimadzu SPD-20A detector at 245 nm wavelength. The MS data was acquired on a Shimadzu LCMS-8050 triple quadrupole mass spectrometer equipped with a DUIS source. The MS instrument was operated in Multiple Reaction Monitoring mode monitoring ions in both positive (177>95, 177>85) and negative (175 >115, 175>71, 175>59) ionisation modes. The product ion 115 *m/z* was used as a quantifier. A calibration curve in the range of 1.95 to 250 ppm of L-ascorbic acid analytical standard (Sigma-Aldrich) was used to determine absolute concentration of analyte in the extracts.

3.2.6. Preparation of germinated brown rice

Following surface sterilization, brown rice was washed three times with sterile dH₂O and spread in a petri dish with moist sterile filter paper (Whatman) for 24, 48, and 72 h in the dark at 28°C. The germinated brown rice was then snap-frozen in liquid nitrogen, lyophilized, and ground using a Tube Mill with a 40 mL grinding chamber (IKA, Germany) prior to ascorbate measurement.

3.2.7. Quantitative reverse transcription-PCR analysis

Total RNA was purified from leaf tissue using the Direct-zol™ RNA MiniPrep Kit (Zymo Research, CA, USA) according to manufacturer's instructions. Quantification, DNase treatment, and reverse transcription of the RNA was performed using the QuantiFluor® RNA System (Promega, WI, USA), RQ1 RNase-Free DNase (Promega), and the Tetro cDNA Synthesis Kit (Bioline), respectively, according to manufacturer's instructions. The primers used for the quantitative reverse transcription-PCR (qRT-PCR) analysis of the *OsGPP*, *OsGME1*, *OsGME2*, and *OsGPP* transcripts are listed in Table 3.1. Primer efficiency was ≥93% for all primer pairs. The qRT-PCR analysis was carried out on a CFX Connect™ Real-Time System (Bio-Rad, CA, USA) with Hard-Shell® 96-Well PCR Plates (Bio-Rad). The qRT-PCR amplification cycles consisted of 1 cycle = 3 min 95°C; 40 cycles = 20 s 95°C, 20 s 60°C, 20 s 72°C and were followed by a melt curve from 65°C to 95°C in 0.5°C steps for 5 s/step.

The qRT-PCRs were performed in a final volume of 10 μ L for KAPA SYBR® FAST (Kapa Biosystems, MA, USA) according to manufacturer's instructions, each with three technical replicates. Absolute quantification of transcript copy number was determined using a 10-fold serial dilution of 10^8 – 10^2 copies for each PCR product and a standard curve generated with the Bio-Rad CFX Manager 3.1 software (Bio-Rad). The PCR products were purified and quantified with a DNA Clean & Concentrator™-5 Kit (Zymo Research) and the QuantiFluor® ONE dsDNA System (Promega), respectively, prior to serial dilution. The geometric mean of the transcript levels of three optimal housekeeping genes selected from *OsGAPDH*, *OsELF1*, *OsACT1*, and *OsUBQ5* using the NormFinder Excel add-in [209] was used to normalize the transcript levels of the *OsGGP*, *OsGME1*, *OsGME2*, and *OsGPP* genes.

3.2.8. Assessment of salt tolerance and plant growth and development with automated imaging

Surface sterilized rice grain were germinated in plastic boxes on moist paper towels (January 9, 2019, denoted DAP 0). At DAP 6, three uniformly germinated grain per line were transplanted to 2.8 L pots filled with 2.6 kg of UC Davis Mix, placed into square containers for water to collect, and grown on benches in the NW Smarthouse maintained at 26°C and 70% relative humidity at The Plant Accelerator, Adelaide, Australia under manual watering (not to weight). At DAP 14, plants were thinned to one seedling per pot and loaded onto the NW Smarthouse for automated imaging. A split-unit design was used to randomize pots (i.e. line and treatment combinations) to cart positions within the Smarthouse. The plants were automatically watered daily so that approximately 500 mL of water was maintained in the soil. The salt plants were treated with saline solution on DAP 20 and 23 (100 mL of 225mM NaCl) to reach a final concentration of 90 mM NaCl in the soil solution, while the control plants received equivalent water applications as previously described [196]. Imaging was carried out daily from DAP 14 to 40 inclusive. From these images the PSA of the plant, as viewed using an RGB camera, was obtained as previously described [210]. For analysis purposes, PSA was defined as the sum of the areas as measured (in kilopixels) from six side/oblique camera views plus twice the shoot area (top view). For the assessment of plant growth and development over the course of a full lifecycle, imaging commenced at DAP 14 and was carried out approximately weekly (DAP 14 to 128 inclusive). During the first half of the experiment (DAP 14 to 65), the plants were loaded on the NW Smarthouse and automatically watered-to-weight daily. A split-unit design was used to randomize pots (i.e. line and treatment combinations) to cart positions within the Smarthouse. Thereafter, the plants were grown on benches in a

conventional greenhouse, also maintained at 26°C and 70% relative humidity, under manual watering (not to weight) and manual imaging (conducted weekly in the NW Smarthouse). From these images the PSA of the plant, as viewed using an RGB camera, was obtained as previously described [210]. For analysis purposes, PSA was defined as the sum of the areas as measured (in kilopixels) from two side views and the top view. The data was prepared for analysis using growthPheno [211], a package for the R statistical computing environment [212].

3.2.9. Elemental analysis of brown rice

Brown rice was washed three times in a 0.1% Tween 20 (Sigma-Aldrich) solution, rinsed with dH₂O, and oven dried for 48 h at 60°C prior to grinding with a Tube Mill with a 40 mL grinding chamber (IKA). Inductively coupled plasma optical emission spectrometry (ICP-OES) of the ground brown rice samples was carried out by the Melbourne Trace Analysis for Chemical, Earth and Environmental Sciences Platform at The University of Melbourne (<https://chemicalanalysis.unimelb.edu.au/>).

3.2.10. Analysis of potential off-target edits

Potential off-target sites of the pUbi-Cas9-uORF sgRNA were predicted using Cas-OFFinder (<http://www.rgenome.net/cas-offinder/>) with the following parameters: mismatch number $\leq$ 2, DNA bulge size $\leq$ 0, and RNA bulge size $\leq$ 0. Three T₂ homozygous, transgene-free WT, *uorf_{OsGGP-1}*, *uorf_{OsGGP-2}*, and *uorf_{OsGGP-3}* plants were used as representatives for the analysis. Genomic DNA was extracted and target DNA amplified with the primers listed in Table 3.1 prior to Sanger sequencing.

3.2.11. Statistical analysis

Statistically significant differences between the WT and *uorf_{OsGGP}* plants were detected with the two-sample t-test using the 5% level of significance and calculated with Minitab 17.1.0 (<http://www.minitab.com/en-us/>). Arcsine and log transformations were performed on percentage and ratio values, respectively, prior to detecting statistically significant differences. A mixed model analysis was performed on the data for each trait from the imaged plants using the R packages ASReml-R [216] and asremlPlus [217]. Estimated means were obtained from the analysis and the significances of their differences were investigated using the Least Significant Difference at the 5% significance level.

3.3. Results

3.3.1. Production of rice with CRISPR/Cas9-induced mutations in the *OsGGP* uORF

Two CRISPR/Cas9 vectors, p35S-Cas9-uORF and pUbi-Cas9-uORF, were transformed into rice cv. Nipponbare via *A. tumefaciens* (Figure 3.1). A total of 168 independent p35S-Cas9-uORF transformation events were recovered from tissue culture. From each independent p35S-Cas9-uORF transformation event, DNA was extracted from shoots regenerating from hygromycin-resistant callus and the presence of the p35S-Cas9-uORF T-DNA was confirmed by multiplex-PCR of the *hptII* and *OsACT1* genes. The presence of CRISPR/Cas9-induced mutations in the *OsGGP* uORF were screened via Sanger sequencing for all 168 independent p35S-Cas9-uORF transformation events, but no mutations were detected (0% mutation frequency). The 168 independent p35S-Cas9-uORF transformation events were destroyed and not studied further in this chapter. Next, a total of 456 independent pUbi-Cas9-uORF transformation events were recovered from tissue culture. From each independent pUbi-Cas9-uORF transformation event, DNA was again extracted from shoots regenerating from hygromycin-resistant callus and the presence of the pUbi-Cas9-uORF T-DNA was confirmed by multiplex-PCR of the *hptII* and *OsACT1* genes. The presence of CRISPR/Cas9-induced mutations in the *OsGGP* uORF were screened via Sanger sequencing for all 456 independent pUbi-Cas9-uORF transformation events and two plants with mutations in the *OsGGP* uORF were detected (1.63% mutation frequency). Approximately five weeks later when the pUbi-Cas9-uORF transformation events were transferred to Jiffy peat pellets for greenhouse acclimation, DNA was once again extracted from shoots from a subsample of 49 independent pUbi-Cas9-uORF transformation events and the presence of CRISPR/Cas9-induced mutations in the *OsGGP* uORF screened via Sanger sequencing. This time, an additional 29 plants with mutations in the *OsGGP* uORF were detected (59.18% mutation frequency). A summary of the mutation types and frequencies detected in the *OsGGP* uORF in the 31 plants are summarized in Figure 3.2.

3.3.2. Isolation of homozygous, transgene-free *uorf_{OsGGP}* mutants

Three independent T₀ *uorf_{OsGGP}* mutants recovered from CRISPR/Cas9-targeted mutagenesis with the pUbi-Cas9-uORF were pursued for further analysis in this chapter: one that was bi-allelic for a single adenine insertion and a single thymine insertion, one that was bi-allelic for a single thymine insertion and a 5-bp deletion, and one that was heterozygous for a 63-bp

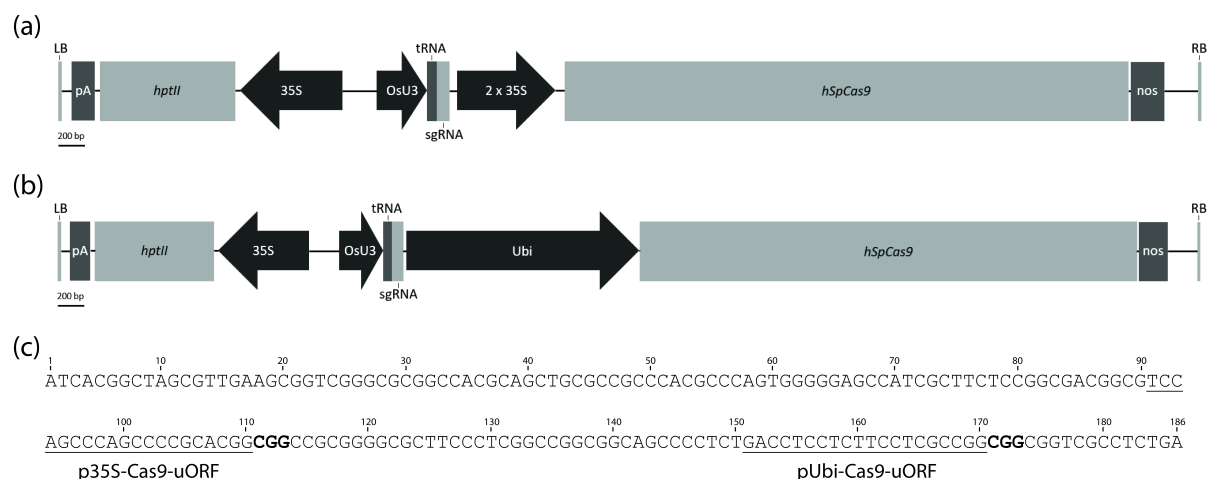


Figure 3.1. CRISPR/Cas9-targeted mutagenesis of the *OsGGP* uORF. Schematic representation of the (a) p35S-Cas9-uORF and (b) pUbi-Cas9-uORF T-DNA. LB, left border; pA, CaMV poly(A) signal; *hptII*, hygromycin phosphotransferase II; 35S, CaMV 35S promoter; OsU3, rice U3 promoter; tRNA, transfer RNA; sgRNA, single guide RNA; 2 x 35S, dual CaMV 35S promoter; Ubi, maize ubiquitin promoter; *hSpCas9*, human codon optimized *Streptococcus pyogenes* Cas9; nos, nopaline synthase terminator; RB, right border. (c) Nucleotide sequence of the *OsGGP* uORF. The sgRNA target sites are underlined and the protospacer-adjacent motifs are shown in bold.

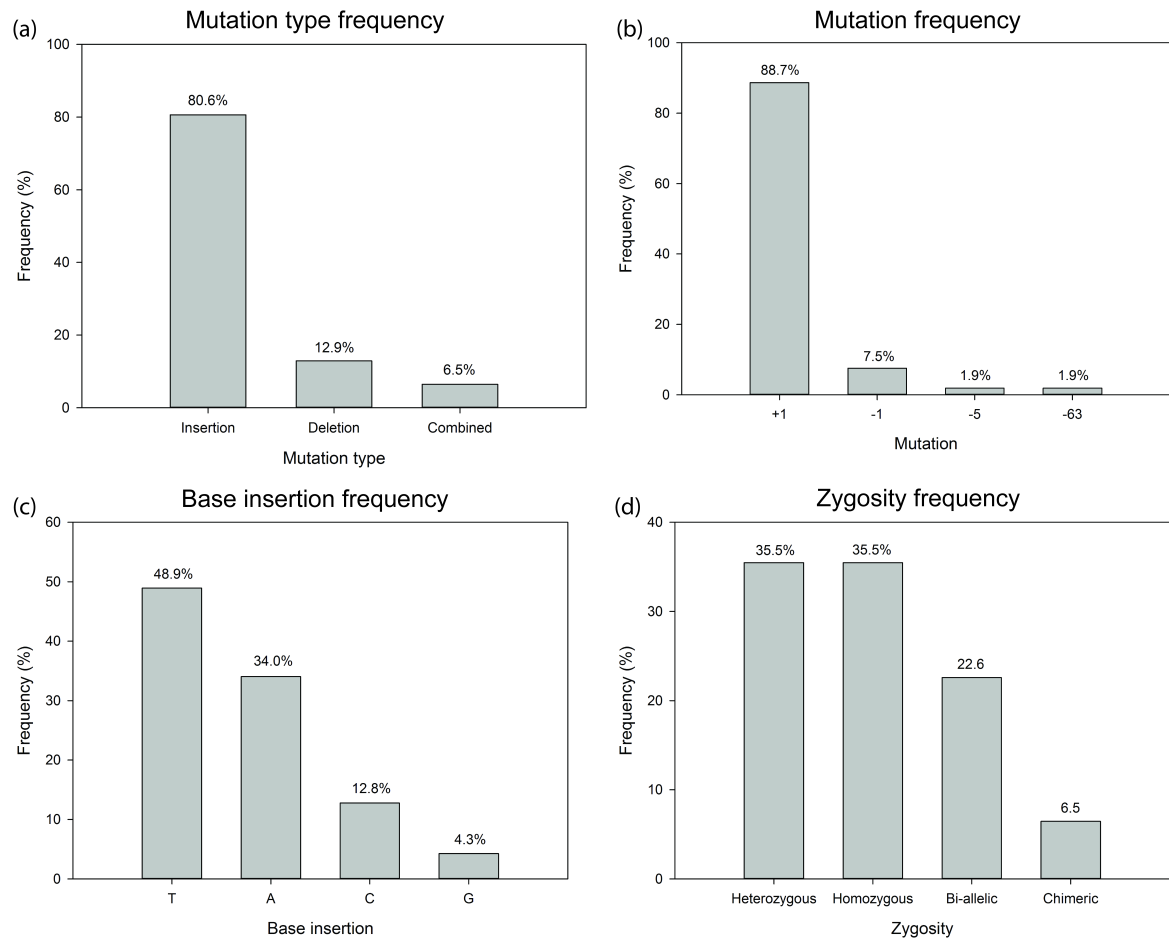


Figure 3.2. CRISPR/Cas9-induced mutation types and frequencies in the 31 plants with mutations detected in the *OsGGP* uORF transformed with the pUbi-Cas9-uORF T-DNA. (a) Mutation type frequency, (b) mutation frequency, (c) base insertion frequency in 1-bp insertion mutants, and (d) zygosity frequency.

deletion (Figure 3.3). From these three parental T₀ mutants, three independent homozygous, transgene-free *uorfOsGGP* mutants were subsequently isolated through PCR-based genotyping: one with a single adenine insertion (*uorfOsGGP-1*), one with a 5-bp deletion (*uorfOsGGP-2*), and one with a 63-bp deletion (*uorfOsGGP-3*) (Figure 3.4a). Consequently, at the peptide level, the *uorfOsGGP-1* mutation disrupts the reading frame replacing the final 5 residues and introduces a delayed stop codon extending the peptide from 61 to 72 residues in length (Figure 3.4b). The *uorfOsGGP-2* mutation deletes two residues and disrupts the reading frame replacing the final 5 residues and introduces a delayed stop codon extending the peptide from 61 to 70 residues in length (Figure 3.4b). The *uorfOsGGP-3* mutation deletes 21 residues whilst maintaining the reading frame reducing the peptide from 61 to 40 residues in length (Figure 3.4b). An independent plant recovered from CRISPR/Cas9-targeted mutagenesis with the pUbi-Cas9-uORF with no mutations detected in the *OsGGP* uORF was used as a wild-type (WT) control in this chapter. The *uorfOsGGP-1* mutant and WT control were studied in further detail in this chapter as they were the most advanced lines.

3.3.3. Ascorbate concentrations of the *uorfOsGGP* mutants

Foliar ascorbate concentrations were significantly increased 1.2- and 1.3-fold in the T₂ homozygous, transgene-free *uorfOsGGP-2* and *uorfOsGGP-3* mutants, respectively, relative to WT (Figure 3.5a). Similarly, foliar ascorbate concentrations were increased 1.2-fold in the T₂ homozygous, transgene-free *uorfOsGGP-1* mutant relative to WT but did not differ significantly (Figure 3.5a). Foliar ascorbate concentrations were also increased 1.2- to 1.7-fold in a range of T₁ heterozygous, homozygous, and bi-allelic, transgene-free *uorfOsGGP-2* and *uorfOsGGP-3* mutants relative to WT (Figure 3.5b). Ascorbate concentrations were negligible in both T₃ homozygous, transgene-free WT and *uorfOsGGP-1* brown rice and did not differ significantly (Figure 3.5c). Ascorbate concentrations were significantly decreased 1.4-fold in *uorfOsGGP-1* brown rice germinated for 24 h relative to WT, but otherwise did not differ significantly in brown rice germinated for 48 and 72 h.

3.3.4. Transcript analysis of the *uorfOsGGP* mutants

Transcript levels of the *OsGGP* gene did not differ significantly between the T₂ homozygous, transgene-free WT and *uorfOsGGP* plants (Figure 3.6a). Similarly, transcript levels of the *GDP-D-mannose-3',5'-epimerase 1* and 2 (*OsGME1* and *OsGME2*) genes, which encode the enzyme responsible for the fourth enzymatic step of the L-galactose pathway, and the *L-galactose-1-phosphate phosphatase* (*OsGPP*) gene, which encodes the sixth enzymatic step of the L-

WT	1 10 20 30 40 50 60 70 80 90	ATCACGGCTAGCGTTGAAGCGGTCGGGCGCGGCCACGCAGCTGCGCCGCCACGCCCAGTGGGGGAGCCATCGCTTCTCCGGCGACGGCGTCC
<i>uorf_{OsGGP}</i> -1 (+A, +T) Bi-allelic	{	ATCACGGCTAGCGTTGAAGCGGTCGGGCGCGGCCACGCAGCTGCGCCGCCACGCCCAGTGGGGGAGCCATCGCTTCTCCGGCGACGGCGTCC ATCACGGCTAGCGTTGAAGCGGTCGGGCGCGGCCACGCAGCTGCGCCGCCACGCCCAGTGGGGGAGCCATCGCTTCTCCGGCGACGGCGTCC
<i>uorf_{OsGGP}</i> -2 (+T, -5) Bi-allelic	{	ATCACGGCTAGCGTTGAAGCGGTCGGGCGCGGCCACGCAGCTGCGCCGCCACGCCCAGTGGGGGAGCCATCGCTTCTCCGGCGACGGCGTCC ATCACGGCTAGCGTTGAAGCGGTCGGGCGCGGCCACGCAGCTGCGCCGCCACGCCCAGTGGGGGAGCCATCGCTTCTCCGGCGACGGCGTCC
<i>uorf_{OsGGP}</i> -3 (-63, WT) Hetero		ATCACGGCTAGCGTTGAAGCGGTCGGGCGCGGCCACGCAGCTGCGCCGCCACGCCCAGTGGGGGAGCCATCGCTTCTCCGGCGACGGCGTCC
WT	100 110 120 130 140 150 160 170 180 183	AGCCCAGCCCCGCACGGCGGCCGCGGGGCGCTTCCCTCGGCCGGCGGCAGCCCCTCT <u>GACCTCCTCTTCCTCGCCGG</u> CGG CGGTCGCCTCTGA
<i>uorf_{OsGGP}</i> -1 (+A, +T) Bi-allelic	{	AGCCCAGCCCCGCACGGCGGCCGCGGGGCGCTTCCCTCGGCCGGCGGCAGCCCCTCTGACCTCCTCTTCCTCGC A CGGCGGCGGTTCGCCTCTGA AGCCCAGCCCCGCACGGCGGCCGCGGGGCGCTTCCCTCGGCCGGCGGCAGCCCCTCTGACCTCCTCTTCCTCGC T CGGCGGCGGTTCGCCTCTGA
<i>uorf_{OsGGP}</i> -2 (+T, -5) Bi-allelic	{	AGCCCAGCCCCGCACGGCGGCCGCGGGGCGCTTCCCTCGGCCGGCGGCAGCCCCTCTGACCTCCTCTTCCTCGC T CGGCGGCGGTTCGCCTCTGA AGCCCAGCCCCGCACGGCGGCCGCGGGGCGCTTCCCTCGGCCGGCGGCAGCCCCTCTGACCTCCTCTTC-----CGGCGGCGGTTCGCCTCTGA
<i>uorf_{OsGGP}</i> -3 (-63, WT) Hetero		AGCCCAGCCCCGCA-----CGGCGGTCGCCTCTGA

Figure 3.3. Nucleotide sequence alignment of the parental T₀ CRISPR/Cas9-induced *uorf_{OsGGP}* mutants used in this chapter. The sgRNA target site is underlined. The protospacer-adjacent motif, sequence changes, and deleted nucleotides are shown in bold, red, and as hyphens, respectively.



Figure 3.4. Nucleotide and peptide sequence alignment of the CRISPR/Cas9-induced homozygous, transgene-free *uorf_{OsGGP}* mutants used in this chapter. **(a)** Nucleotide sequence alignment. The sgRNA target site is underlined. The protospacer-adjacent motif, sequence changes, and deleted nucleotides are shown in bold, red, and as hyphens, respectively. **(b)** Peptide sequence alignment. Sequence changes and deleted residues are shown in red and as hyphens, respectively.

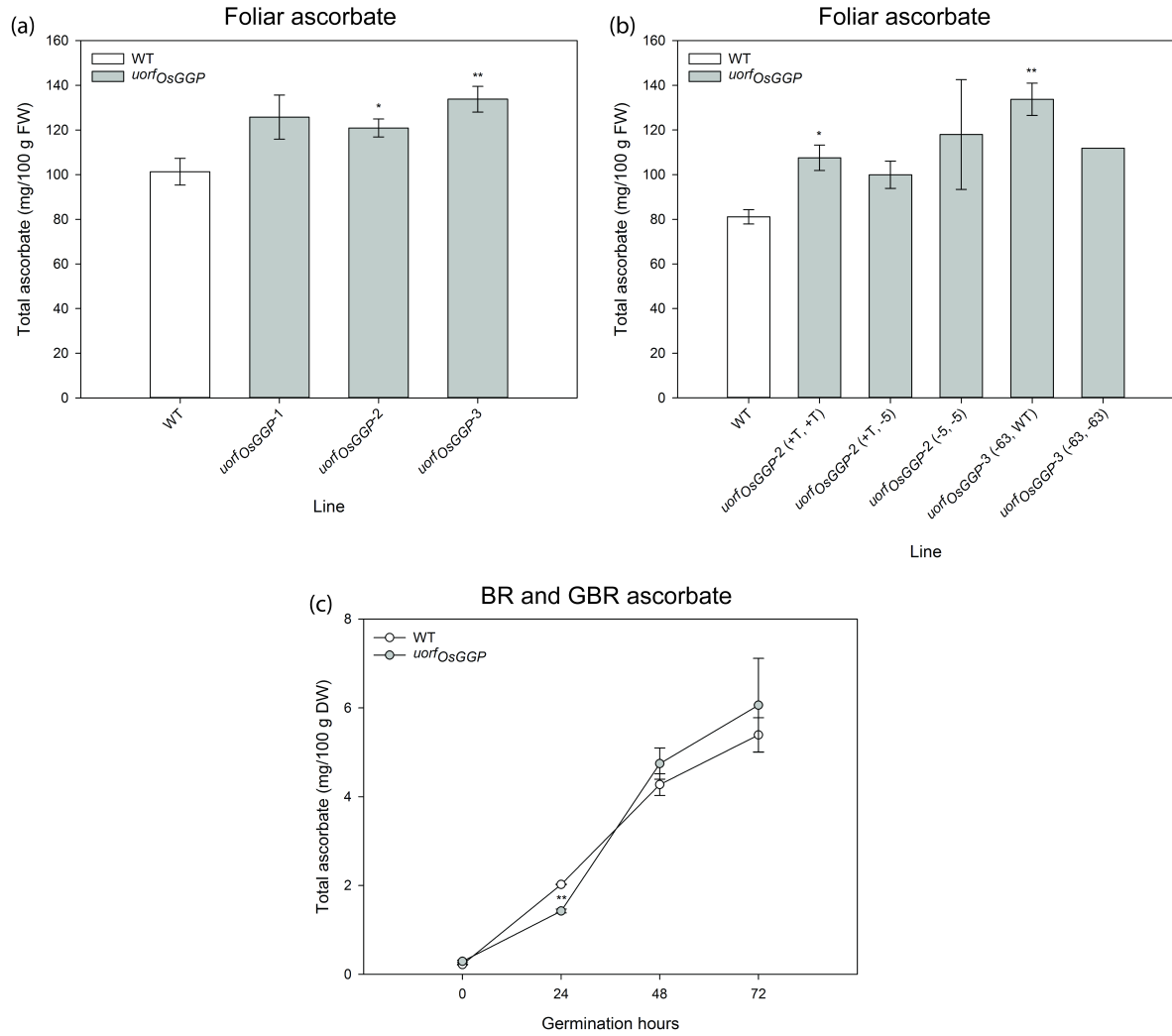


Figure 3.5. Ascorbate concentrations of WT and *uorfOsGGP* plants. **(a)** Foliar ascorbate concentrations of T₂ homozygous, transgene-free WT and *uorfOsGGP* plants. Bars represent mean $\pm$ SEM of six biological replicates. **(b)** Foliar ascorbate concentrations of T₁ heterozygous, homozygous, and bi-allelic, transgene-free WT and *uorfOsGGP* plants. Bars represent mean $\pm$ SEM of one to five biological replicates. **(c)** Ascorbate concentrations of T₃ homozygous, transgene-free WT and *uorfOsGGP*-1 brown rice and germinated brown rice. Values represent mean $\pm$ SEM of three independent replicates of approximately 50 grain. Asterisks indicate statistically significant differences between WT and *uorfOsGGP* plants (two-sample t-test; * p-value ≤ 0.05 ; ** p-value ≤ 0.01).

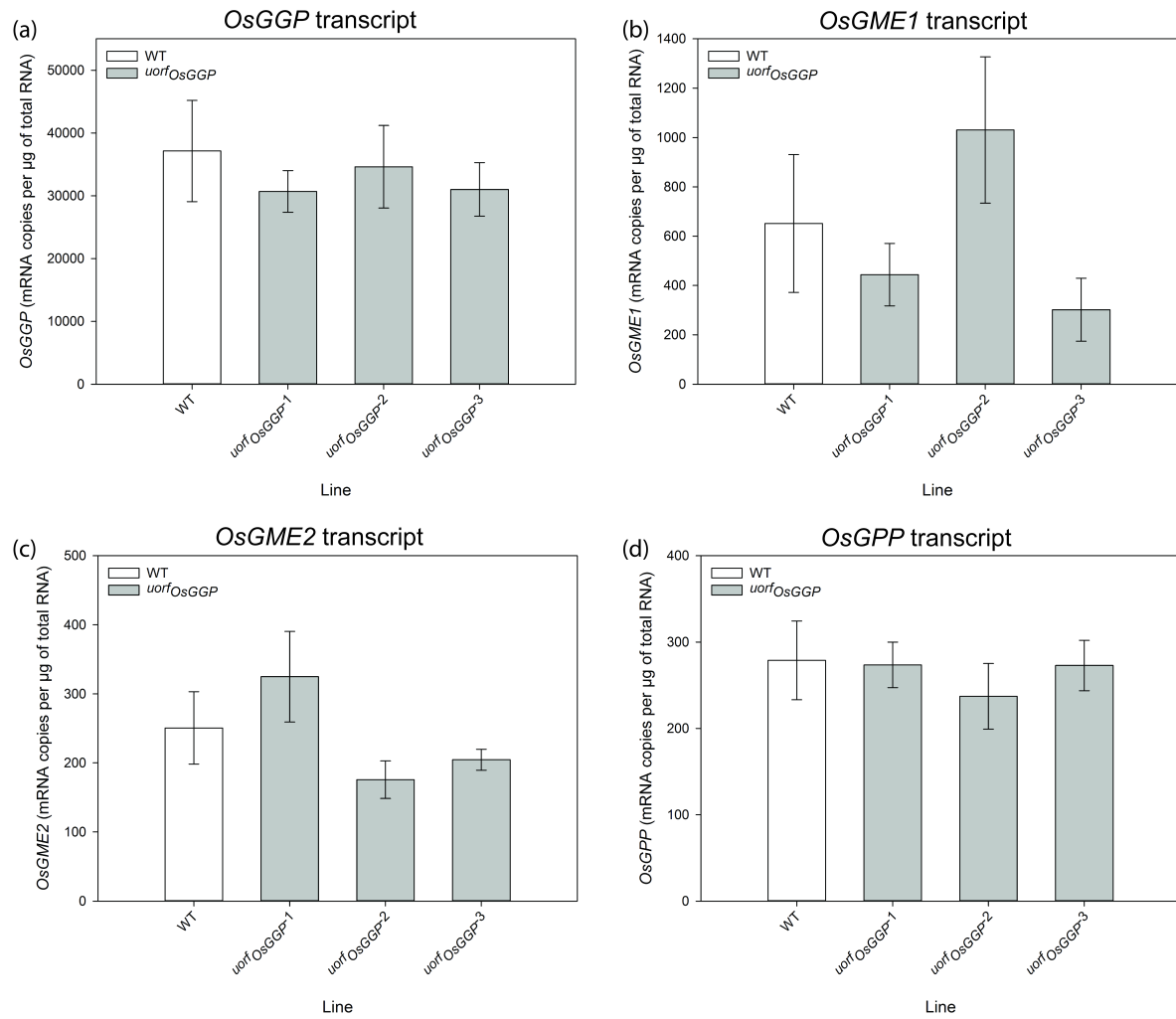


Figure 3.6. Transcript analysis of WT and *uorfOsGPP* plants. Transcript levels of the (a) *OsGPP*, (b) *OsGME1*, (c) *OsGME2*, and (d) *OsGPP* genes in T₂ homozygous, transgene-free WT and *uorfOsGPP* plants. Bars represent mean $\pm$ SEM of six biological replicates. No statistically significant differences were detected between WT and *uorfOsGPP* plants (two-sample t-test).

galactose pathway, did not differ significantly between the WT and *uorfo_{sGGP}* plants (Figure 3.6c-d).

3.3.5. Shoot growth measurements of the *uorfo_{sGGP}*-1 mutant

Shoot growth measurements of T₂ homozygous, transgene-free WT and *uorfo_{sGGP}*-1 plants were carried out using automated imaging over the course of a full lifecycle (128 days). Imaging of the plants commenced at DAP 14 and was carried out approximately weekly (DAP 14 to 128 inclusive). The smoothed PSA (sPSA), a strong predictor of shoot biomass, was calculated for DAP 16, 30, 44, 58, 79, 101, and 128. The WT and *uorfo_{sGGP}*-1 plants did not differ significantly with respect to sPSA values at any of the calculated timepoints (Figure 3.7a). From the sPSA values, the sPSA absolute growth rate (AGR) and sPSA relative growth rate (RGR) were calculated at the following DAP intervals: 16-30, 30-44, 44-58, 58-79, 79-101, and 101-128. The WT and *uorfo_{sGGP}*-1 plants did not differ significantly with respect to sPSA AGR values at any of the calculated DAP intervals (Figure 3.7b). The *uorfo_{sGGP}*-1 mutant had a significantly higher sPSA RGR value for the DAP interval 16-30 relative to WT, but otherwise did not differ significantly at any of the other calculated DAP intervals (Figure 3.7c).

3.3.6. Agro-morphological performance of the *uorfo_{sGGP}*-1 mutant

The T₂ homozygous, transgene-free *uorfo_{sGGP}*-1 mutant had a significant 1.2-fold decrease in panicle length, 1.1-fold decrease in the number of primary branches per panicle, and 1.2-fold decrease in the number of spikelets per panicle relative to WT, but had a significant 1.3-fold increase in both the number of panicles per plant and grain weight per plant (Table 3.2). The WT and *uorfo_{sGGP}*-1 plants did not differ significantly with respect to straw biomass, the number of fertile spikelets per panicle, spikelet fertility, and thousand grain weight (Table 3.2). The *uorfo_{sGGP}*-1 mutant had a minor significant increase in the grain length to width ratio relative to WT, but otherwise did not differ significantly with respect to the percentage of chalky grain, grain length, grain width, and grain area (Table 3.3).

3.3.7. Salt stress tolerance assessment of the *uorfo_{sGGP}*-1 mutant

Shoot growth measurements of control and salt-stressed T₂ homozygous, transgene-free WT and *uorfo_{sGGP}*-1 plants were carried out using automated imaging. Salt stress was applied in two steps to the plants at DAP 20 and 23 and the plants were imaged daily from DAP 14 to 40 inclusive. Foliar ascorbate concentrations were significantly increased 1.3-fold in both the

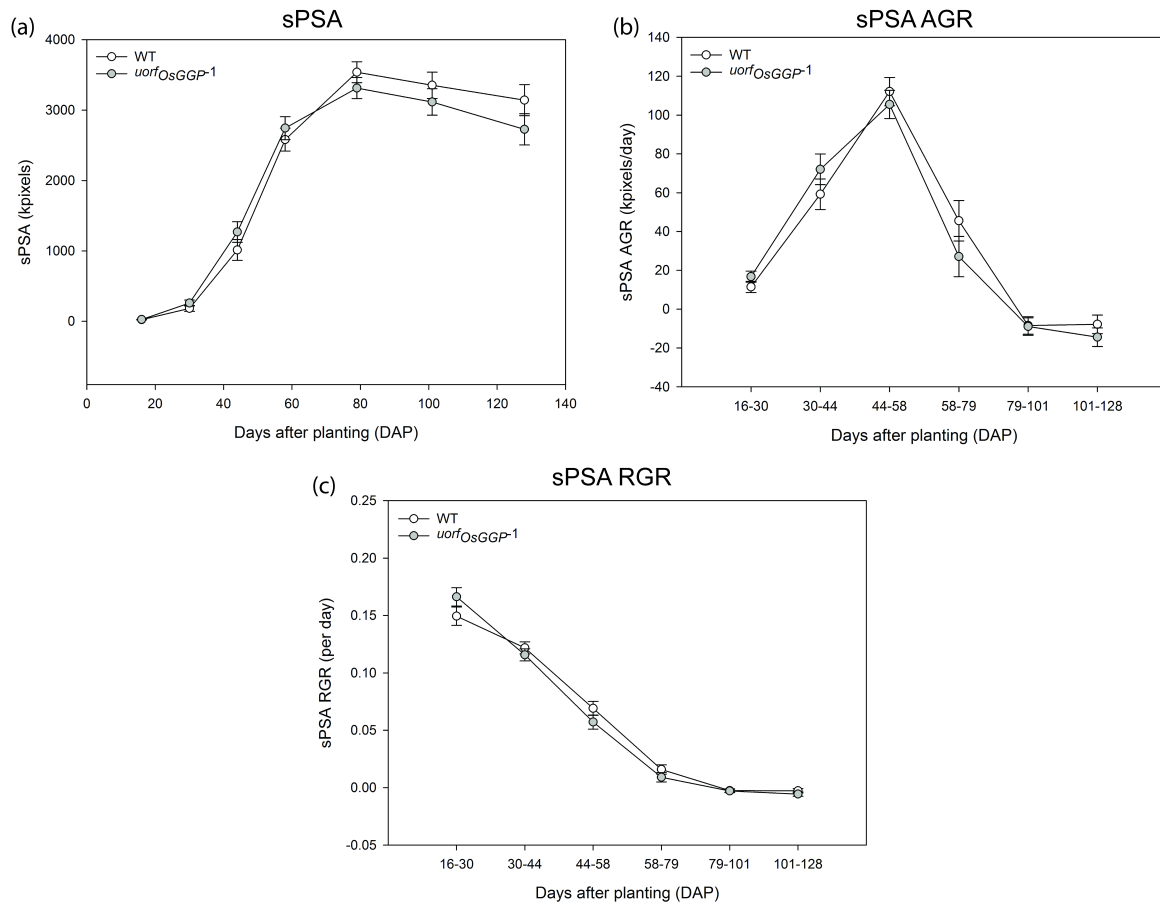


Figure 3.7. Shoot growth measurements of WT and *uorfOsGGP-1* plants using automated imaging over the course of a full lifecycle. The (a) sPSA, (b) sPSA AGR, and (c) sPSA RGR of T₂ homozygous, transgene-free WT and *uorfOsGGP-1* plants. Values represent mean $\pm$ half least significant (5%) pairwise difference of six biological replicates. Non-overlapping error bars indicate significant differences at $\alpha = 0.05$.

Table 3.2. Agro-morphological performance of T₂ homozygous, transgene-free WT and *uorfo_{sGGP}-1* plants.

Genotype	Straw biomass (g DW)	No. of panicles per plant	Panicle length (cm)	No. of primary branches per panicle	No. of spikelets per panicle	No. of fertile spikelets per panicle	Spikelet fertility (%)	Grain weight per plant (g)	Thousand grain weight (g)
WT	36.72 ± 1.05	20.33 ± 1.02	16.45 ± 0.57	7.66 ± 0.26	67.66 ± 4.29	46.13 ± 5.08	61.61 ± 5.57	16.73 ± 1.59	29.37 ± 0.78
<i>uorfo_{sGGP}-1</i>	35.33 ± 0.71	27.33 ± 0.76***	13.49 ± 0.28**	6.91 ± 0.08*	53.46 ± 1.62*	43.09 ± 3.21	77.27 ± 5.18	21.82 ± 1.35*	28.55 ± 1.13

Values represent mean ± SEM of six biological replicates. Asterisks indicate significant differences between the WT and *uorfo_{sGGP}* plants (two-sample t-test, * *p*-value ≤0.05, ** *p*-value ≤0.01, *** *p*-value ≤0.001).

Table 3.3. Quality parameters of T₃ homozygous, transgene-free WT and *uorf_{OsGGP-1}* grain.

Genotype	Chalky grains (%)	Grain length (mm)	Grain width (mm)	Length to width ratio	Grain area (mm²)
WT	4.60 ± 1.32	5.32 ± 0.02	2.86 ± 0.01	1.84 ± 0.00	9.87 ± 0.10
<i>uorf_{OsGGP-1}</i>	3.72 ± 1.23	5.38 ± 0.02	2.86 ± 0.01	1.86 ± 0.02*	9.97 ± 0.07

Values represent mean ± SEM of six biological replicates. Asterisks indicate significant differences between the WT and *uorf_{OsGGP-1}* plants (two-sample t-test, * *p*-value ≤ 0.05).

control and salt-stressed *uorfo_{sGGP}-1* mutants relative to WT (Figure 3.8a, b). From the sPSA values, the sPSA AGR and sPSA RGR were calculated at the following DAP intervals: 16-20, 20-24, 24-30, 30-35, and 35-40 (Figure 3.8 and Figure 3.9). As expected, the sPSA RGR declined more rapidly in salt-stressed plants than control plants (Figure 3.8c, d). However, a similar trend in the sPSA RGR of both the control and salt-stressed *uorfo_{sGGP}-1* mutant relative to WT was observed, with the *uorfo_{sGGP}-1* mutant having a significantly higher sPSA RGR for the DAP intervals 16-20, 20-24, and 24-30 relative to WT, but otherwise did not differ significantly for the DAP intervals 30-35 and 35-40.

3.3.8. Grain Fe concentrations of the *uorfo_{sGGP}-1* mutant

No statistically significant differences were detected for Fe concentrations, or the concentration of any of the other elements that were assessed, between T₃ homozygous, transgene-free WT and *uorfo_{sGGP}-1* brown rice with ICP-OES (Table 3.4).

3.3.9. Analysis of potential off-target mutations in the *uorfo_{sGGP}* mutants

Potential off-target (OT) sites in genic regions with 1- or 2-bp mismatches to the pUbi-Cas9-uORF sgRNA in three representative T₂ homozygous, transgene-free WT, *uorfo_{sGGP}-1*, *uorfo_{sGGP}-2*, and *uorfo_{sGGP}-3* plants were sequenced (Table 3.5). A 23-bp deletion was detected for OT-4 in LOC_Os06g02530 in the *uorfo_{sGGP}-1* mutant (Table 3.6). The LOC_Os06g02530 coding sequence is 720 bp in length and encodes a protein of 240 residues in length. The 23-bp deletion in LOC_Os06g02530 in the *uorfo_{sGGP}-1* mutant disrupts the reading frame replacing the final 137 residues of the protein yet keeps the protein length at 240 residues. The protein product for LOC_Os06g02530 is described on the Rice Genome Annotation Project (<http://rice.plantbiology.msu.edu/>) as an expressed protein, however a nucleotide BLAST analysis of the locus revealed that it is predicted to encode a histone H3.v1 protein.

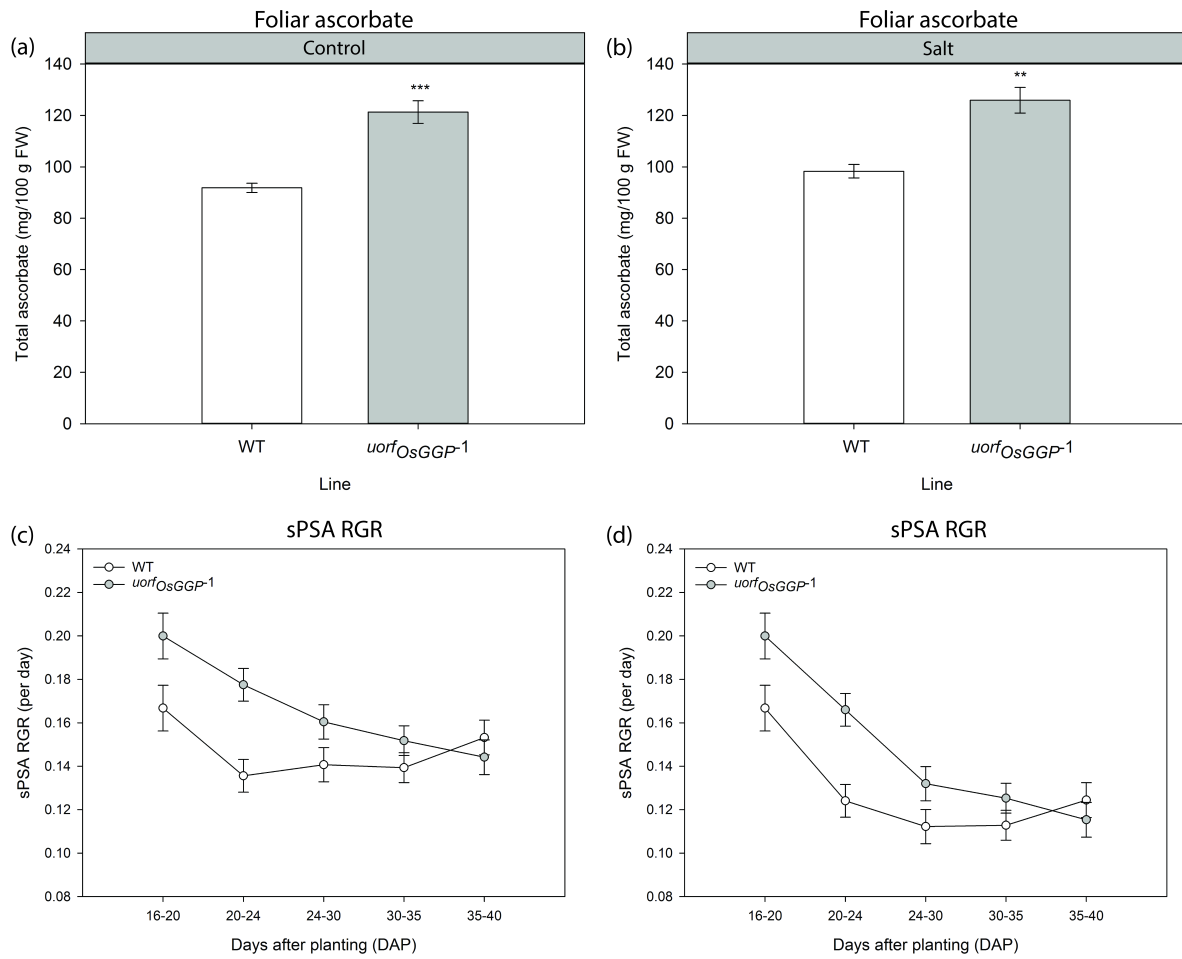


Figure 3.8. Salt stress tolerance assessment of WT and *uorfOsGGP-1* plants. Foliar ascorbate concentrations of (a) control and (b) salt-stressed T₂ homozygous, transgene-free WT and *uorfOsGGP-1* plants at DAP 40. Bars represent mean $\pm$ SEM of six biological replicates. Asterisks indicate statistically significant differences between WT and *uorfOsGGP-1* plants (two-sample t-test; ** p-value ≤ 0.01 ; *** p-value ≤ 0.001). The sPSA RGR of (c) control and (d) salt-stressed T₂ homozygous, transgene-free WT and *uorfOsGGP-1* plants. Salt was applied at DAP 20 and 23. Values represent mean $\pm$ half least significant (5%) pairwise difference of six biological replicates. Non-overlapping error bars indicate significant differences at $\alpha = 0.05$.

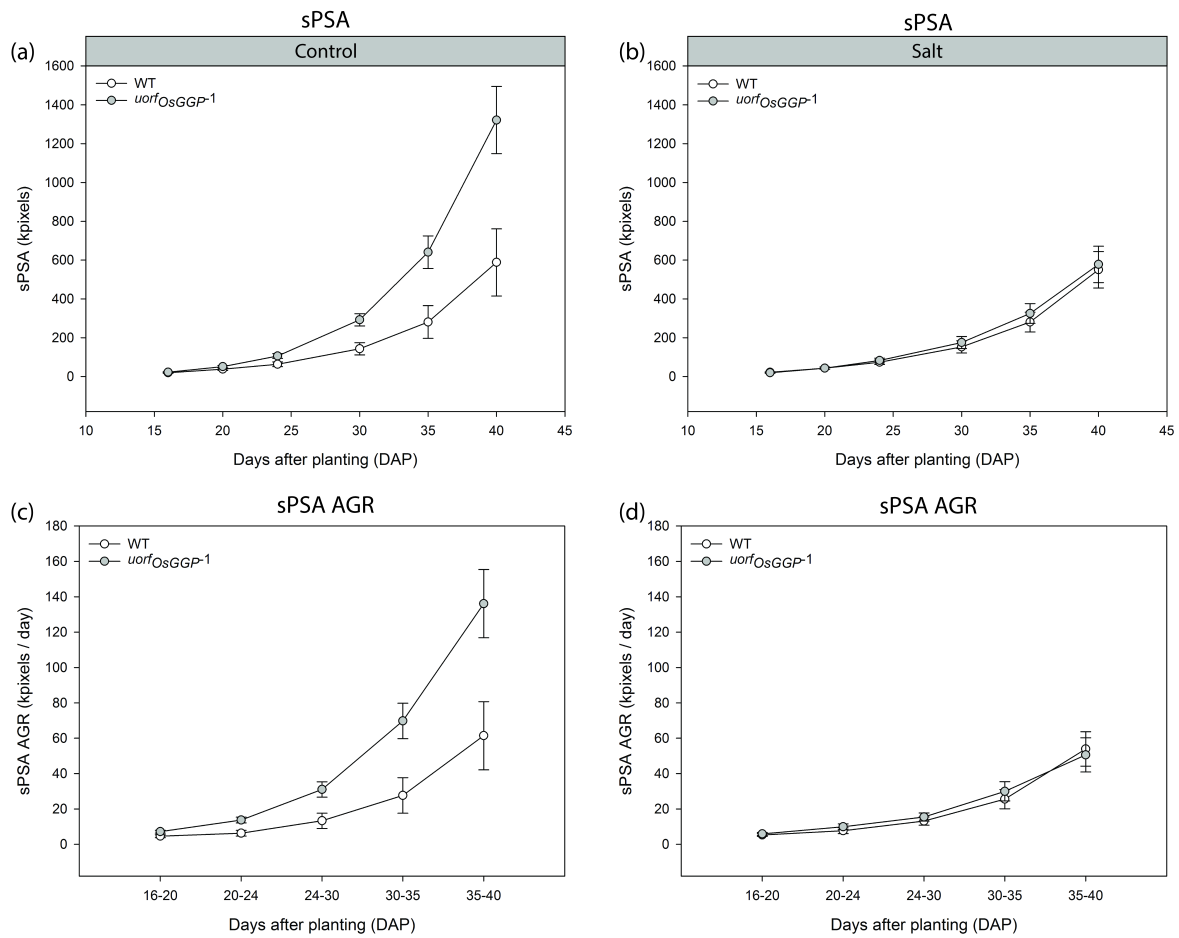


Figure 3.9. The sPSA and sPSA AGR of control and salt-stressed WT and *uorfOsGGP-1* plants. The sPSA of (a) control and (b) salt-stressed T₂ homozygous, transgene-free WT and *uorfOsGGP-1* plants. The sPSA AGR of (c) control and (d) salt-stressed T₂ homozygous, transgene-free WT and *uorfOsGGP-1* plants. Salt was applied at DAP 20 and 23. Values represent mean $\pm$ half least significant (5%) pairwise difference of six biological replicates. Non-overlapping error bars indicate significant differences at $\alpha = 0.05$.

Table 3.4. Elemental concentrations of T₃ homozygous, transgene-free WT and *uorfo_{sGGP}-1* brown rice.

	Fe (mg/kg)	Mg (w/w%)	Na (mg/kg)	K (w/w%)	Ca (w/w%)	P (w/w%)	S (w/w%)	Zn (mg/kg)	Mn (mg/kg)	Cu (mg/kg)
WT	16.52 ± 0.95	0.15 ± 0.00	11.53 ± 1.69	0.39 ± 0.01	0.02 ± 0.00	0.37 ± 0.01	0.12 ± 0.00	36.57 ± 1.75	13.90 ± 0.59	7.50 ± 0.39
<i>uorfo_{sGGP}-1</i>	15.75 ± 0.63	0.15 ± 0.00	10.97 ± 0.72	0.38 ± 0.01	0.02 ± 0.00	0.39 ± 0.01	0.12 ± 0.00	42.15 ± 1.74	14.15 ± 0.56	7.76 ± 0.40

Values represent mean ± SEM of five to six biological replicates. No statistically significant differences were detected between WT and *uorfo_{sGGP}* brown rice (two-sample t-test).

Table 3.5. Potential off-target sites (mismatches ≤ 2 -bp) in genic regions of the pUbi-Cas9-uORF sgRNA.

Target site	Sequence ^a	Locus	Description ^b	Direction	Mismatches (bp)
<i>OsGGP</i> uORF	GACCTCCTCTTCCTCGCCGGCGG	-	-	-	-
OT-1	GACCTCCTCgTCCTCGaCGGCGG	LOC_Os07g09040	peptidyl-prolyl cis-trans isomerase, FKBP-type, putative, expressed	+	2
OT-2	GACgTCCTCgTCCTCGCCGGCGG	LOC_Os02g56950	expressed protein	-	2
OT-3	GACCTCCTCgTCCTCGiCGGAGG	LOC_Os10g35680	acetyltransferase, GNAT family, putative, expressed	+	2
OT-4	GACCTCCTCTTCCTCGiCGGAGG	LOC_Os06g02530	expressed protein	+	1
OT-5	GACCTCCTCgTCCcCGCCGGCGG	LOC_Os09g37650	flavin-containing monooxygenase family protein, putative, expressed	+	2
OT-6	GACCTCCTCgTCgTCGCCGGCGG	LOC_Os11g13840	AP2 domain containing protein, expressed	-	2

^aThe mismatched bases in the potential off-target sites and the protospacer-adjacent motif are shown in lower case letters and bold, respectively. ^bBased on the Rice Genome Annotation Project (<http://rice.plantbiology.msu.edu/>).

Table 3.6. Summary of off-target mutations detected in three representative T₂ homozygous, transgene-free WT, *uorfo_{sGGP}-1*, *uorfo_{sGGP}-2*, and *uorfo_{sGGP}-3* plants with Sanger sequencing.

Line	Replicate	OT-1 (2 bp)	OT-2 (2 bp)	OT-3 (2 bp)	OT-4 (1 bp)	OT-5 (2 bp)	OT-6 (2 bp)
WT	1	-	-	-	-	-	-
	2	-	-	-	-	-	-
	3	-	-	-	-	-	-
<i>uorfo_{sGGP}-1</i>	1	-	-	-	-23	-	-
	2	-	-	-	-23	-	-
	3	-	-	-	-23	-	-
<i>uorfo_{sGGP}-2</i>	1	-	-	-	-	-	-
	2	-	-	-	-	-	-
	3	-	-	-	-	-	-
<i>uorfo_{sGGP}-3</i>	1	-	-	-	-	-	-
	2	-	-	-	-	-	-
	3	-	-	-	-	-	-

The number of mismatches between the sgRNA and off-target sites are presented in brackets. -, no mutation found.

3.4. Discussion

3.4.1. The CRISPR/Cas9 complex was active and generated new mutation as the T₀ rice transformants developed

The mutation frequencies were strikingly low for the independent T₀ p35S-Cas9-uORF (0%) and pUbi-Cas9-uORF (1.64%) transformants when DNA extracted from shoots regenerating from hygromycin-resistant callus was used to screen for mutations, relative to what has been reported in the literature for the CRISPR/Cas9 genome editing system in rice (typically around 80-85%) [173]. However, when DNA extracted from T₀ pUbi-Cas9-uORF transformants transferred to Jiffy peat pellets for greenhouse acclimation was used to screen for mutations, the mutation frequency increased to 59.18%. This result suggests that the CRISPR/Cas9 complex was active and could generate new mutations as the T₀ rice transformants developed. Similar results have been reported with the CRISPR/Cas9 genome editing system in rice, with new mutations reported to be induced in T₁ and T₂ progeny possessing the CRISPR/Cas9 complex [243, 244]. Taken together, these results highlight the importance of timing when screening for CRISPR/Cas9-induced mutations in plants transformed with CRISPR/Cas9 vectors. Based on the mutation frequencies observed in this chapter, it is recommended that T₁ progeny should be screened for CRISPR/Cas9-induced mutations to provide the CRISPR/Cas9 complex the maximal amount of time to induce mutations in the developing T₀ transformants, as well as to confirm that the mutations are heritable. Given that the CRISPR/Cas9 complex was active in developing plants, this also complicates comparisons between sgRNA mutation frequencies that are reported in the literature, as different research groups may screen for CRISPR/Cas9-induced mutations at different stages of development. It is therefore also recommended that the developmental stage at which CRISPR/Cas9-induced mutations are screened are reported. The CRISPR/Cas9-induced mutation types and frequencies in the 31 plants with mutations detected in the *OsGGP* uORF transformed with the pUbi-Cas9-uORF T-DNA are similar to those reported, i.e. a high frequency of 1-bp insertions, followed by 1-bp deletions, and a high frequency of adenine and thymine insertions relative to guanine and cytosine insertions [243].

3.4.2. CRISPR/Cas9-induced mutations increased foliar ascorbate concentrations without affecting transcript levels of the *OsGGP* gene, but did not affect ascorbate concentrations in brown rice or germinated brown rice

Foliar ascorbate concentrations were significantly increased 1.2- to 1.3-fold in all three T₂ homozygous *uorf_{OsGGP}* mutants relative to WT, indicating that all three mutations could disrupt the mechanism of the *OsGGP* uORF. Transcript levels of the *OsGGP* gene were not affected in all three T₂ homozygous *uorf_{OsGGP}* mutants relative to WT, proving that the increased ascorbate concentrations were not due to increased transcription of the *OsGGP* gene, but likely due to alleviating *GGP* uORF-induced ribosomal stalling and enhanced translation of the *OsGGP* mORF. Additional experiments utilizing the dual-luciferase reporter system will be necessary to confirm the effects of the mutated *OsGGP* uORFs on the translation of downstream mORFs, and should be the focus of future research efforts [153, 154, 163, 239]. Foliar ascorbate concentrations were also increased 1.2- to 1.7-fold in a range of T₁ heterozygous, homozygous, and bi-allelic, transgene-free *uorf_{OsGGP}-2* and *uorf_{OsGGP}-3* mutants relative to WT, including a significant 1.7-fold increase in heterozygous *uorf_{OsGGP}-3* mutants relative to WT, proving that mutations in the *GGP* uORF are dominant. Ascorbate concentrations were negligible in both WT and *uorf_{OsGGP}-1* brown rice, likely due to the decline of ascorbate levels during grain maturation reported in cereals [54, 55]. Further understanding this decline in ascorbate concentrations in the maturing grains of cereals may help identify strategies to reduce this decline, and therefore present avenues to biofortify cereals with ascorbate, and ought to be the subject of future research. Ascorbate concentrations were also either reduced or not significantly different in germinating *uorf_{OsGGP}-1* brown rice relative to WT. These findings suggest that disrupting the function of the *OsGGP* uORF may not be a viable strategy to increase ascorbate concentrations in brown rice or germinated brown rice for the purpose of improving the dietary intake of ascorbate in humans.

3.4.3. Insights gained into the function of the *GGP* uORF peptide

Generating mutations that disrupt the reading frame or delete residues in the *GGP* uORF peptide offers insight into the function of the peptide that cannot otherwise be gained when the start codon is disrupted, as was the case for the *GGP* uORFs that were recently edited in *Arabidopsis*, lettuce, and tomato [153, 154]. The *uorf_{OsGGP}-1* mutant, which replaces the five residues (G⁵⁷GGRL⁶¹) at the C-terminal end of the peptide, and the *uorf_{OsGGP}-2* mutant, which deletes two residues (L⁵⁵A⁵⁶)

and also replaces the five residues (G⁵⁷GGRL⁶¹) at the C-terminal end of the peptide, both significantly increased foliar ascorbate concentrations, suggesting that these several residues toward the C-terminal end of the peptide have an active role in ribosome stalling. This is consistent with the L⁵⁵AGGG⁵⁹ string of residues being highly conserved in a wide range of species [163]. The *uorf_{OsGGP}-3* mutant, which deletes 21 highly conserved residues (G³⁷GRGALPSAGGSPDLLFLAG⁵⁷) in a wide range of species [163], resulted in the highest fold-increase in ascorbate concentrations observed for any of the *uorf_{OsGGP}* mutants, and is consistent with it being the most significant disruption to the peptide sequence. Taken together, this data suggests that these highly conserved residues toward the C-terminal end of the peptide are important for ribosome stalling, and is in agreement with ribosomes being most strongly associated (i.e. stalled) toward the end of the *GGP* uORF [163]. Positively charged residues in the *GGP* uORF peptide have been implicated in ribosome stalling, as positively charged residues have been reported to reduce ribosomal velocity due to interacting with the negatively charged ribosomal exit tunnel [163, 245]. However, both the *uorf_{OsGGP}-1* and *uorf_{OsGGP}-2* mutants, which both have four more positively charged residues in the peptide relative to WT, increased ascorbate concentrations, indicating that it is not the frequency of positively charged residues in the peptide, but rather the sequence of residues that is important for ribosome stalling. How ascorbate interacts with the *GGP* uORF mechanism to influence translation of the *GGP* uORF or mORF still remains an area for future research. One hypothesis is that the *GGP* uORF peptide and the ribosomal exit tunnel wall may create an ascorbate binding pocket whereby ascorbate inactivates the peptidyl transferase centre, causing the ribosome to stall, and thereby preventing downstream translation of the *GGP* mORF, as has been reported for the TnaC and ermCL ribosomal stalling systems in *Escherichia coli* [220, 246-249]. This hypothesis could be tested, for example, by employing near-atomic cryo-electron microscopy to resolve the structure of the stalled *GGP* uORF peptide-ribosome complex by the addition of free ascorbate [247-249]. Another hypothesis is that the *GGP* uORF mechanism may respond to the redox state of the cell whereby the phosphorylation status of eukaryotic initiation factors are altered, which in turn influences whether ribosomes preferentially initiate from the *GGP* uORF with a weaker Kozak sequence or the *GGP* mORF with a stronger Kozak sequence, as has been reported for uORF-containing stress genes in mammalian cells [250]. This hypothesis could be tested, for example, by co-infiltrating genes encoding redox proteins (e.g. superoxide dismutases, catalases, or thioredoxins) or genes encoding proteins that

increase the biosynthesis of antioxidants (e.g. carotenoids, tocopherols, or flavonoids) with a luciferase reporter system construct containing the *GGP* uORF upstream of the luciferase reporter gene into *N. benthamiana* to determine whether the redox state of the cell influences luciferase translation independent of ascorbate concentrations [163, 251].

3.4.4. Increased ascorbate levels in the *uorf_{OsGGP}-1* mutant were associated with increased grain weight per plant

Automated imaging revealed that the WT and *uorf_{OsGGP}-1* mutant did not differ significantly regarding the accumulation of shoot biomass, except for a slight increase in the sPSA RGR for the DAP interval 16-30. An assessment of the agro-morphological performance of the *uorf_{OsGGP}-1* mutant revealed that increased ascorbate levels were associated with a significant 1.3-fold increase in grain weight per plant relative to WT. The increased grain weight per plant may be attributed to the 1.3-fold increase in the number of panicles per plant, as well as a trend towards increased spikelet fertility in the *uorf_{OsGGP}-1* mutant (77%) relative to WT (62%). This agro-morphological performance data agrees with previous studies that have reported that increased ascorbate concentrations in rice has been associated with increased grain yield. One study, for example, reported that increased ascorbate concentrations in rice via constitutive overexpression of the rice *dehydroascorbate reductase 1* (*OsDHAR1*) gene was associated with increased grain yield when grown under paddy field conditions, in part due to an increase in the number of panicles and spikelet fertility [138]. Another study reported that increased ascorbate concentrations in rice plants via the constitutive overexpression of the rice *L-galactono-1,4-lactone dehydrogenase* (*OsL-GalLDH*) gene was associated with increased spikelet fertility [79]. The enhanced agro-morphological performance reported for these ascorbate-enriched rice plants was largely attributed to enhanced photosynthetic performance and antioxidant activity [79, 138]. Investigating whether the *uorf_{OsGGP}* mutants have enhanced photosynthetic performance or antioxidant activity should therefore also be the subject of future research. One major limitation of this experiment is that only one *uorf_{OsGGP}* mutant line was studied over the span of only one generation. Therefore, additional experiments investigating the agro-morphological performance of all three *uorf_{OsGGP}* mutants over multiple generations, including under both greenhouse and paddy field conditions, is warranted to determine whether the increased grain weight per plant in the *uorf_{OsGGP}-1* mutant is reproducible.

3.4.5. Increased ascorbate levels in the *uorfo_{OsGGP}-1* mutant did not affect salt tolerance in terms of shoot biomass accumulation

Previous studies have reported that non-destructive measurements such as sPSA is a strong predictor of shoot biomass and that the sPSA RGR can be used effectively to screen for salt tolerance in rice [194-197]. However, despite a 1.3-fold increase in foliar ascorbate concentrations in the *uorfo_{OsGGP}-1* mutant relative to WT under both control and salt stress conditions, a similar trend in the sPSA RGR was observed for both the control and salt-stressed *uorfo_{OsGGP}-1* mutants relative to WT. This result suggests that elevated ascorbate levels from CRISPR/Cas9-induced mutations in the *OsGGP* uORF in rice does not affect salt tolerance in terms of the accumulation of shoot biomass in response to a final concentration of 90 mM NaCl at the vegetative growth phase. This result contradicts many studies that have reported enhanced salt tolerance in plants with increased expression of ascorbate biosynthetic genes [64, 68, 70, 80, 89-91, 93, 99, 118, 187, 198, 252]. A recent study, however, has reported that increased expression of the kiwifruit (*Actinidia chinensis* L.) *AcGGP* gene in rice enhanced salt tolerance only in terms of visible symptoms and lipid peroxidation, but not shoot biomass [187]. Therefore, investigating oxidative stress markers such as lipid peroxidation in salt-stressed *uorfo_{OsGGP}* mutants may be required to determine whether increased ascorbate levels due to CRISPR/Cas9-induced mutations in the *OsGGP* uORF in rice ameliorates salt-induced oxidative stress.

3.4.6. Increased ascorbate levels in the *uorfo_{OsGGP}-1* mutant did not affect grain Fe levels

Ascorbate has been reported to play an important role in the reduction and transport of Fe in the developing seeds of *Arabidopsis* and pea (*Pisum sativum* L.) plants [7, 218]. Developing embryos were reported to secrete high amounts of ascorbate to reduce Fe^{3+} from citrate-malate complexes allowing for the uptake of Fe^{2+} by the embryo [7, 218]. This mechanism was validated genetically with the ascorbate-deficient *Arabidopsis ggp1 (vtc2)* and *ggp2 (vtc5)* mutants which had lower Fe reduction activity and lower seed Fe levels relative to WT [7]. Plants with increased ascorbate concentrations may therefore have increased Fe reduction activity and seed Fe levels. However, increased ascorbate levels in the *uorfo_{OsGGP}-1* mutant did not affect the concentration of Fe in brown rice, indicating that increased ascorbate levels due to CRISPR/Cas9-induced mutations in the

OsGGP uORF does not facilitate a greater loading of Fe into the developing rice grain. It is worth noting, however, that ascorbate has only been documented to play a role in the reduction and transport of Fe in the developing seeds of *Arabidopsis* and pea plants, which are both non-graminaceous species and therefore this mechanism may not be present in graminaceous species, such as rice. Whether increased ascorbate levels can facilitate a greater loading of Fe into seeds should therefore be tested in ascorbate-enriched *Arabidopsis* and pea plants, where this mechanism has been genetically validated. Further, investigating whether grain Fe levels are altered in ascorbate-deficient graminaceous mutants, such as rice, may shed light on the presence of this mechanism in graminaceous species.

3.4.7. An off-target 23-bp deletion was detected in the *uorf_{OsGGP-1}* mutant

A 23-bp deletion was detected in the OT-4 site (LOC_Os06g02530) in the *uorf_{OsGGP-1}* mutant with Sanger sequencing. The OT-4 site has a 1-bp mismatch (cytosine to thymine) 4-bps upstream of the protospacer-adjacent motif of the pUbi-Cas9-uORF sgRNA. While the occurrence of off-target mutations are seldom reported in rice, previous studies have detected them, particularly when sgRNAs with only 1-bp mismatches are used [253, 254]. Future research in the field of CRISPR/Cas9 genome editing in rice should therefore utilize high-fidelity sgRNAs, with at least 2-bp mismatches to off-target sites to reduce the occurrence of off-target mutations. Based on a nucleotide BLAST analysis, LOC_Os06g02530 is predicted to encode a histone H3.v1 protein. Histone proteins are core components of the nucleosome—the basic structural unit of DNA packaging in eukaryotic cells—and play a key role in gene regulation and epigenetic mechanisms [255]. The 23-bp deletion detected in LOC_Os06g02530 in the *uorf_{OsGGP-1}* mutant disrupts the reading frame of the coding sequence replacing the final 137 residues of the 240-residue long protein. Mutations disrupting the deposition of histones or the methylation of histones have reported to impair plant growth and development [256-258], however no detrimental effects on plant growth and development in the *uorf_{OsGGP-1}* mutant were observed. It is therefore possible that this locus is partially redundant with other loci predicted to encode histone H3 proteins in the rice genome, and that mutations in this locus may be neutral, having neither a beneficial nor detrimental effect on rice plant growth and development. This off-target 23-bp deletion was only detected after the fact of conducting the experiments pertaining to the plant growth and development, agro-morphological performance, salt tolerance, and grain Fe levels of the

uorf_{OsGGP-1} mutant and therefore may be a confounding factor in the interpretation of these results presented in this chapter. The off-target 23-bp deletion in the *uorf_{OsGGP-1}* mutant will need to be eliminated by backcrossing to determine whether this off-target mutation is responsible for any of the phenotypes observed.

3.5. Conclusion

The results of this chapter found that CRISPR/Cas9 induced mutations in the *OsGGP* uORF significantly increased foliar ascorbate levels but did not affect ascorbate levels in brown rice or germinating brown rice. The CRISPR/Cas9 induced mutations in the *OsGGP* uORF did not affect transcript levels of the *OsGGP* gene, indicating that the increased ascorbate levels were due to enhanced translation of the *GGP* mORF. Increased ascorbate levels in the *uorf_{OsGGP}-1* mutant was associated with a significant increase in grain weight per plant but did not affect salt stress tolerance or grain Fe levels. An off-target 23-bp deletion was detected in the *uorf_{OsGGP}-1* mutant that will need to be eliminated by backcrossing to determine whether this mutation was responsible for any of the phenotypes observed for the *uorf_{OsGGP}-1* mutant. Disrupting the function of the *GGP* uORF with the CRISPR/Cas9 genome editing system represents a promising transgene-free strategy to increase ascorbate biosynthesis, and possibly grain weight per plant, not only in rice, but other crops.

Chapter 4.

Genome-Wide Identification and Characterization of the *GDP-L-Galactose Phosphorylase* Gene Family in Bread Wheat

4.1. Abstract

Background: Ascorbate is a powerful antioxidant in plants and an essential micronutrient for humans. The *GDP-L-galactose phosphorylase* (*GGP*) gene encodes the rate-limiting enzyme of the L-galactose pathway—the dominant ascorbate biosynthetic pathway in plants—and is a promising gene candidate for increasing ascorbate in crops. In addition to transcriptional regulation, *GGP* production is regulated at the translational level through an upstream open reading frame (uORF) in the long 5'-untranslated region (5'UTR). The *GGP* genes have yet to be identified in bread wheat (*Triticum aestivum* L.), one of the most important food grain sources for humans.

Results: Bread wheat chromosomal groups 4 and 5 were found to each contain three homoeologous *TaGGP* genes on the A, B, and D subgenomes (*TaGGP2-A/B/D* and *TaGGP1-A/B/D*, respectively) and a highly conserved uORF was present in the long 5'UTR of all six genes. Phylogenetic analyses demonstrated that the *TaGGP* genes separate into two distinct groups and identified a duplication event of the *GGP* gene in the ancestor of the *Brachypodium*/Triticeae lineage. A microsynteny analysis revealed that the *TaGGP1* and *TaGGP2* subchromosomal regions have no shared synteny suggesting that *TaGGP2* may have been duplicated via a transposable element. The two groups of *TaGGP* genes have distinct expression patterns with the *TaGGP1* homoeologs broadly expressed across different tissues and developmental stages and the *TaGGP2* homoeologs highly expressed in anthers. Transient transformation of the *TaGGP* coding sequences in *Nicotiana benthamiana* leaf tissue increased ascorbate concentrations more than five-fold, confirming their functional role in ascorbate biosynthesis *in planta*.

Conclusions: We have identified six *TaGGP* genes in the bread wheat genome, each with a highly conserved uORF. Phylogenetic and microsynteny analyses highlight that a transposable element may have been responsible for the duplication and specialized expression of *GGP2* in anthers in the *Brachypodium*/Triticeae lineage. Transient transformation of the *TaGGP* coding sequences in *N. benthamiana* demonstrated their activity *in planta*. The six *TaGGP* genes and uORFs identified in this study provide a valuable genetic resource for increasing ascorbate concentrations in bread wheat.

4.2. Introduction

Ascorbate, also known as vitamin C, is the most abundant water-soluble antioxidant in plant cells and plays important roles in photosynthetic function and stress tolerance [23, 259]. Ascorbate can mitigate the harmful effects of reactive oxygen species (ROS) produced by normal or stressed cellular metabolism, either directly as an ROS scavenger or indirectly as a substrate for the enzyme ascorbate peroxidase [21]. Increased biosynthesis of ascorbate in model and crop species has demonstrated that even minor increases in ascorbate can confer enhanced tolerance to a broad range of abiotic stresses [192, 260]. Additional roles for ascorbate within plant cells include: serving as an enzymatic cofactor [261] and precursor for organic acids [34], as well as influencing the cell cycle [36], cell wall expansion [262], and flowering time and senescence [38].

In humans, ascorbate is an essential micronutrient due to loss-of-function mutations within the *L-gulonolactone oxidase* gene which encodes the enzyme responsible for catalysing the last step in vertebrate ascorbate biosynthesis [39]. Beyond its well-established role in preventing the disease scurvy, ascorbate plays key roles in many physiological processes important for human health, including promoting and regulating the uptake of dietary iron in the human digestive process [9] and serving as a cofactor for enzymes involved in epigenetic programming [42, 192]. Although ascorbate deficiency is considered a historical disease, suboptimal intake of ascorbate is currently present in both developing and developed countries [48], and the re-emergence of scurvy has been documented in Australia due to poor dietary habits [51].

Four pathways for ascorbate biosynthesis have been proposed in plants: the L-galactose, L-gulose, *myo*-inositol, and D-galacturonate pathways. Despite evidence for each of these pathways in plants, the L-galactose pathway, which converts D-fructose-6-P to ascorbate via eight enzymatic steps, is the dominant pathway leading to ascorbate biosynthesis in *Arabidopsis thaliana* [12, 13, 107], tomato (*Solanum lycopersicum* L.) [15], rice (*Oryza sativa* L.) [14], and green algae (*Chlamydomonas reinhardtii*) [16]. The final gene to be identified in the L-galactose pathway, *GDP-L-galactose phosphorylase* (GGP, also known as *VTC2*), was discovered and functionally characterized in *Arabidopsis* in 2007 [12, 108, 112]. The GGP enzyme catalyzes the conversion of GDP-L-galactose to L-galactose-1-P and represents the first committed step toward ascorbate biosynthesis. Several studies have now identified the GGP enzyme as the rate-limiting step in the L-galactose pathway [62, 64, 71, 108] and as the key regulatory enzyme in ascorbate

biosynthesis [17]. Overexpression of *GGP* genes has increased ascorbate concentrations by up to 4.2-fold in *Arabidopsis* [62, 71], 3.1-fold in potato (*Solanum tuberosum* L.) [73], 2.1-fold in strawberry (*Fragaria x ananassa*) [73], 6.2-fold in tomato [67, 73], 1.4-fold in tobacco (*Nicotiana tabacum* L.) [74], and 2.6-fold in rice [64, 187].

A highly conserved upstream open reading frame (uORF)—a class of small ORFs located upstream of protein-coding major ORFs (mORFs) in the 5'-untranslated region (5'UTR) of mRNAs—of the *GGP* gene has been shown to regulate *GGP* translation and ascorbate concentrations [153, 154, 163, 219]. Although uORFs are abundant in plant genomes with up to 60% of genes estimated to contain an uORF, less than 1% are conserved between plant species [263, 264]. The translational regulation of mORFs by highly conserved uORFs in response to cellular metabolite levels has been documented in several plant studies. For example, uORFs have been shown to regulate translation of the *Arabidopsis* *S-adenosylmethionine decarboxylase* and *polyamine oxidase 2* genes in response to polyamines [221-224], the *Arabidopsis* S1 class *basic-leucine zipper* and tomato *ornithine decarboxylase* genes in response to sucrose [231-233], the *Arabidopsis* *phosphoethanolamine N-methyltransferase* gene in response to phosphocholine [236], and the *Arabidopsis* *heat shock transcription factor B1* gene in response to galactinol [238]. In the case of the *GGP* uORF, the encoded peptide which is predicted to initiate from a non-canonical AUC or ACG start-codon is proposed to cause ribosomal stalling, thereby preventing translation of the downstream *GGP* mORF under high ascorbate concentrations, whereas under low ascorbate concentrations the uORF is skipped and the *GGP* mORF translated [163]. The precise mechanism of how ascorbate may influence ribosomal initiation at the uORF or mORF has yet to be elucidated. Disruption of the *GGP* uORF through mutation of key residues was shown to interfere with ascorbate feedback regulation of translation and increase ascorbate concentration when transiently transformed in *Nicotiana benthamiana* [163]. Recently, CRISPR/Cas9 genome editing has been used to disrupt the *GGP* uORF in *Arabidopsis*, lettuce (*Lactuca sativa*), and tomato to increase ascorbate concentrations and enhance oxidative stress tolerance [153, 154].

Despite the importance of the *GGP* gene in ascorbate biosynthesis and regulation, characterisation of *GGP* genes within graminaceous species has been limited to rice and maize (*Zea mays* L.). Both rice and maize have one copy of the *GGP* gene present in their genomes, located on chromosomes Os11 and Zm6, respectively. Complete knock-out of the *OsGGP* gene in

rice severely reduced foliar ascorbate concentrations by 80% and was correlated with reduced photosynthetic efficiency, lower biomass production, and reduced tolerance to ozone stress and zinc deficiency [14]. In maize, the *ZmGGP* gene is expressed during endosperm development and expression levels have been shown to be genotype-dependent [265].

Wheat is the world's most cultivated crop in terms of land area and accounts for one fifth of the calories consumed by humans [266]. Abiotic stresses such as drought, salinity, and high temperature represent the major limiting factors in global wheat productivity [267]. Increased ascorbate concentrations in wheat could help to mitigate yield losses associated with these abiotic stresses. Like many other cereals, wheat contains negligible levels of ascorbate in the mature grain [52], therefore, ascorbate enrichment of wheat grain could also provide a novel means of improving human dietary intakes of this essential vitamin.

Here we expand upon our knowledge of *GGP* genes in graminaceous species by describing the *TaGGP* gene family in bread wheat. We describe the chromosomal location, confirm the presence of the highly conserved uORFs, present phylogenetic relationship, gene synteny, and tissue and developmental expression patterns, and demonstrate activity of the *TaGGP* coding sequences by transient transformation in *N. benthamiana*.

4.3. Results

4.3.1. Six *TaGGP* genes were identified on bread wheat chromosomal groups 4 and 5

A total of six *TaGGP* genes were identified within the bread wheat genome and their genomic locations were determined. Chromosomal groups 4 and 5 each contain three homoeologous *TaGGP* genes located on the A, B, and D subgenomes (*TaGGP2-A/TaGGP2-B/TaGGP2-D* and *TaGGP1-A/TaGGP1-B/TaGGP1-D*, respectively). All six genes are located on the short arm of their respective chromosomal group except *TaGGP2-A* which is located on the long arm. In addition to the six *TaGGP* genes, we identified two *HvGGP* genes within the barley (*Hordeum vulgare*) genome on chromosomes 4 and 5 (*HvGGP2* and *HvGGP1*, respectively), two *BdGGP* genes within the *Brachypodium distachyon* (*Brachypodium* hereafter) genome on chromosome 4 (*BdGGP1* and *BdGGP2*), two *AetGGP* genes within the *Aegilops tauschii* genome on chromosome 4 and 5 (*AetGGP2* and *AetGGP1*, respectively), and one *SbGGP* gene within the *Sorghum bicolor* L. genome on chromosome 8.

The six *TaGGP* genomic sequences range in length from 2922 bp to 4163 bp mainly due to differences in intron sequence length (Figure 4.1) and are similar in size and structure to other graminaceous *GGP* genes (Figure S4.1). Genomic sequence comparisons of the six *TaGGP* genes revealed that they share between 36.0% to 94.9% identity. The *TaGGP1* and *TaGGP2* proteins are 430 and 431 amino acids in length, respectively, and sequence comparisons of the six *TaGGP* proteins revealed that they share between 78.1% to 99.3% identity. Alignment of the amino acid sequence of the wheat, barley, rice, maize, sorghum, *Brachypodium*, and *A. tauschii* GGP proteins identified highly conserved regions within the proteins, as well as the conserved histidine triad (HIT) motif of the HIT protein superfamily, the conserved KKRP nuclear localization signal (NLS), and 13 residues that distinguish GGP2 proteins from GGP1 proteins within these species (Figure 4.2).

4.3.2. Each *TaGGP* gene contains a highly conserved uORF

For all six *TaGGP* genes, a highly conserved uORF was identified in the long 5'UTR of the mRNA initiating either from the non-canonical AUC (isoleucine) or ACG (threonine) start-codon. The uORF peptides of the *TaGGP1* homoeologs and *TaGGP2* homoeologs are 56 and 47

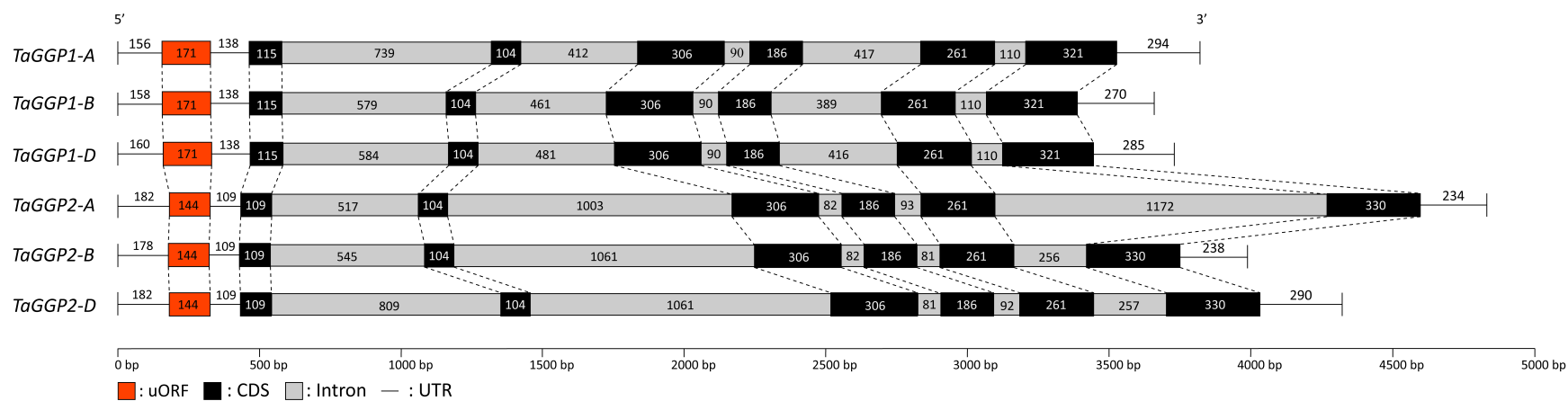


Figure 4.1. Gene structure of the *TaGGP* genes from bread wheat cv. Chinese Spring. The uORF (orange box), coding sequence (CDS, black box), introns (grey box), and UTR (lines) of the *TaGGP* genes are depicted and the length (bp) of each section indicated.

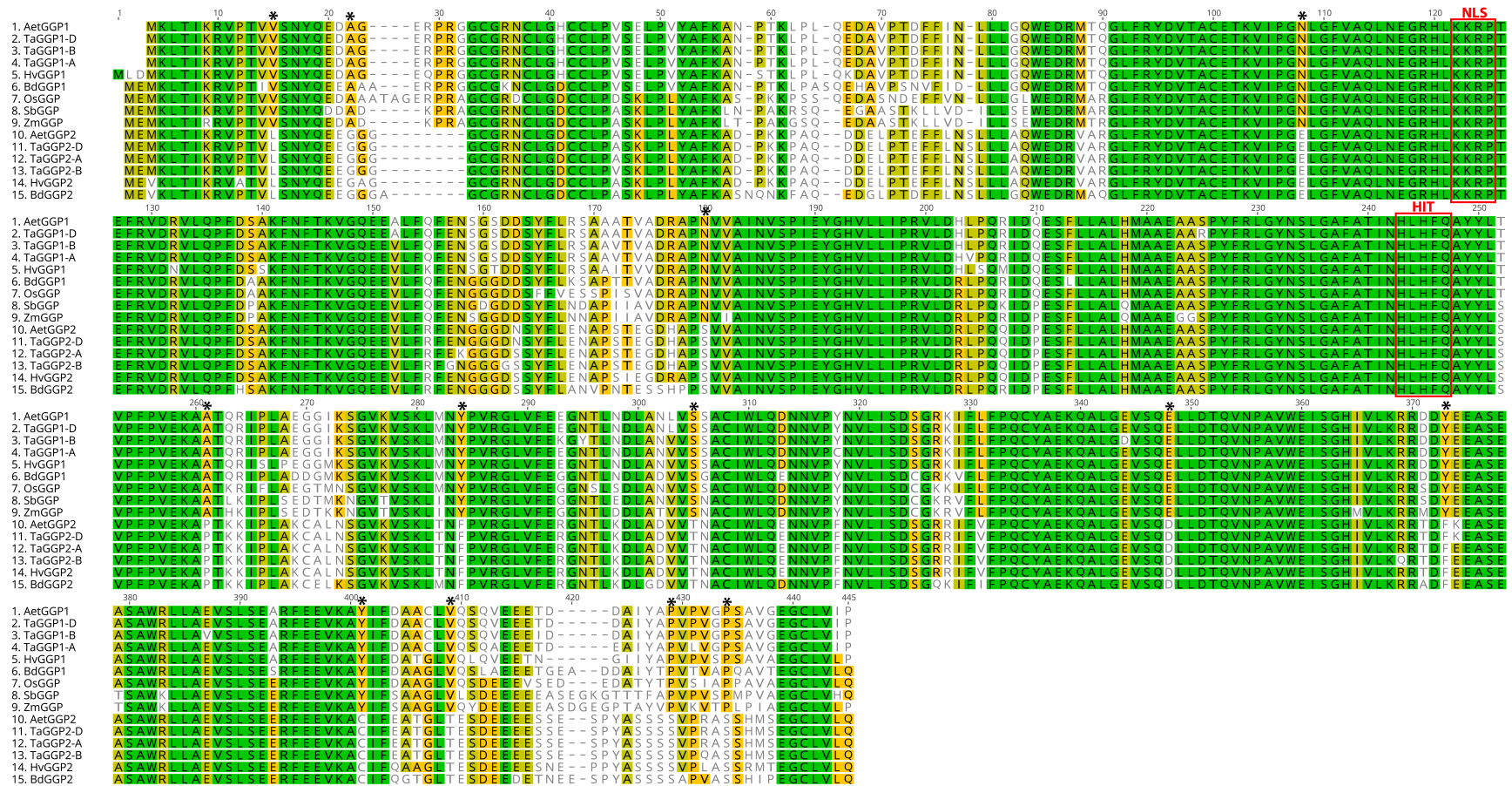


Figure 4.2. Amino acid sequence alignment of GGP proteins from a range of graminaceous species. Green, olive green, yellow and white background colour represents 100%, 80 to 100%, 60 to 80%, and less than 60% conservation of amino acids between species, respectively. The HIT motif (HφHφQ, where φ is a hydrophobic amino acid) of the HIT protein superfamily and the KKRPNLS are outlined in red. The 13 residues across the protein that distinguish GGP1 proteins from GGP2 proteins within the graminaceous species are indicated by an asterisk. The prefixes for the graminaceous species are as follows: Aet is *Aegilops tauschii*; Bd is *Brachypodium distachyon*; Hv is *Hordeum vulgare*; Os is *Oryza sativa*; Sb is *Sorghum bicolor*; Ta is *Triticum aestivum*; and Zm is *Zea mays*.

amino acids in length, respectively, and sequence comparisons of the six *TaGGP* uORF peptides revealed that they share between 67.9% to 100% identity. Alignment of the amino acid sequence of the *GGP* uORF peptides from a range of graminaceous and non-graminaceous species identified highly conserved regions within the peptides, as well as a lysine at the first residue in rice instead of a highly conserved isoleucine in all other species, a truncation of 11 residues in the *TaGGP2*, *HvGGP2*, *BdGGP2*, and *AetGGP2* uORF peptides relative to the *TaGGP1*, *HvGGP1*, *BdGGP1*, and *AetGGP1* uORF peptides, and a single amino acid change that distinguishes graminaceous species (alanine) from non-graminaceous species (glutamic acid) (Figure 4.3).

4.3.3. Graminaceous GGP proteins separate into two groups

Phylogenetic analysis of the GGP protein sequences from a range of graminaceous and non-graminaceous plant species identified two clades of GGP proteins, belonging to either graminaceous or non-graminaceous species (Figure 4.4). The graminaceous clade of GGP proteins further separates into two groups, with the *TaGGP1* proteins forming one group with *HvGGP1*, *BdGGP1*, *AetGGP1*, *OsGGP*, *ZmGGP*, and *SbGGP*, and the *TaGGP2* proteins forming a second group with *HvGGP2*, *BdGGP2*, and *AetGGP2*. This classification of group 1 and group 2 graminaceous GGP proteins is further supported by a phylogenetic analysis of the graminaceous *GGP* coding sequences which clearly demonstrates that the *TaGGP1* homoeologs, *HvGGP1*, *BdGGP1*, and *AetGGP1* are more closely related to the *OsGGP*, *ZmGGP*, and *SbGGP* in terms of evolutionary distance than the *TaGGP2* homoeologs, *HvGGP2*, *BdGGP2*, and *AetGGP2* (Figure S4.2). Phylogenetic analysis of the *GGP* uORF peptide sequences revealed similar relationships to those observed for the GGP protein sequences (Figure S4.3).

4.3.4. The *TaGGP1* and *TaGGP2* subchromosomal regions do not share synteny

Selection of ten genes upstream and downstream of the respective *TaGGP* genes on each of the six chromosomes, referred to as the *TaGGP* subchromosomal regions (SR), were analysed for gene density and synteny (Figure 4.5 and Table S4.1). The physical length of the *TaGGP1-A*, *TaGGP1-B*, and *TaGGP1-D* SRs were 18, 14.8, and 8.7 Mbp, respectively, and the *TaGGP2-A*, *TaGGP2-B*, and *TaGGP2-D* SRs were 3.4, 4.8, and 3 Mbp, respectively. No shared synteny was found between the *TaGGP1* and *TaGGP2* SRs on chromosomal groups 4 and 5 except for the

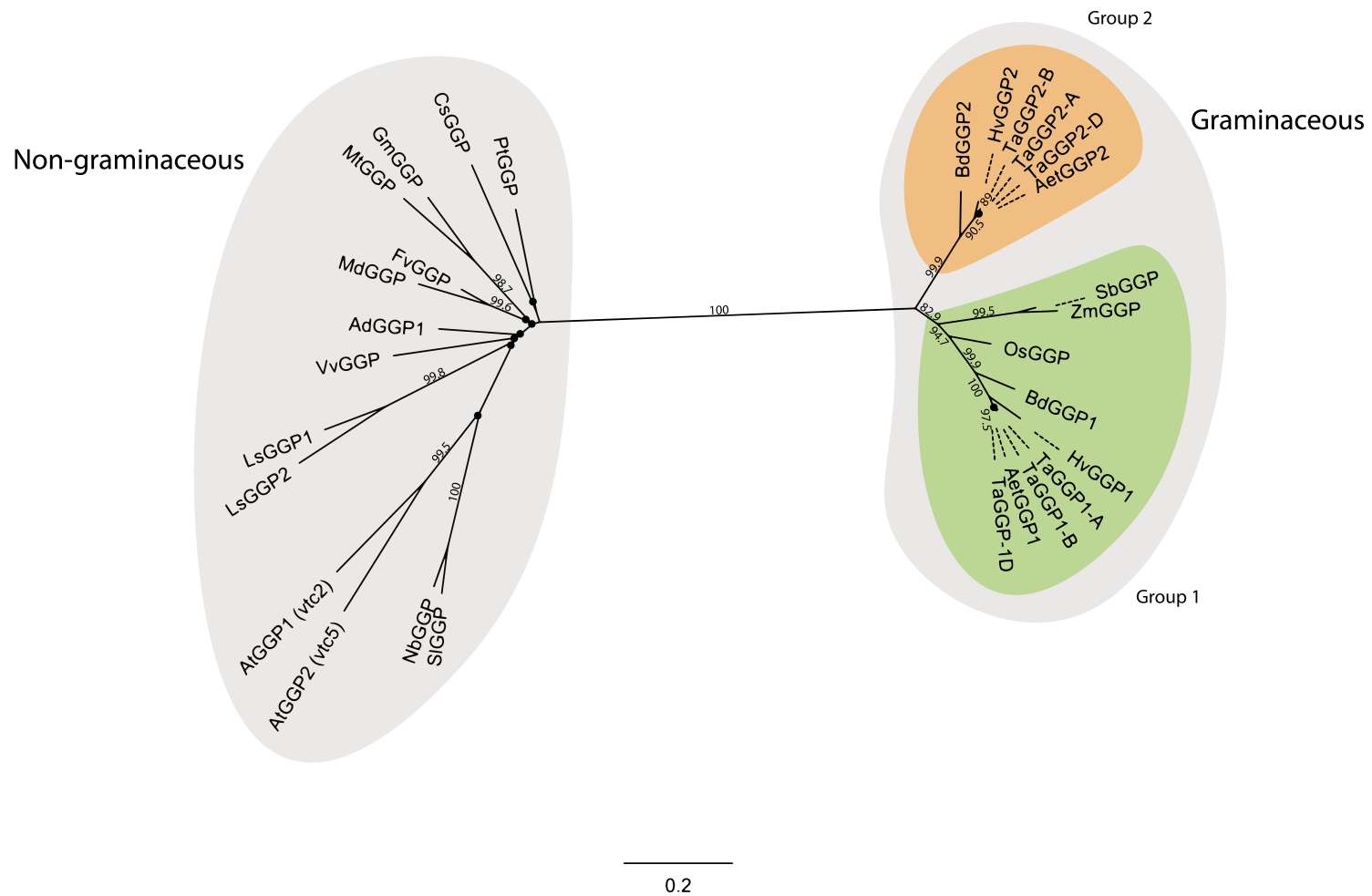


Figure 4.4. An unrooted phylogenetic tree of GGP proteins from a range of graminaceous and non-graminaceous species. Black nodes (●) represent weak bootstrap values (<75%). The scale bar corresponds to evolutionary distance in substitutions per site and the numbers correspond to bootstrap percentage. The prefixes for the species are the same as those presented in Figure 4.1 and 4.2.

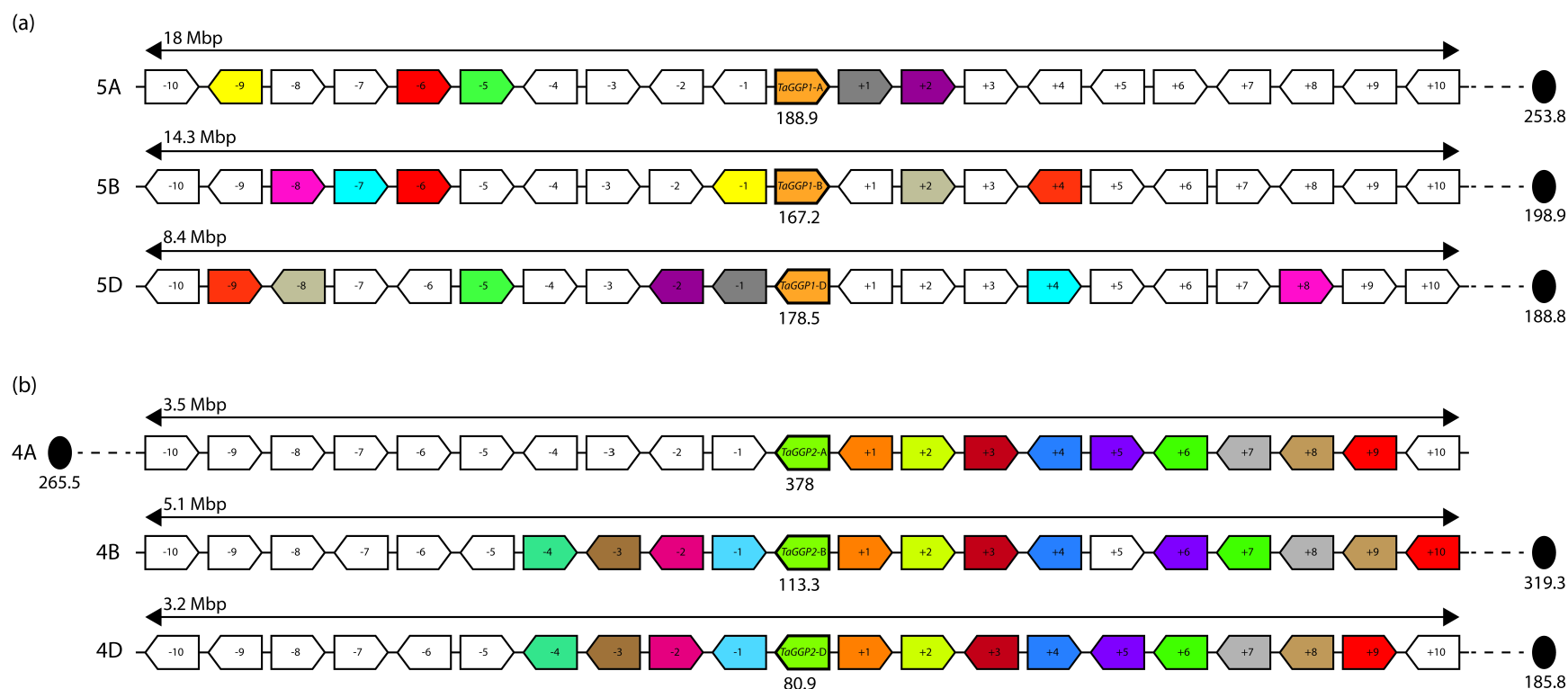


Figure 4.5. Microsynteny analysis of the *TaGGP* genes from bread wheat cv. Chinese Spring. (a) *TaGGP1* and (b) *TaGGP2* homoeologous SRs on chromosomal groups 5 and 4, respectively. Syntenic genes are represented by arrow colour. The orientation of the genes is indicated by the direction of the arrows. The physical length (Mbp) of each of the *TaGGP* SRs is provided for each chromosome. The physical position (Mbp) of the *TaGGP* genes and the centromeres (black oval) are also presented. The physical distance (Mbp) between the proximal gene of the SR and the respective centromere (black dotted line) is 58.2, 22.2 and 4.4 for the *TaGGP1-A*, *TaGGP1-B*, and *TaGGP1-D* SRs, respectively, and 127.5, 203.8 and 104.6 for the *TaGGP2-A*, *TaGGP2-B*, and *TaGGP2-D* SRs, respectively. The list of genes used in this analysis and their respective position and gene function are provided in Table S1

TaGGP genes themselves. Synteny between the *TaGGP1* homoeologous SRs was much lower than that of the *TaGGP2* homoeologous SRs with only 19 genes having shared synteny—the *TaGGP1* homoeologs were the only genes to share synteny on all three homoeologous SRs—whereas the *TaGGP2* homoeologous SRs have 37 genes with shared synteny. The *TaGGP1* homoeologs are found in the centromeric region of their respective chromosomes, whereas the *TaGGP2* homoeologs are found in the interstitial region of their respective chromosomes between the pericentromeric and telomeric regions.

4.3.5. The two groups of *TaGGP* genes have distinct expression patterns in bread wheat tissues

Quantitative reverse transcription-PCR analysis of a range of bread wheat tissues and developmental stages revealed distinct patterns of expression for the *TaGGP1* and *TaGGP2* genes (Figure 4.6). The *TaGGP1* homoeologs were broadly expressed across different tissues and developmental stages, with the highest levels of expression detected in the seedling leaf, bracts, and anthers (Figure 4.6a). The *TaGGP1* homoeologs displayed a trend of nonbalanced expression with the *TaGGP1-D* homoeolog having lower expression levels across all tissues relative to *TaGGP1-A* and *TaGGP1-B*. In contrast to the *TaGGP1* homoeologs, the *TaGGP2* homoeologs were highly expressed in the anthers, with approximately 50- to 580-fold higher expression in anthers relative to other tissues (Figure 4.6b). The *TaGGP2* homoeologs also displayed a trend of nonbalanced expression with the *TaGGP2-B* homoeolog having lower expression levels across all tissues relative to *TaGGP2-A* and *TaGGP2-D*. Taken together, the *TaGGP* genes were most highly expressed in anthers, seedling leaf, and bracts and most lowly expressed in the crown, seedling root, and caryopses 3-5 DAP.

We searched for anther/pollen specific cis-regulatory elements in the 1-kb promoter region of the *TaGGP* genes to identify whether the high expression of the *TaGGP2* homoeologs in anthers could be explained due to differences in these regulatory elements (Figure S4.4). However, the high level of expression of the *TaGGP2* homoeologs in anthers does not appear to correlate with the density of anther/pollen cis-regulatory elements surrounding the transcription start site (TSS), with a higher density of anther/pollen cis-regulatory elements instead being present in regions surrounding the TSS of the *TaGGP1* homoeologs (Figure S4.4).

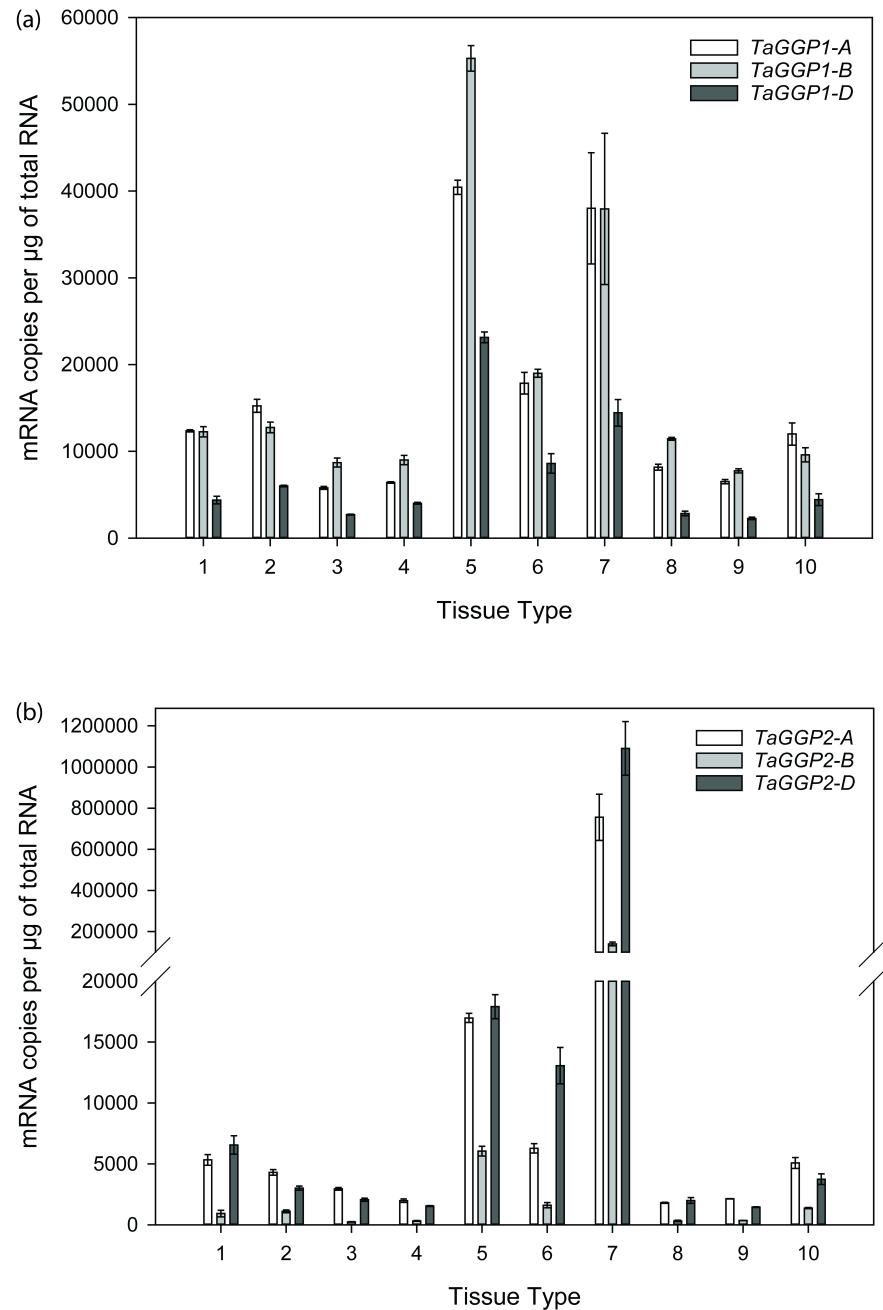


Figure 4.6. Quantitative reverse transcription-PCR analysis of the *TaGGP* genes from bread wheat cv. Chinese Spring. Relative expression of the (a) *TaGGP1* and (b) *TaGGP2* homoeologs is provided in: (1) embryonic root; (2) mesocotyl; (3) seedling root; (4) crown; (5) seedling leaf; (6) bracts; (7) anthers; (8) pistil; (9) caryopsis 3-5 DAP; and (10) embryo 22 DAP. The geometric mean of *TaActin*, *TaCyclophilin*, and *TaELF* were used as normalisation factors. Error bars indicate SEM of three technical replicates derived from a bulk of three independent biological samples.

4.3.6. Transient transformation of *TaGGP* genes in *N. benthamiana* leaf increases ascorbate concentration

Coding sequences of the *TaGGP1-A*, *TaGGP1-B*, *TaGGP1-D*, *TaGGP2-A*, and *TaGGP2-B* genes increased ascorbate concentrations in transiently transformed *N. benthamiana* leaf relative to the control (Figure 4.7). Transformation with *TaGGP1-D* produced the greatest increase in ascorbate concentration (5.3-fold) relative to the control. Similar fold increases in ascorbate concentrations were also observed with transformation of *TaGGP1-B* (4.1-fold), *TaGGP2-A* (4.4-fold), and *TaGGP2-B* (4.5-fold). Transformation with *TaGGP1-A* produced a 2.4-fold increase relative to the control, which was the lowest increase in ascorbate concentration observed. *Agrobacterium tumefaciens* containing the *TaGGP2-D* construct were unable to be recovered. The higher fold increase in ascorbate observed for some of the *TaGGP* genes relative to the positive control *AcGGP* gene from kiwifruit may be explained by the strength of the promoters used to drive transgene expression: the dual 35S promoter used to drive the *TaGGP* genes is a stronger promoter than the single 35S promoter used to drive *AcGGP*.

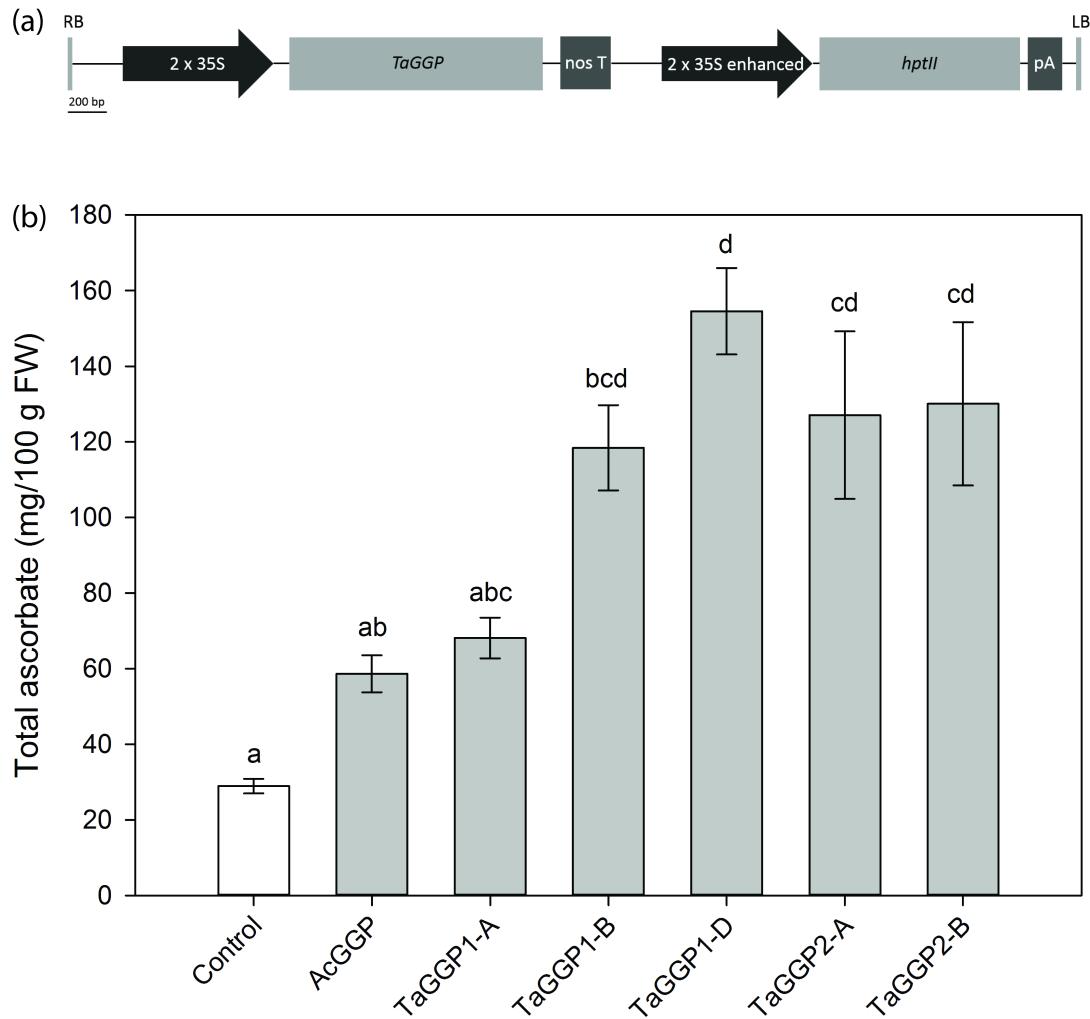


Figure 4.7. Transient transformation of the *TaGGP* coding sequences in *N. benthamiana*. (a) Schematic representation of the T-DNAs used for constitutive overexpression of the *TaGGP* genes in *N. benthamiana*. RB, right border; 2 x 35S, dual CaMV 35S promoter; *TaGGP*, coding sequence of *TaGGP1-A* (1,293 bp), *TaGGP1-B* (1,293 bp), *TaGGP1-D* (1,293 bp), *TaGGP2-A* (1,296 bp), and *TaGGP2-B* (1,296 bp); nos T, nopaline synthase terminator; 2 x 35S enhanced, dual CaMV 35S promoter enhanced; *hptII*, hygromycin phosphotransferase II; pA, CaMV poly(A) signal; LB, left border. (b) Ascorbate concentrations of *N. benthamiana* leaves co-infiltrated with *A. tumefaciens* (GV3101 MP90) containing the constructs of interest and a P19 construct to prevent post-transcriptional gene silencing. The control was infiltrated with *A. tumefaciens* containing the P19 construct alone. The *AcGGP* gene from kiwifruit driven by a single 35S promoter in the pGreen vector system was used as a positive control. Bars represent mean $\pm$ SEM of three infiltrated young leaves. Means that do not share a letter are significantly different (one-way ANOVA followed by Tukey post-hoc test with 95% confidence level).

4.4. Discussion

4.4.1. The *TaGGP* genes are highly conserved, functional, and have an uORF amenable to genome editing

We identified six *TaGGP* genes in the bread wheat genome, comprising of three *TaGGP1* homoeologs on chromosomal group 5 and three *TaGGP2* homoeologs on chromosomal group 4. Five of the genes are located on the short arm of their respective chromosomal group; the *TaGGP2-A* gene is located on the long arm of chromosome 4A due to a known pericentric inversion between the long and short chromosome arms [268]. We have also identified two copies of the *GGP* gene in the barley, *Brachypodium*, and *A. tauschii* genomes, as well as one copy in the sorghum genome. To our knowledge, this is the first time that the *HvGGP1*, *HvGGP2*, *BdGGP1*, *BdGGP2*, *AetGGP1*, *AetGGP2*, and *SbGGP* gene have been reported on.

The *TaGGP* protein sequences are highly conserved, particularly between homoeologs, with up to 99.3% shared identity (Figure 4.2). This high level of conservation is striking given the genetic redundancy of the hexaploid wheat genome, which generally displays accelerated rates of coding sequence evolution relative to diploid organisms such as *Brachypodium* and suggests that all six *TaGGP* genes play important functional roles during wheat development and as such may be resistant to genetic drift [269]. This hypothesis is consistent with our finding that all five *TaGGP* genes transiently transformed into *N. benthamiana* leaf increased leaf ascorbate concentrations relative to the control, confirming their functional role in ascorbate biosynthesis *in planta* (Figure 4.7). We can also conclude from the transient transformation that the 13 residues that distinguish the graminaceous GGP2 proteins from GGP1 proteins do not appear to affect the function of the GGP2 proteins. The substantial increase in ascorbate concentrations in *N. benthamiana* leaf observed for the *TaGGP* genes is consistent with GGP as the rate-limiting step towards ascorbate biosynthesis in plants [62, 64, 71, 108], and demonstrates the potential of utilising these genes, particularly *TaGGP1-D*, to increase ascorbate concentrations not only in wheat but other crops.

All six *TaGGP* proteins contained the highly conserved H ϕ H ϕ Q motif (where ϕ is a hydrophobic amino acid) that is characteristic of the galactose-1-phosphate uridylyltransferase branch of the HIT superfamily (Fig 4.2) [108, 112, 270]. They also all possess the highly conserved KKRP NLS, which may be responsible for localizing the *TaGGP* proteins to the nucleus, in

addition to the cytoplasm where most of the ascorbate biosynthesis occurs (Fig 4.2) [271]. Fusion proteins of *AtGGP1*:YFP, *SlGGP1*:GFP, and *SlGGP2*:GFP have been found to localize to the cytoplasm and nucleus suggesting that GGP may be a dual-function protein serving both enzymatic and regulatory roles [271-273]. Determining the extent to which this NLS influences subcellular localization of GGP and what role GGP may be playing in the nucleus warrants further investigation.

We also identified a highly conserved uORF in the long 5'UTR of the mRNA for each of the six *TaGGP* genes (Figure 4.1 and 4.3). Notably, the uORF of the *TaGGP2* homoeologs had a truncation of 11 residues relative to the *TaGGP1* homoeologs and the same truncation of 11 residues was also observed for the *BdGGP2* and *HvGGP2* uORFs relative to the *BdGGP1* and *HvGGP2* uORFs, respectively, indicating that this truncation occurred prior to the evolutionary divergence of *Brachypodium* and the Triticeae tribe (Figure 4.3). If these graminaceous *GGP2* uORFs are still functional this would suggest that these 11 residues are not critical to the functionality of the peptide. Disruption of the *GGP* uORF through genome editing has recently been shown as a viable strategy to increase ascorbate concentrations in a variety of plant species [153, 154]. For example, CRISPR/Cas9-targeted mutagenesis of the *AtGGP1* uORF in *Arabidopsis* and the *LsGGP1* and *LsGGP2* uORFs in lettuce to disrupt their function increased foliar ascorbate concentrations 1.7-, 1.4-, and 2.6-fold, respectively. Additionally, the increased ascorbate concentrations in lettuce conferred enhanced tolerance to methyl viologen-induced oxidative stress [153]. Similarly, CRISPR/Cas9-targeted mutagenesis of the *SlGGP2* uORF in tomato to disrupt its function increased foliar ascorbate concentrations approximately 1.4-fold [154]. Due to the plasticity of the hexaploid wheat genome, manipulating individual or multiple *TaGGP* uORFs with genome editing tools such as CRISPR/Cas9 represents a promising strategy to produce ascorbate-enriched wheat plants whilst also minimizing any pleiotropic effects on plant development that may arise due to disruption of the *GGP* uORF.

4.4.2. A *Mutator*-like transposable element may be responsible for the duplication and specialized gene expression of *GGP2* in the *Brachypodium*/Triticeae lineage

We have categorized the graminaceous *GGP* genes described in this manuscript as group 1 or group 2 based on phylogeny. The protein and coding sequences of the *TaGGP1* homoeologs, *AetGGP1*, *HvGGP1*, and *BdGGP1* are more closely related to *OsGGP*, *ZmGGP*, and *SbGGP* in

terms of evolutionary distance than that of the *TaGGP2* homoeologs, *AetGGP2*, *HvGGP2*, and *BdGGP2* (Fig 4.4). This grouping of graminaceous *GGP1* and *GGP2* genes is also consistent with the chromosomal location of the *GGP* genes and the known evolutionary history of grass genomes. For example, the wheat chromosomal group Ta5, barley chromosome Hv5, *Brachypodium* chromosome Bd4, rice chromosome Os12, and sorghum chromosome Sb8, which each harbour *GGP1* genes, are all descended from the ancestral grass chromosome A12 [274]. Whereas, the wheat chromosomal group Ta4, barley chromosome Hv4, and *Brachypodium* chromosome Bd4, which each harbour *GGP2* genes, are all descended from the ancestral grass chromosome A11 (in *Brachypodium* A11 and A12 were fused to form the single chromosome Bd4) [274]. Given that rice, maize, and sorghum each have a single copy of *GGP*, while wheat, barley, *Brachypodium*, and *A. tauschii* each have two copies of *GGP*, it appears that *GGP* may have been duplicated from A12 to A11 in the ancestor of the *Brachypodium*/Triticeae lineage. This duplication event would have occurred after the evolutionary divergence of the *Brachypodium*/Triticeae lineage from rice approximately 33.5 million years ago (MYA) but prior to the evolutionary divergence of *Brachypodium* and the Triticeae tribe approximately 24.5 MYA [274].

Our microsynteny analysis of the *TaGGP* genes revealed that the *TaGGP1* and *TaGGP2* SRs have no shared synteny except for the *TaGGP* genes themselves (Fig 4.5). We also did not find any shared synteny between the *BdGGP1* and *BdGGP2* SRs in *Brachypodium* (Table S2) or between the *HvGGP1* and *HvGGP2* SRs in barley (Table S3). This lack of synteny between the *GGP1* and *GGP2* SRs in wheat, *Brachypodium*, and barley suggests that *GGP2* may have been duplicated via transposition in the ancestor of the *Brachypodium*/Triticeae lineage since other methods of gene duplication preserve gene synteny [275]. Given that the introns and gene structure of the wheat, *Brachypodium*, and barley *GGP2* genes have been preserved, a Class II transposable element (DNA transposon) that transposes via a DNA intermediate may have been responsible for such a transposition. We propose that a *Mutator*-like transposable element (MULE)—a family of DNA transposons that are prevalent in plants and capable of capturing, mobilizing, and amplifying gene fragments—may have been responsible for the duplication of *GGP2* [276]. However, given the proposed evolutionary time since *GGP2* was duplicated (24.5 to 33.5 MYA) any features of the transposable element would have degenerated by now and no longer be detectable [275].

We identified distinct expression patterns for the *TaGGP1* and *TaGGP2* genes with the *TaGGP1* homoeologs broadly expressed across different tissues and developmental stages, whereas the *TaGGP2* homoeologs were highly specialized with approximately 50- to 580-fold higher expression in anthers relative to other tissues (Figure 4.6). Our *TaGGP* gene expression results are supported by gene expression data extracted from <http://bar.utoronto.ca/> which also demonstrates that the *TaGGP1* homoeologs are broadly expressed across different tissues and development stages, and that the *TaGGP2* homoeologs are highly expressed in anthers and leaves relative to other tissues (Figure S4.5). Similar *GGP* expression profiles were also observed for other graminaceous species in terms of higher expression in anthers and/or leaves relative to other tissues (Figure S4.5 and S4.6). To our knowledge this is the first time that specialized expression of a *GGP* gene in anthers has been documented, however, a previous study found high ascorbate concentrations correlated with high RNA concentrations in the developing anthers of jimsonweed (*Datura stramonium* L.) [277]. Together this expression data suggests that ascorbate may play an important role in anther development in wheat and other graminaceous species, for example protecting pollen from ROS induced by environmental stresses [278]. Determining whether the high expression of *GGP* in anthers is correlated with higher ascorbate concentrations relative to other tissues warrants further study.

In regards to the origin of the *cis*-regulatory elements responsible for the high expression of *GGP2* genes in the anthers relative to other tissues in wheat, *Brachypodium*, and barley, we propose two hypotheses: (i) following duplication, the regulatory region of the *GGP2* gene diverged due to a reduction in selection pressure, which subsequently generated *de novo cis*-regulatory elements, or (ii) *cis*-regulatory elements in the regulatory region were inherited from a transposable element that may have been responsible for the duplication of *GGP2*. In support of the first hypothesis, sequence alignments of the 1-kb promoter region of the wheat, *Brachypodium*, and barley *GGP* genes demonstrate that the promoter region of the *GGP2* genes have commonalities with the promoter region of the *GGP1* genes but have otherwise diverged significantly (Figure S4.7). It is possible that sequence changes to the promoter region of *GGP2* genes have generated *de novo cis*-regulatory elements responsible for the high expression of *GGP2* genes in anthers. On the other hand, transposable elements are known to influence the expression of transposed genes [275]. For example, MULEs—including any sequences that have been captured, mobilized, and amplified by a MULE—have been shown to be preferentially expressed

in reproductive tissues [275, 279]. The terminal inverted repeat (TIR) of MULEs are proposed to serve as the regulatory region for the expression of MULEs, including any transposed gene fragments [275, 279]. If *GGP2* was transposed by a MULE in the ancestor of the *Brachypodium*/Triticeae lineage, it is possible that *GGP2* inherited *cis*-regulatory elements responsible for the high expression of *GGP2* in anthers directly from the TIR of the MULE.

4.4.3. Applications towards plant breeding

Abiotic stresses, such as high temperature, are responsible for major losses in global wheat productivity and are predicted to intensify due to climate change [280]. Identifying genetic strategies to mitigate or offset these losses will be essential for maintaining global wheat production levels. Exogenous application of ascorbate to wheat plants has been shown to alleviate salt- and drought-induced oxidative stress in many studies [281-286]. The development of high-ascorbate bread wheat cultivars through overexpression of the *TaGGP* genes or genome editing of one or more *TaGGP* uORFs to disrupt their function could help to mitigate wheat production losses caused by abiotic stress without the need for exogenous applications.

With respect to human nutrition, the development of bread wheat products with elevated levels of ascorbate may be limited due to the decline of ascorbate to negligible levels during grain maturation that may be difficult to overcome [54, 55]. However, wheat products made from germinated grain may provide a vehicle for ascorbate biofortification since germinated grain contains considerably higher levels of ascorbate relative to non-germinated grain [287, 288]. Germinated grain from bread wheat cultivars with an enhanced capacity to biosynthesize ascorbate could represent a novel way for improving human dietary intake of ascorbate.

4.5. Conclusions

We identified six *TaGGP* genes in the bread wheat genome, each with a highly conserved uORF. Phylogenetic and microsynteny analyses revealed that *GGP2* may have been duplicated via a transposable element in the ancestor of the *Brachypodium*/Triticeae lineage. The *TaGGP2* homoeologs had specialized expression in anthers which may have been inherited from the TIR of a MULE. Transient transformation of the *TaGGP* coding sequences in *N. benthamiana* leaf increased ascorbate concentrations up 5.3-fold, confirming their activity *in planta*. We propose that the six *TaGGP* genes and uORFs identified in this study will provide a valuable genetic resource for increasing ascorbate biosynthesis in bread wheat to improve wheat abiotic stress tolerance and human nutrition.

4.6. Methods

4.6.1. Identification and naming of the *GGP* genes and uORFs

The *OsGGP* coding sequence (LOC4351698) was used as a query in BLASTN searches against the International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v1.0 assembly of the wheat genome cv. Chinese Spring [266] to identify *TaGGP* coding sequences. Each *TaGGP* gene was assigned a unique name based on homoeologous grouping and their respective subgenomes and followed the recommended rules for gene symbolization in wheat (<http://wheat.pw.usda.gov/ggpages/wgc/98/Intro.htm>). Gene prediction models including predicted coding and proteins sequences were obtained from the IWGSC database. The *OsGGP* coding sequence was also used as a query in BLASTN searches against the genome sequence databases for barley (IBSC_v2), *Brachypodium* (*Brachypodium_distachyon_v3.0*), *A. tauschii* (*Aet_v4.0*) from <https://plants.ensembl.org> and for sorghum (*Sorghum_bicolor_NCBIv3*) from <https://www.ncbi.nlm.nih.gov>. The structure of the *GGP* genes was determined with Splign (<https://www.ncbi.nlm.nih.gov/sutils/splign/splign.cgi>) [289]. The *GGP* uORF coding and peptide sequences were manually identified in the 5'UTR of the *GGP* genes starting from the highly conserved AUC codon encoding isoleucine as previously described [163].

4.6.2. Cloning and gene sequencing

Full-length gene sequences of *TaGGP1-A* and *TaGGP2-A* could not be retrieved from the IWGSC RefSeq v1.0 assembly of the wheat genome cv. Chinese Spring [266] due to each gene having a gap in intron 2 where sequence was missing. PCR primers were designed to amplify gene fragments spanning the gap of the *TaGGP1-A* and *TaGGP2-A* genes from cv. Chinese Spring genomic DNA using the respective IWGSC gene sequences and Primer3 software [207, 208]. The *TaGGP1* gene fragment was PCR amplified with forward primer 5' - GGCAATCTGTTAGGCAAGCA -3' and reverse primer 5' - TGTCAAAAACAGGTATCAGCAATTT - 3'. The *TaGGP2* gene fragment was PCR amplified with forward primer 5' - AGTTCTTCCTAAATTCTCTCCTCCTT - 3' and reverse primer 5' - TTGAGGAATCACCTCACCTTA - 3'. PCR amplification cycles consisted of 1 cycle = 1 m 95°C; 35 cycles = 20 s 95°C, 30 s 61°C, 30 s 72°C; 1 cycle = 5 m 72°C. PCRs were performed in a final volume of 20 µL for MyTaq™ HS Red DNA Polymerase (Bioline, MA, USA) according

to manufacturer's instructions. The PCR products were purified with DNA Clean & Concentrator-5 (Zymo Research, CA, USA) and cloned into the pGEM®-T Vector System (Promega, WI, USA) according to manufacturer's instructions. Restriction enzyme digestion of the cloned PCR products in pGEM®-T with PvuII (New England Biolabs, MA, USA) according to the manufacturer instructions was used to discriminate the *TaGGP1-A* and *TaGGP2-A* gene fragments from their respective B and D homoeologous gene fragments prior to Sanger sequencing at the Australian Genome Research Facility Ltd (<http://www.agrf.org.au/>).

4.6.3. Sequence alignment and phylogenetic analysis of the *GGP* and uORF sequences

Sequence alignment and phylogenetic analysis of the *GGP* coding, protein, and uORF peptide sequences were performed in Geneious 10.2.2 using ClustalW alignment with default parameters and the PhyML Geneious plugin [290] with a K80 substitution model for nucleotide sequences or a LG substitution model for amino acid sequences and a bootstrap value of 1,000. The *GGP* genes used in the sequence alignments and phylogenetic analyses are provided in Table S4.

4.6.4. Microsynteny analysis of the *TaGGP*, *BdGGP*, and *HvGGP* genes

Ten high confidence genes upstream and downstream of the six *TaGGP*, two *BdGGP*, and two *HvGGP* genes were retrieved from the IWGSC RefSeq v1.0 assembly of the wheat genome cv. Chinese Spring [266], the *Brachypodium distachyon* v3.0 assembly of the *Brachypodium* genome (<https://plants.ensembl.org>), and the IBSC_v2 assembly of the barley genome (<https://plants.ensembl.org>), respectively, excluding genes associated with retrotransposons and retrovirus related polyproteins, and used as query sequences in BLASTN searches against the respective databases with an E-value threshold of $1e^{-10}$. The genes were considered syntenic based on their genomic location and sequence identity ($\geq 90\%$). The list of the genes used in the microsynteny analysis of the six *TaGGP*, two *BdGGP*, and two *HvGGP* genes, and their respective position and gene function, are provided in Tables S1, S2, and S3, respectively.

4.6.5. Quantitative reverse transcription-PCR analysis of the *TaGGP* genes

The cDNA library of 10 different tissues and development stages of bread wheat cv. Chinese Spring was previously generated as described [291]. The quantitative reverse transcription-PCR analysis of the six *TaGGP* genes was carried out using subgenome specific PCR primer pairs

(Table S5) for each *TaGGP* gene as previously described [292]. Subgenome specificity of primer pairs was verified using cv. Chinese Spring nulli-tetrasomic DNA [293]. Primer efficiency was $\geq 97\%$ for all primer pairs. A three gene normalisation factor of the most stable housekeeping genes *TaActin*, *TaCyclophilin*, and *TaELF* was used to normalise qRT-PCR expression data as previously described [291, 294]. Supplementary *GGP* gene expression data was extracted using the corresponding gene identifiers provided in Table S4 from <http://bar.utoronto.ca/> for wheat [295], barley [296], and maize [297], and from <https://www.ebi.ac.uk> for *Brachypodium*, rice, and sorghum [298].

4.6.6 Analysis of *TaGGP* promoter regions

For each *TaGGP* gene, the first exon as well as the 1-kb upstream region, which was considered the promoter region, was extracted from the genome sequence databases for *T. aestivum* (IWGSC) from <https://plants.ensembl.org>. These sequences were used as a query in BLASTN searches against the NCBI expressed sequence tag (EST) database for *T. aestivum* (<https://blast.ncbi.nlm.nih.gov>) to identify *TaGGP* cap-trapped full-length cDNA sequences and the TTS. The identification of anther/pollen *cis*-acting elements was performed in Geneious 10.2.2 as previously described [292]. Sequence alignment of the 1-kb promoter region of the *GGP* genes was performed in Geneious 10.2.2 using the MAFFT Geneious plugin [299] with a L-INS-i algorithm.

4.6.7. Transient transformation of *N. benthamiana* leaves

Bread wheat cv. Gladius was sourced from the Australian Grains Genebank (<https://www.seedpartnership.org.au/associates/agg>). PCR primers were designed to amplify the coding sequence of the *TaGGP1* homoeologs and *TaGGP2* homoeologs from bread wheat cv. Gladius cDNA using Primer3 software [207, 208]. The *TaGGP1* homoeologs were PCR amplified with forward primer 5' – ATGAAGCTGACGATTAAGAGGGTA -3' and reverse primer 5' – TCACGGAATTACGAGGCAG - 3'. The *TaGGP2* homoeologs were PCR amplified with forward primer 5' – ATGGAGATGAAGCTGACGAT - 3' and reverse primer 5' – TCACTGAAGGACAAGGCAAC - 3'. PCR amplification cycles consisted of 1 cycle = 30 s 98°C; 35 cycles = 10 s 98°C, 15 s 61°C, 45 s 72°C; 1 cycle = 5 m 72°C. PCRs were performed in a final volume of 50 µL for Phusion High-Fidelity DNA Polymerase (New England Biolabs) according

to the manufacturer's instructions. The PCR products were purified with DNA Clean & Concentrator-5 (Zymo Research), A overhangs added with MyTaqTM HS DNA Polymerase (Bioline), ligated into the pCR8/GW/TOPO vector system using the pCRTM8/GW/TOPO[®] TA Cloning Kit (Invitrogen, CA, USA), and recombined into the pMDC32 vector system [203] using GatewayTM LR ClonaseTM II Enzyme mix (Invitrogen) according to the manufacturer's instructions. Cloning of the correct gene family member and gene orientation was confirmed by Sanger sequencing at the Australian Genome Research Facility Ltd. *A. tumefaciens* (GV3101 pMP90) containing the constructs of interest (except *TaGGP2-D*) were co-infiltrated with *A. tumefaciens* containing a P19 construct [300] to prevent post-transcriptional gene silencing into the younger leaves of 6-week-old *N. benthamiana* plants (LAB strain) grown in a plant growth room (16 h daylength, 24°C) [301]. Infiltration conditions were conducted as previously described [302]. Infiltration with *A. tumefaciens* containing the P19 construct alone was used as a control. Infiltration with *A. tumefaciens* containing the kiwifruit *AcGGP* gene in the pGreen vector system [71, 303] was used as a positive control. Leaves were harvested for ascorbate measurements 7 d after infiltration and frozen on dry ice.

4.6.8. Ascorbate measurement

Total ascorbate (ascorbate + dehydroascorbate) was extracted and measured as previously described with modifications [205]. Briefly, total ascorbate was extracted from ground, lyophilized tissue homogenised in extraction fluid containing 8% metaphosphoric acid, 2 mM EDTA, and 2 mM TCEP (Sigma-Aldrich, MO, USA) at 40°C for 2 h. The extract was centrifuged and 5 µl of supernatant was injected onto a C18 3 µm 33 x 7 mm Alltima Rocket column (Hichrom Limited, UK) with a flow rate of 1 mL/min. The concentration was determined by reverse phase HPLC with UV detection at 245 nm on a Shimadzu (Japan) 8050 instrument (RT 1.43 min) using ascorbic acid as a standard (Sigma-Aldrich). The elution procedure was gradient with mobile phases A: MS grade water with 1% formic acid (Sigma-Aldrich) and B: MS grade acetonitrile with 1% formic acid (Sigma-Aldrich). The peak was validated as ascorbic acid by using L-ascorbic acid analytical standard (Sigma-Aldrich) with HPLC-MS Multi Reaction Monitoring (MRM) scan, qualifier transition $m/z (+) = 177.05 > 95.00$.

5.6.9. Statistical analysis

Statistically significant differences between means were detected using a one-way ANOVA followed by a Tukey post-hoc test with a 95% confidence level using Minitab® 17.1.0 (<http://www.minitab.com/en-us/>).

5.8. Supplementary information

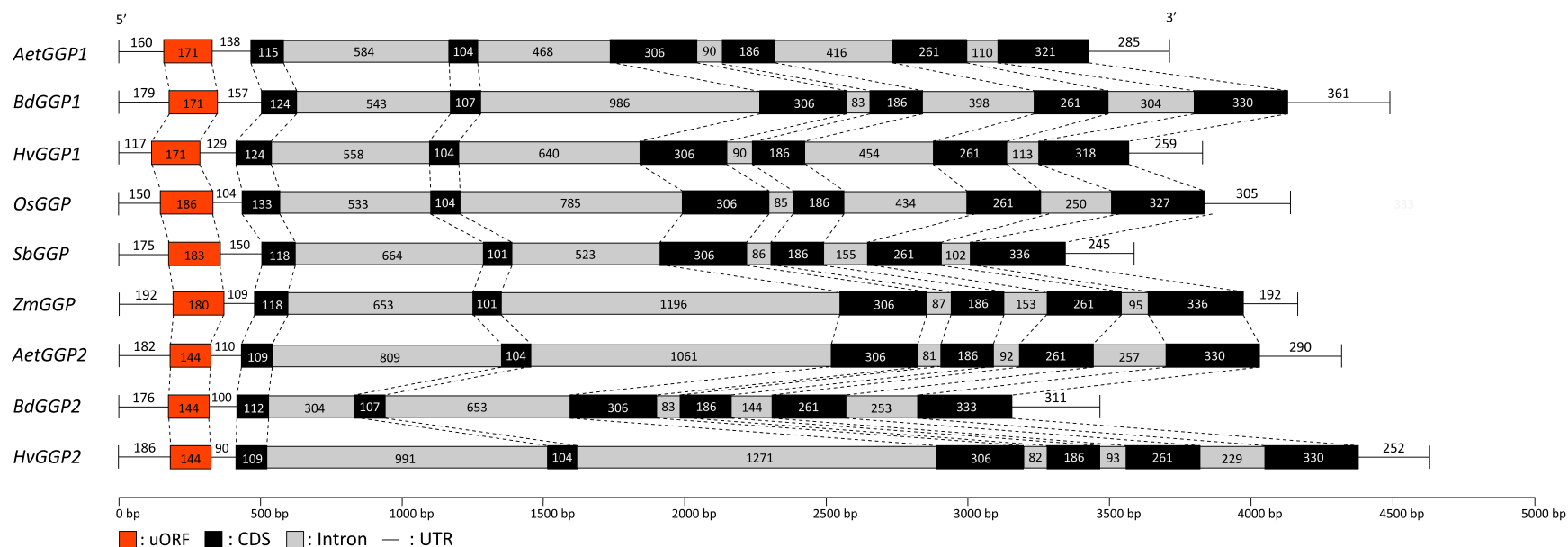


Figure S4.1. Gene structure of the *AetGPP*, *BdGPP*, *HvGPP*, *OsGPP*, *SbGPP*, and *ZmGPP* genes. The uORF (orange box), coding sequence (CDS, black box), introns (grey box), and UTR (lines) of the *GPP* genes are depicted and the length (bp) of each section indicated. The prefixes for the graminaceous species are as follows: Aet is *Aegilops tauschii*; Bd is *Brachypodium distachyon*; Hv is *Hordeum vulgare*; Os is *Oryza sativa*; Sb is *Sorghum bicolor*; Ta is *Triticum aestivum*; and Zm is *Zea mays*.

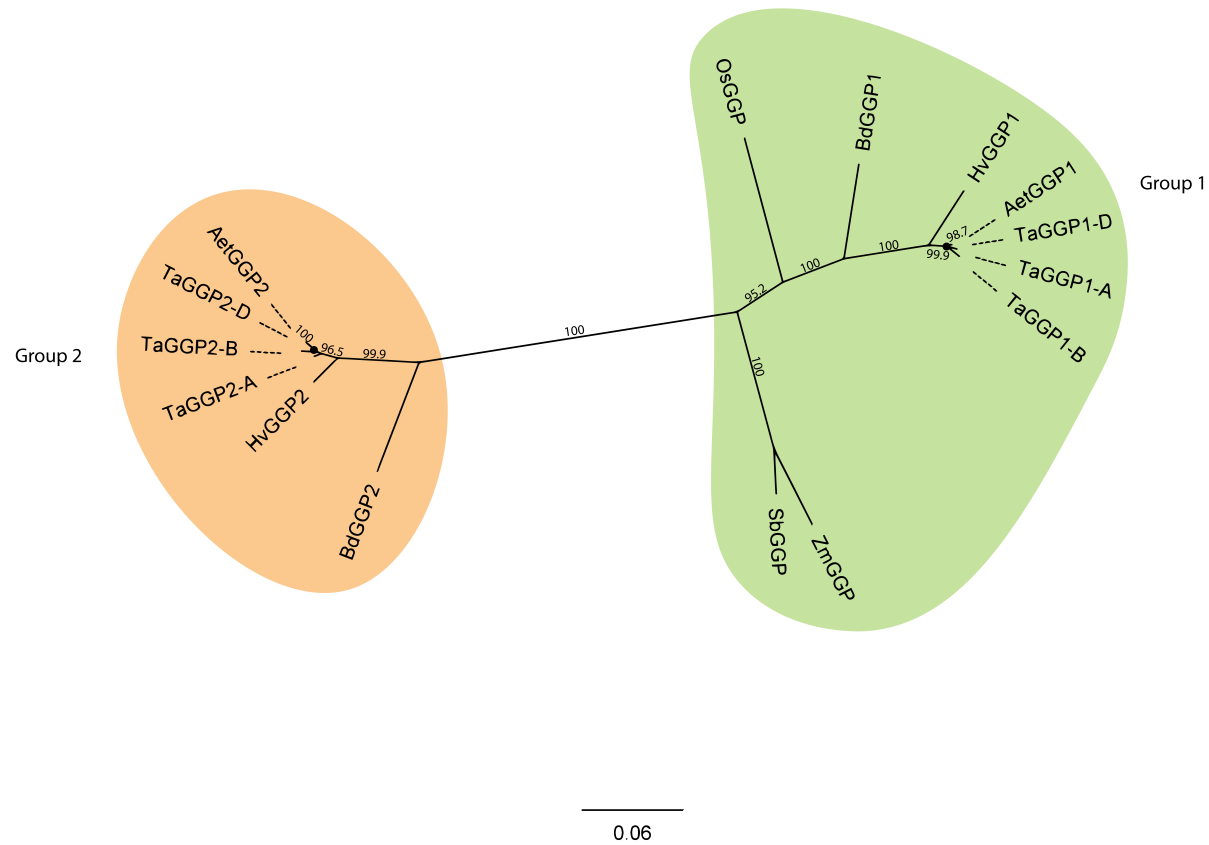


Figure S4.2. An unrooted phylogenetic tree of *GGP* coding sequences from a range of graminaceous species. Black nodes (●) represent weak bootstrap values (<75%). The scale bar corresponds to evolutionary distance in substitutions per base and the numbers correspond to bootstrap percentage. The prefixes for the species are the same as those presented in Figure 4.1 and 4.2.

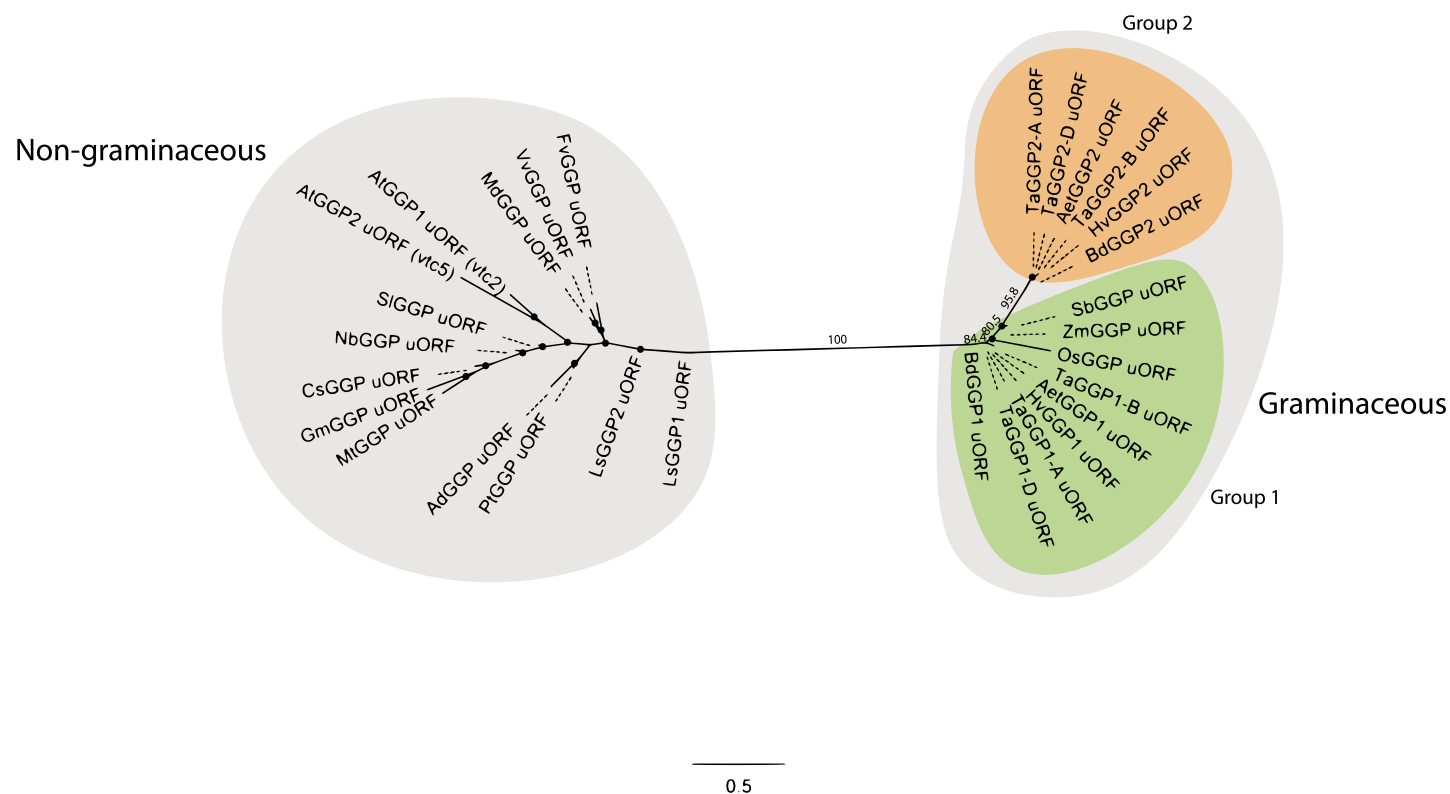


Figure S4.3. An unrooted phylogenetic tree of *GGP* uORF peptides from a range of graminaceous and non-graminaceous species. Black nodes (●) represent weak bootstrap values (<75%). The scale bar corresponds to evolutionary distance in substitutions per site and the numbers correspond to bootstrap percentage. The prefixes for the species are the same as those presented in Figure 4.1 and 4.2.

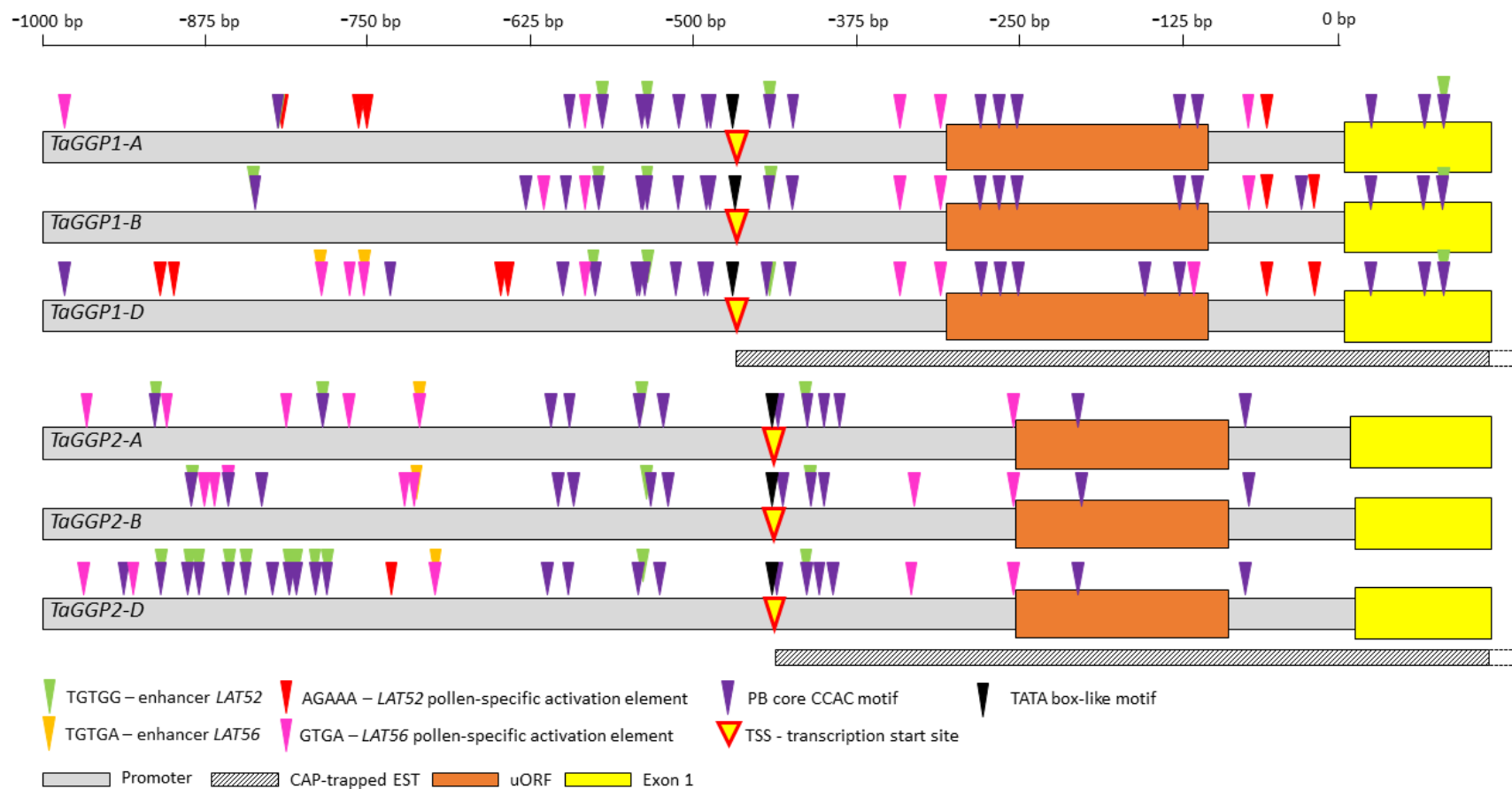


Figure S4.4. Annotation of anther/pollen *cis*-acting elements within the 1-kb promoter region of the *TaGGP* genes. The promoter (grey), uORF (orange), and first exon (yellow) of the *TaGGP* genes are depicted. The CAP-trapped EST (striped) for the *TaGGP1* homoeologs and *TaGGP2* homoeologs are also depicted.

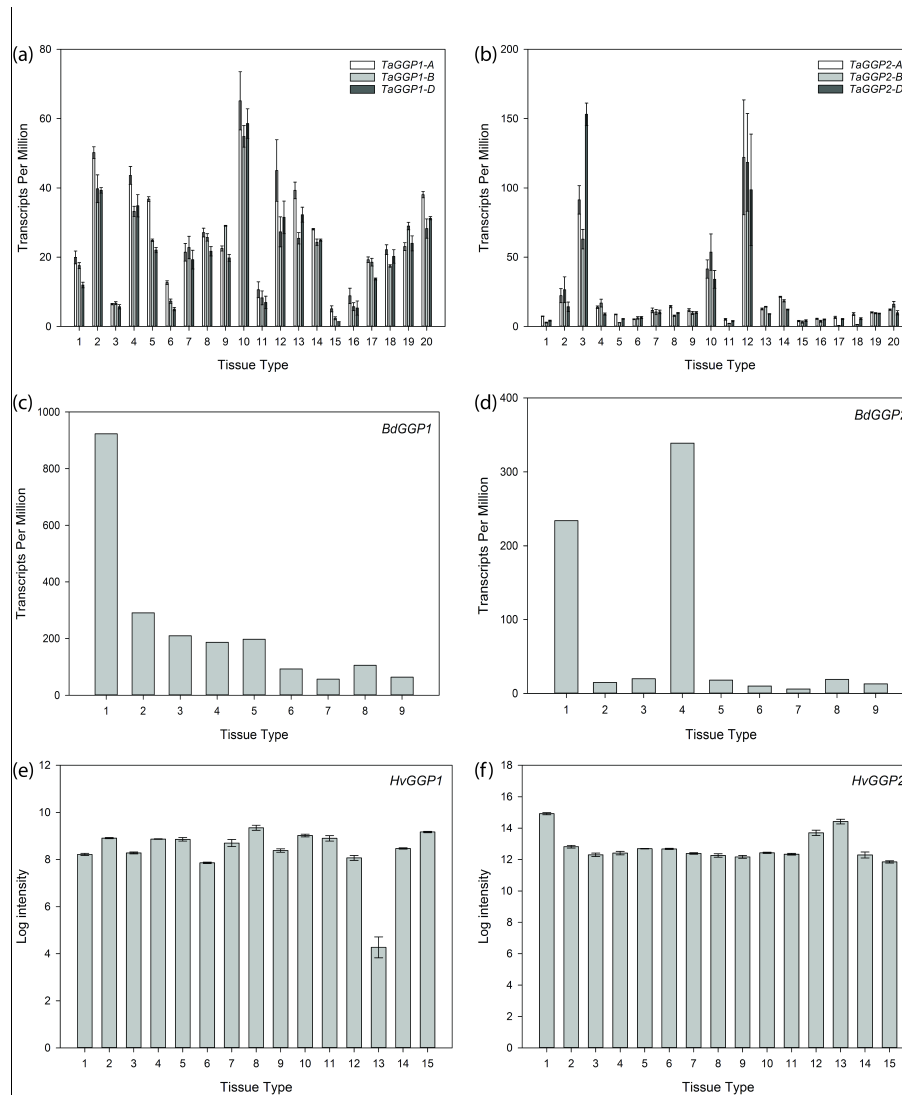


Figure S4.5. GGP gene expression data extracted from <http://bar.utoronto.ca/> for wheat and <https://www.ebi.ac.uk> for *Brachypodium*. Relative expression of (a) *TaGGP1-A/B/D* (TraesCS5A01G107800, TraesCS5B01G111700, and TraesCS5D01G122900, respectively) and (b) *TaGGP2-A/B/D* (TraesCS4A01G201200, TraesCS4B01G105300 and TraesCS4D01G102200, respectively) is provided in: (1) grain; milk grain stage, (2) first leaf blade; seedling stage, (3) anther; anthesis, (4) coleoptile; seedling stage, (5) stigma and ovary; anthesis, (6) grain; ripening stage, (7) awns; ear emergence, (8) glumes; ear emergence, (9) flag leaf blade; 50 percent spike, (10) fifth leaf blade; flag leaf stage, (11) grain; soft dough, (12) flag leaf blade; flag leaf stage, (13) shoot axis; first leaf stage, (14) roots; seedling stage, (15) endosperm; dough stage, (16) grain; hard dough, (17) roots; tillering stage, (18) shoot axis; full boot, (19) fifth leaf blade; ear emergence, and (20) first leaf sheath; seedling stage. Error bars indicate SEM of three biological replicates. Relative expression of (c) *BdGGP1* (Bradi4g40740) and (d) *BdGGP2* (Bradi4g22700) is provided in: (1) leaves, (2) early inflorescence, (3) emerging inflorescence, (4) anther, (5) pistil, (6) seed 5 DAP, (7) seed 10 DAP, (8) embryo 25 DAP, and (9) endosperm 25 DAP. Relative expression of (e) *HvGGP1* (AK355342) and (f) *HvGGP2* (AK358028) is provided in: (1) leaf, (2) crown, (3) root, (4) coleoptile, (5) mesocotyl, (6) radicle, (7) caryopsis without embryo, (8) embryo 22 DAP, (9) caryopsis 5 DAP, (10) caryopsis 10 DAP, (11) caryopsis 16 DAP, (12) floral bracts, (13) anthers, (14) pistil, and (15) immature inflorescence. Error bars indicate SEM of three biological replicates.

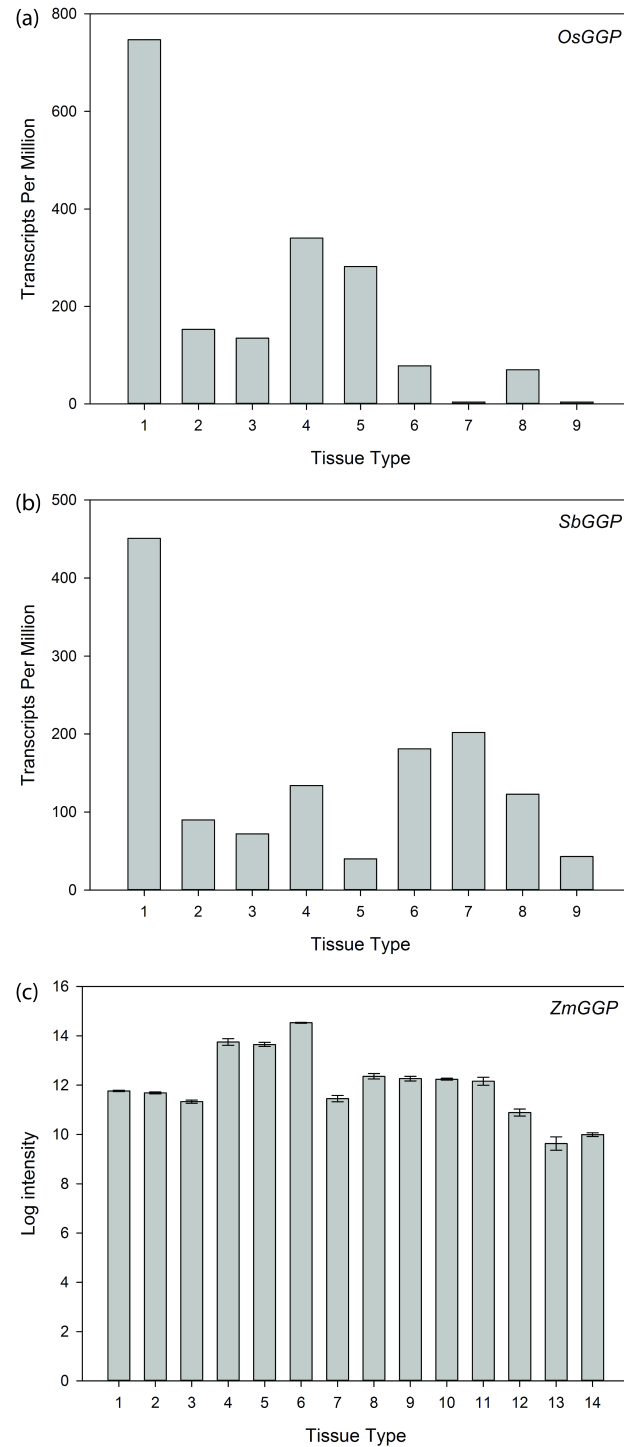


Figure S4.6. *GGP* gene expression data extracted from <https://www.ebi.ac.uk> for rice and sorghum and <http://bar.utoronto.ca/> for maize. Relative expression of (a) *OsGGP* (LOC_Os12g08810) and (b) *SbGGP* (SORBI_3008G064700) is provided in: (1) leaves, (2) early inflorescence, (3) emerging inflorescence, (4) anther, (5) pistil, (6) seed 5 DAP, (7) seed 10 DAP, (8) embryo 25 DAP, and (9) endosperm 25 DAP. Relative expression of (c) *ZmGGP* (GRMZM2G057186_T01) is provided in: (1) germinating seed 24 h, (2) coleoptile 6 DAS, (3) primary root 6 DAS, (4) anthers R1, (5) whole seedling VE, (6) pooled leaves V1, (7) pre-pollination cob R1, (8) silks R1, (9) embryo 16 DAP, (10) embryo 20 DAP, (11) embryo 24 DAP, (12) endosperm 16 DAP, (13) endosperm 20 DAP, and (14) endosperm 24 DAP. Error bars indicate SEM of three biological replicates.



Figure S4.7. Nucleotide sequence alignment of the 1-kb promoter regions of wheat, *Brachypodium*, and barley *GGP* genes. (a) *GGP1* genes, (b) *GGP2* genes, and (c) *GGP1* and *GGP2* genes. Green, olive green, yellow, and white background colour represents 100%, 80 to 100%, 60 to 80%, and less than 60% conservation of amino acids between species, respectively. The TATA box-like motifs and uORF are outlined in red and black, respectively.

Table S4.1. Ten high confidence genes upstream and downstream of the six *TaGGP* genes.

SR	Position	Gene ID	Gene Function
<i>TaGGP1-A</i>	-10	TraesCS5A01G106800	Myb/SANT-like DNA-binding domain protein
	-9	TraesCS5A01G106900	Protein transport protein Sec24-like
	-8	TraesCS5A01G107000	Ribosomal RNA large subunit methyltransferase F
	-7	TraesCS5A01G107100	kinesin 3
	-6	TraesCS5A01G107200	mediator of RNA polymerase II transcription subunit-like protein
	-5	TraesCS5A01G107300	Protein kinase, putative
	-4	TraesCS5A01G107400	uncharacterized
	-3	TraesCS5A01G107500	Cell division topological specificity factor-like protein
	-2	TraesCS5A01G107600	Transducin/WD40 repeat protein
	-1	TraesCS5A01G107700	GRF1-interacting factor-like protein
	0	TraesCS5A01G107800	GDP-L-galactose phosphorylase 1
	1	TraesCS5A01G107900	Calcium-dependent protein kinase
	2	TraesCS5A01G108000	Fructose-bisphosphate aldolase
	3	TraesCS5A01G108100	HSP20-like chaperones superfamily protein
	4	TraesCS5A01G108200	F-box/LRR-repeat protein 3
	5	TraesCS5A01G108300	Isopentenyl-diphosphate delta-isomerase
	6	TraesCS5A01G108400	Chaperone protein DnaJ
	7	TraesCS5A01G108500	Cyclin family protein
	8	TraesCS5A01G108600	WWE protein-protein interaction domain protein family
	9	TraesCS5A01G108700	Nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase family protein
	10	TraesCS5A01G108800	phragmoplast orienting kinesin 1
<i>TaGGP1-B</i>	-10	TraesCS5B01G110700	AT-rich interactive domain-containing protein 2
	-9	TraesCS5B01G110800	Diacylglycerol kinase
	-8	TraesCS5B01G110900	Methyltransferase
	-7	TraesCS5B01G111000	Ubiquitin carboxyl-terminal hydrolase, putative
	-6	TraesCS5B01G111100	mediator of RNA polymerase II transcription subunit-like protein
	-5	TraesCS5B01G111200	Chaperone DnaK
	-4	TraesCS5B01G111300	Exocyst complex component Sec10, putative, expressed
	-3	TraesCS5B01G111400	Adenylate kinase
	-2	TraesCS5B01G111500	RNA-directed DNA polymerase (Reverse transcriptase); Ribonuclease H
	-1	TraesCS5B01G111600	Protein transport protein Sec24-like family
	0	TraesCS5B01G111700	GDP-L-galactose phosphorylase 1
	1	TraesCS5B01G111800	Thioredoxin
	2	TraesCS5B01G111900	E3 ubiquitin-protein ligase RGLG2
	3	TraesCS5B01G112000	SPFH/Band 7/PHB domain-containing membrane-associated protein family
	4	TraesCS5B01G112100	Cyclin T1
	5	TraesCS5B01G112200	30S ribosomal protein S4
	6	TraesCS5B01G112300	Photosystem II reaction center PsbP family protein
	7	TraesCS5B01G112400	phospholipid:diacylglycerol acyltransferase
	8	TraesCS5B01G112500	ATP-dependent Clp protease ATP-binding subunit

	9	TraesCS5B01G112600	SGF29 tudor-like domain-containing protein
	10	TraesCS5B01G112700	Allantoinase
<i>TaGGPI-D</i>	-10	TraesCS5D01G121800	Nucleotide-binding protein PFLU_0879
	-9	TraesCS5D01G121900	Cyclin family protein
	-8	TraesCS5D01G122000	E3 ubiquitin-protein ligase RGLG2
	-7	TraesCS5D01G122200	RNA-directed DNA polymerase (Reverse transcriptase)
	-6	TraesCS5D01G122300	Aminopeptidase
	-5	TraesCS5D01G122400	Protein kinase, putative
	-4	TraesCS5D01G122500	DNA repair protein RadA
	-3*	TraesCS5D01G122600	Ribosome maturation factor RimP
	-2*	TraesCS5D01G122700	Fructose-bisphosphate aldolase
	-1	TraesCS5D01G122800	Calcium-dependent protein kinase
	0	TraesCS5D01G122900	GDP-L-galactose phosphorylase 1
	1	TraesCS5D01G123000	Protein reticulata-related 1, chloroplastic
	2	TraesCS5D01G123100	U-box domain-containing protein
	3	TraesCS5D01G123200	Quinolate synthase A
	4	TraesCS5D01G123300	Ubiquitin carboxyl-terminal hydrolase, putative
	5	TraesCS5D01G123400	HAUS augmin-like complex subunit 6
	6	TraesCS5D01G123500	Regulator of chromosome condensation (RCC1) family protein
	7	TraesCS5D01G123600	CCR4-NOT transcription complex subunit 11
	8	TraesCS5D01G123700	Histone-lysine N-methyltransferase setd3
	9	TraesCS5D01G123800	Katanin p60 atpase-containing subunit a1
	10	TraesCS5D01G123900	tRNA-specific 2-thiouridylase MnmA
<i>TaGGP2-A</i>	-10	TraesCS4A01G200200	Two-component response regulator-like protein
	-9	TraesCS4A01G200300	5'-AMP-activated protein kinase beta-2 subunit protein
	-8	TraesCS4A01G200400	Apoptosis-inducing factor
	-7	TraesCS4A01G200500	Pentatricopeptide repeat-containing protein family
	-6	TraesCS4A01G200600	Digalactosyldiacylglycerol synthase 1, chloroplastic
	-5	TraesCS4A01G200700	Transcription factor-related family protein
	-4	TraesCS4A01G200800	BEL1-like homeodomain protein 1
	-3	TraesCS4A01G200900	cysteine-rich RLK (RECEPTOR-like protein kinase) 40
	-2	TraesCS4A01G201000	Phox-associated domain,Phox-like,Sorting nexin isoform 3
	-1	TraesCS4A01G201100	B3 domain-containing protein
	0	TraesCS4A01G201200	GDP-L-galactose phosphorylase 2
	1	TraesCS4A01G201300	Phosphatidylinositol 4-kinase
	2	TraesCS4A01G201400	Cysteine desulfurase
	3	TraesCS4A01G201600	E3 ubiquitin-protein ligase
	4	TraesCS4A01G201700	Peroxidase
	5	TraesCS4A01G201800	Receptor lectin kinase
	6	TraesCS4A01G201900	Fusarium resistance orphan protein
	7	TraesCS4A01G202000	Expressed protein-RZ53
	8	TraesCS4A01G202100	Alcohol dehydrogenase
	9	TraesCS4A01G202200	Alcohol dehydrogenase

<i>TaGGP2-B</i>	10	TraesCS4A01G202300	Alcohol dehydrogenase
	-10	TraesCS4B01G104300	Transcriptional corepressor SEUSS
	-9	TraesCS4B01G104400	Transcriptional corepressor SEUSS
	-8	TraesCS4B01G104500	Endoglucanase
	-7	TraesCS4B01G104600	Transcription factor, putative
	-6	TraesCS4B01G104700	Ninja-family protein
	-5	TraesCS4B01G104800	Exosome complex component RRP41
	-4	TraesCS4B01G104900	Kinase
	-3	TraesCS4B01G105000	exosome complex exonuclease
	-2	TraesCS4B01G105100	Heavy-metal-associated domain-containing protein, putative
	-1	TraesCS4B01G105200	MYB transcription factor
	0	TraesCS4B01G105300	GDP-L-galactose phosphorylase 2
	1	TraesCS4B01G105400	Phosphatidylinositol 4-kinase
	2	TraesCS4B01G105600	Cysteine desulfurase
	3	TraesCS4B01G105700	E3 ubiquitin-protein ligase
	4	TraesCS4B01G105800	Peroxidase
	5	TraesCS4B01G105900	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A-like protein 1
	6	TraesCS4B01G106000	Protein kinase
	7	TraesCS4B01G106100	Fusarium resistance orphan protein
	8	TraesCS4B01G106200	Expressed protein-RZ53
	9	TraesCS4B01G106300	Alcohol dehydrogenase
	10	TraesCS4B01G106400	Alcohol dehydrogenase
<i>TaGGP2-D</i>	-10	TraesCS4D01G101100	Coiled-coil domain-containing protein 18, putative isoform 1
	-9	TraesCS4D01G101200	Coiled-coil domain-containing protein 18, putative isoform 1
	-8	TraesCS4D01G101400	Myb/SANT-like DNA-binding domain protein
	-7	TraesCS4D01G101500	Transcriptional corepressor SEUSS
	-6	TraesCS4D01G101600	Premnaspirodiene oxygenase
	-5	TraesCS4D01G101700	Transcriptional corepressor SEUSS
	-4	TraesCS4D01G101800	protein kinase family protein
	-3	TraesCS4D01G101900	exosome complex exonuclease
	-2	TraesCS4D01G102000	Heavy-metal-associated domain-containing protein, putative, expressed
	-1	TraesCS4D01G102100	MYB transcription factor
	0	TraesCS4D01G102200	GDP-L-galactose phosphorylase 2
	1	TraesCS4D01G102300	Phosphatidylinositol 4-kinase
	2	TraesCS4D01G102400	Cysteine desulfurase
	3	TraesCS4D01G102500	E3 ubiquitin-protein ligase
	4	TraesCS4D01G102600	Peroxidase
	5	TraesCS4D01G102700	Kinase, putative
	6	TraesCS4D01G102800	Fusarium resistance orphan protein
	7	TraesCS4D01G102900	Expressed protein-RZ53
	8	TraesCS4D01G103000	Alcohol dehydrogenase
	9	TraesCS4D01G103100	Alcohol dehydrogenase

10	TraesCS4D01G103200	SWAP (Suppressor-of-White-APricot)/surp domain-containing protein
----	--------------------	---

*Both genes were detected as sense-antisense gene pair.

Table S4.2. Ten high confidence genes upstream and downstream of the two *BdGGP* genes.

SR	Position	Gene ID	Gene Function
<i>BdGGP1</i>	-10	bradi4g40840	unknown
	-9	bradi4g40830	BTB/POZ domain-containing protein
	-8	bradi4g40810	Phosphoribosyltransferase-like
	-7	bradi4g40800	Thioredoxin M-type, chloroplastic
	-6	bradi4g40790	Armadillo-type fold domain containing protein
	-5	bradi4g40785	Tetratricopeptide repeat-like superfamily protein
	-4	bradi4g40780	Photosystem I reaction centre subunit N
	-3	bradi4g40770	YUCCA-like flavin monooxygenase
	-2	bradi4g40760	TPR domain protein
	-1	bradi4g40750	flavin-containing monooxygenase YUCCA10-like
	0	bradi4g40740	GDP-L-galactose phosphorylase 1
	1	bradi4g40727	putative disease resistance protein
	2	bradi4g40720	growth regulator related protein
	3	bradi4g40710	photosystem II reaction center PsbP family protein
	4	bradi4g40700	cysteine desulfurase-like
	5	bradi4g40690	cysteine desulfurase-like
	6	bradi4g40680	Peroxidase superfamily protein
	7	bradi4g40675	Syntaxin-related protein KNOLLE (Syntaxin 111)
	8	bradi4g40670	Phosphomethylpyrimidine kinase type-1 domain containing protein
	9	bradi4g40660	BREVIS RADIX-like
	10	bradi4g40647	vignain-like
<i>BdGGP2</i>	-10	bradi4g22777	Transposase, IS4-like domain containing protein
	-9	bradi4g22770	SEUSS transcriptional co-regulator
	-8	bradi4g22765	uncharacterised
	-7	bradi4g22760	STE_MEKK_ste11_MAP3K.20
	-6	bradi4g22750	unknown
	-5	bradi4g22745	reverse transcriptase
	-4	bradi4g22740	NB-ARC domain-containing protein
	-3	bradi4g22730	MYB transcritcion factor
	-2	bradi4g22715	flavin-containing monooxygenase YUCCA10-like
	-1	bradi4g22707	unknown
	0	bradi4g22700	GDP-L-galactose phosphorylase 2
	1	bradi4g22690	phosphatidylinositol kinase
	2	bradi4g22681	No apical meristem protein domain containing protein
	3	bradi4g22670	NifS-like protein
	4	bradi4g22660	Peroxidase superfamily protein
	5	bradi4g22658	unknown function
	6	bradi4g22656	unknown function
	7	bradi4g22653	uncharacterised

8	bradi4g22651	uncharacterised
9	bradi4g22650	transposon protein
10	bradi4g22645	uncharacterised

Table S4.3. Ten high confidence genes upstream and downstream of the two *HvGGP* genes.

SR	Position	Gene ID	Gene Function
<i>HvGGP1</i>	-10	HORVU5Hr1G032240	uncharacterised
	-9	HORVU5Hr1G032230	uncharacterised
	-8	HORVU5Hr1G032220	uncharacterised
	-7	HORVU5Hr1G032210	uncharacterised
	-6	HORVU5Hr1G032170	Zn-Finger MYM-type protein like
	-5	HORVU5Hr1G032100	uncharacterised
	-4	HORVU5Hr1G032050	iron binding dioxygenase
	-3	HORVU5Hr1G031990	2-oxoglutarate-dependent dioxygenase-like
	-2	HORVU5Hr1G031980	ubiquitin carboxyl-terminal hydrolase
	-1	HORVU5Hr1G031970	uncharacterised
	0	HORVU5Hr1G031870	GDP-L-galactose phosphorylase 1
	1	HORVU5Hr1G031770	uncharacterised
	2	HORVU5Hr1G031730	uncharacterised
	3	HORVU5Hr1G031720	uncharacterised
	4	HORVU5Hr1G031690	cyclin-dependent kinase inhibitor
	5	HORVU5Hr1G031650	RNA-directed DDNA polymerase
	6	HORVU5Hr1G031630	40S ribosomal protein S19 family
	7	HORVU5Hr1G031600	peroxisomal adenine nucleotide carrier
	8	HORVU5Hr1G031590	uncharacterised
	9	HORVU5Hr1G031560	neomenthol dehydrogenase-like
	10	HORVU5Hr1G031550	LIM domain containing protein
<i>HvGGP2</i>	-10	HORVU4Hr1G017170	SUPPRESSOR OF GAMMA RESPONSE 1-like
	-9	HORVU4Hr1G017160	transcriptional corepressor SEUSS-like
	-8	HORVU4Hr1G017140	transcriptional corepressor SEUSS-like
	-7	HORVU4Hr1G017110	uncharacterised
	-6	HORVU4Hr1G017100	STE_MEKK_ste11_MAP3K.3 - like
	-5	HORVU4Hr1G017090	Serine Carboxypeptidase
	-4	HORVU4Hr1G017080	interactor of constitutive active ROPs 2, chloroplastic
	-3	HORVU4Hr1G017070	heavy metal-associated isoprenylated plant protein 23
	-2	HORVU4Hr1G017060	protein GPR107-like
	-1	HORVU4Hr1G017040	MYB transcription factor
	0	HORVU4Hr1G017030	GDP-L-galactose phosphorylase 2
	1	HORVU4Hr1G017010	phosphatidylinositol 4-kinase beta 1-like
	2	HORVU4Hr1G017000	methyltransferases superfamily protein
	3	HORVU4Hr1G016990	uncharacterised
	4	HORVU4Hr1G016980	ATP-dependent zinc metalloprotease FTSH, chloroplastic
	5	HORVU4Hr1G016970	F-box domain containing protein
	6	HORVU4Hr1G016950	E3 ubiquitin-protein ligase
	7	HORVU4Hr1G016940	Peroxidase superfamily protein

8	HORVU4Hr1G016920	uncharacterised
9	HORVU4Hr1G016880	lectin protein kinase family protein
10	HORVU4Hr1G016860	uncharacterised

Table S4.4. The *GGP* genes utilized in the manuscript for sequence alignments and phylogenetic analyses.

Gene Name	Gene Identifier	Gene Description	Source Website
<i>AdGGP1</i>	Achn155031	GDP-L-galactose phosphorylase	http://bioinfo.bti.cornell.edu
<i>AetGGP1</i>	LOC109780373	GDP-L-galactose phosphorylase 2-like	https://www.ncbi.nlm.nih.gov
<i>AetGGP2</i>	LOC109741548	GDP-L-galactose phosphorylase 2-like	https://www.ncbi.nlm.nih.gov
<i>AtGGP1 (vtc2)</i>	At4g26850	GDP-L-galactose phosphorylase 1	https://www.ncbi.nlm.nih.gov
<i>AtGGP2 (vtc5)</i>	At5g55120	GDP-L-galactose phosphorylase VITAMIN C DEFECTIVE 5 (VTC5)	https://www.ncbi.nlm.nih.gov
<i>BdGGP1</i>	LOC100842186	GDP-L-galactose phosphorylase 2	https://www.ncbi.nlm.nih.gov
<i>BdGGP2</i>	LOC100843203	GDP-L-galactose phosphorylase 2	https://www.ncbi.nlm.nih.gov
<i>CsGGP</i>	LOC101205337	GDP-L-galactose phosphorylase 1	https://www.ncbi.nlm.nih.gov
<i>FvGGP</i>	LOC101314355	GDP-L-galactose phosphorylase 2	https://www.ncbi.nlm.nih.gov
<i>GmGGP</i>	LOC100800738	GDP-L-galactose phosphorylase 1	https://www.ncbi.nlm.nih.gov
<i>HvGGP1</i>	HORVU5Hr1G031870	Predicted protein	https://plants.ensembl.org
<i>HvGGP2</i>	HORVU4Hr1G017030	Predicted protein	https://plants.ensembl.org
<i>LsGGP1</i>	LOC111899617	GDP-L-galactose phosphorylase 2	https://www.ncbi.nlm.nih.gov
<i>LsGGP2</i>	LOC111907963	GDP-L-galactose phosphorylase 2-like	https://www.ncbi.nlm.nih.gov
<i>MdGGP</i>	LOC103427292	GDP-L-galactose phosphorylase 2-like	https://www.ncbi.nlm.nih.gov
<i>MtGGP</i>	LOC11423375	GDP-L-galactose phosphorylase 1	https://www.ncbi.nlm.nih.gov
<i>NbGGP</i>	Niben101Scf08035g00004.1	GDP-L-galactose phosphorylase 1	https://solgenomics.net
<i>OsGGP</i>	LOC4351698	GDP-L-galactose phosphorylase 2	https://www.ncbi.nlm.nih.gov
<i>PtGGP</i>	LOC7483618	GDP-L-galactose phosphorylase 2	https://www.ncbi.nlm.nih.gov
<i>SbGGP</i>	LOC8084736	GDP-L-galactose phosphorylase 2	https://www.ncbi.nlm.nih.gov
<i>SIGGP</i>	Solyc06g073320	GDP-L-galactose hexose-1-phosphate guanylyltransferase	https://solgenomics.net
<i>VvGGP</i>	LOC100253859	GDP-L-galactose phosphorylase 2	https://www.ncbi.nlm.nih.gov
<i>ZmGGP</i>	LOC100283852	VTC2	https://www.ncbi.nlm.nih.gov
<i>TaGGP1-A</i>	TraesCS5A02G107800	n/a	https://plants.ensembl.org
<i>TaGGP1-B</i>	TraesCS5B02G111700	n/a	https://plants.ensembl.org
<i>TaGGP1-D</i>	TraesCS5D02G122900	n/a	https://plants.ensembl.org
<i>TaGGP2-A</i>	TraesCS4A02G201200	n/a	https://plants.ensembl.org
<i>TaGGP2-B</i>	TraesCS4B02G105300	n/a	https://plants.ensembl.org
<i>TaGGP2-D</i>	TraesCS4D02G102200	n/a	https://plants.ensembl.org

Table S4.5. Primers used for the quantitative reverse transcription-PCR analysis of the six *TaGGP* and housekeeping genes.

Gene name	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	PCR product length (bp)	Annealing temperature (°C)
<i>TaGGP1-A</i>	GAAGCACGCTTTGAGGAAGT	CAGCTGAAGGGCCAACCA	121	63
<i>TaGGP1-B</i>	AAGCAGGCTCTGGGCGATG	GCTTCCGACAGAGAGACCAC	164	63
<i>TaGGP1-D</i>	TTGGGTGCCTTTGCAACTAT	GCAAGCACTGGAAACCAAA	222	60
<i>TaGGP2-A</i>	GTTCTCTTTCGCTTTGAGAAAGGT	CGGTCAAGGACACGTGGA	155	65
<i>TaGGP2-B</i>	AAGCATAAGTGGGAGGACAGA	ACGGTCAACACGGAACCTCA	156	66
<i>TaGGP2-D</i>	AGGCTGCAAGCCATACTTC	TTCCCTCTCTCAAACACCAAA	220	60
<i>TaActin</i>	GACAATGGAACCGGAATGGTC	GTGTGATGCCAGATTTTCTCCAT	236	60
<i>TaCyclophilin</i>	CAAGCCGCTGCACTACAAGG	AGGGGACGGTGCAGATGAA	227	60
<i>TaELF</i>	CAGATTGGCAACGGCTACG	CGGACAGCAAAACGACCAAG	227	60

Chapter 5.

General Discussion

Ascorbate is an essential multifunctional compound for both plants and mammals, capable of donating electrons enabling it to act as a reactive oxygen species scavenger, function as an enzymatic co-factor, or reduce higher oxidation states of iron (Fe) to Fe²⁺ [5]. In plants, the role of ascorbate in photosynthetic functions and stress tolerance is well documented, but it has also recently been reported to facilitate the loading of Fe into seeds [7, 22, 23]. In humans, ascorbate is an essential micronutrient that must be obtained through diet due to loss-of-function mutations in the gene encoding the last enzymatic step of the vertebrate ascorbate biosynthetic pathway, and plays crucial roles in the formation of collagen, epigenetic programming, and the uptake of Fe in the gut [9, 39, 42, 192]. The *GDP-L-galactose phosphorylase* (*GGP* or *vtc2/5*) gene encodes the rate-limiting enzymatic step of the L-galactose pathway—the predominant ascorbate biosynthetic pathway in plants—and is the most promising candidate for increasing ascorbate levels in crops [192]. In addition to transcriptional regulation, the translation of *GGP* is regulated through a highly conserved *cis*-acting upstream open reading frame (uORF) in the 5' leader sequence of the *GGP* mRNA [17, 163]. Developing strategies to increase ascorbate levels in rice (*Oryza sativa* L.) and wheat (*Triticum aestivum* L.), two of the world's most important staple crops, has the potential to improve both food security and crop productivity through enhanced photosynthetic performance, antioxidant activity, and nutritional value. This PhD project employed two distinct metabolic engineering strategies to increase ascorbate levels in rice: (i) constitutive overexpression of the *OsGGP* coding sequence (Chapter 2), and (ii) CRISPR/Cas9-targeted mutagenesis of the *OsGGP* uORF (Chapter 3). This PhD project also identified and characterized the *TaGGP* gene family in the hexaploid bread wheat genome (Chapter 4).

Several studies have increased the expression of ascorbate biosynthetic or recycling genes to increase foliar ascorbate levels in rice, but whether these strategies affect ascorbate levels in the grain was unclear [64, 79, 138, 187]. Characterization of the 35S-*OsGGP* and *uorf_{OsGGP}* rice plants presented in Chapter 2 and Chapter 3 revealed that ascorbate levels were negligible in both 35S-*OsGGP* and *uorf_{OsGGP}* brown rice. This is most likely due to the decline of ascorbate levels in maturing grain reported in cereals [54, 55]. Overcoming the decline in ascorbate levels during grain maturation represents a significant challenge to ascorbate biofortification of cereal species, such as rice and wheat. Further understanding this decline in ascorbate concentrations in the maturing grains of cereals may help identify strategies to reduce this decline, and therefore present avenues to biofortify cereals with ascorbate, and ought to be the subject of future research. Alternative, future research in ascorbate biofortification may

benefit from focusing on crops that contain meaningful ascorbate levels at harvest (i.e. fresh fruits and vegetables).

Characterization of the 35S-*OsGGP* rice plants presented in Chapter 2 revealed that constitutive overexpression of the *OsGGP* CDS significantly increased ascorbate levels in germinated brown rice and tissues at the vegetative growth phase, but decreased ascorbate levels at the reproductive growth phase. The decrease in ascorbate levels at the reproductive growth phase was dependent on homozygosity of the 35S-*OsGGP* T-DNA and was associated with reduced transcript levels of the endogenous *OsGGP* gene, suggesting the occurrence of gene silencing. Similar results were reported when the kiwifruit (*Actinidia chinensis* L.) *AcGGP* gene was expressed in *Arabidopsis thaliana* [71]. Characterization of the *uorf_{OsGGP}* mutants presented in Chapter 3 revealed that CRISPR/Cas9-induced mutations in the *OsGGP* uORF significantly increased foliar ascorbate levels without affecting the transcript levels of the *OsGGP* gene. Elevated ascorbate levels in the *uorf_{OsGGP}* mutants are therefore likely due to alleviated ribosomal stalling on the *OsGGP* uORF and enhanced translation of the *GGP* mORF. Future research should utilize the dual-luciferase reporter system to confirm the effects of the mutated *OsGGP* uORFs on the translation of downstream mORFs [153, 154, 163, 239]. Editing the *OsGGP* uORF in rice offers many advantages over constitutive overexpression of the *OsGGP* CDS, for example by avoiding the occurrence of gene silencing or the random insertion of T-DNAs in the genome. Moreover, since the *uorf_{OsGGP}* mutants do not rely on the presence of foreign DNA, they are more likely to be accepted by the public and subject to less regulatory oversight. Disrupting the function of the *GGP* uORF with the CRISPR/Cas9 genome editing system is a promising transgene-free strategy to increase ascorbate biosynthesis not only in rice, but other crops.

Increased expression of ascorbate biosynthetic genes has been reported to enhance salt tolerance in a wide range of model and crop species. [64, 68, 70, 80, 89-91, 93, 99, 118, 187, 198, 252]. It was therefore hypothesized that the 35S-*OsGGP* and *uorf_{OsGGP}* rice plants may promote salt tolerance. However, characterization of the 35S-*OsGGP* and *uorf_{OsGGP}* rice plants presented in Chapter 2 and Chapter 3 revealed that neither the 35S-*OsGGP* nor the *uorf_{OsGGP}* rice plants had enhanced salt tolerance in terms of the relative growth rate of projected shoot area—a strong predictor of shoot biomass—in response to a final concentration of 90 mM NaCl at the vegetative growth phase. It is unclear why the 35S-*OsGGP* and *uorf_{OsGGP}* rice plants did not display enhanced salt tolerance given the established role of ascorbate in alleviating oxidative stress, however, it was recently reported that rice plants expressing the kiwifruit

AcGGP in rice only displayed enhanced salt tolerance in terms of visible symptoms and oxidative stress markers, but not shoot biomass [187]. Therefore, it may also be necessary to investigate oxidative stress markers, such as lipid peroxidation, in salt-stressed 35S-*OsGGP* and *uorfOsGGP* rice plants to determine whether the increased ascorbate levels in these plants mitigate the damaging effects of salt-induced oxidative stress. Future research efforts should also focus on assessing the tolerance of the *uorfOsGGP* mutants to other stresses that limit global rice productivity besides salt stress, such as drought, flooding, and cold stress [175]. Assaying enzymatic activities, such as ascorbate peroxidase activity [133], or assessing photosynthetic performance parameters, such as non-photochemical bleaching [27, 304], CO₂ assimilation [139, 305], maximum photochemical efficiency [95, 124], and photosynthetic electron transport rates [305], could also prove fruitful in teasing apart whether the *uorfOsGGP* mutants have enhanced antioxidant or photoprotective activity.

Ascorbate is a powerful enhancer of Fe bioavailability in the human gut and can overcome Fe absorption inhibitors such as phytate and polyphenolic compounds [44, 182-185]. The ascorbate-enriched 35S-*OsGGP* germinated brown rice produced in Chapter 2 was used to test whether engineering ascorbate-enriched crops could be a suitable strategy to improve Fe bioavailability in human diets utilizing the Caco-2 cell line model—an accurate predictor of human Fe absorption [67]. The Fe bioavailability assessment revealed that the 35S-*OsGGP* germinated brown rice slightly increased ferritin levels in Caco-2 cells, albeit non-significantly, and that a greater increase in ascorbate levels would likely be required to counteract the high levels of phytate to reveal a significant effect. Ascorbate levels—but not Fe or phytate levels—were positively correlated with Caco-2 cell ferritin levels. Similar results were reported with an *in vitro* bioavailability study of Andean potato clones with varying Fe, ascorbate, and phenolic levels [201]. Taken together, these results indicate that increasing ascorbate levels in crops may be able to improve Fe bioavailability in human diets, but further *in vivo* studies of ascorbate-enriched crops on Fe bioavailability are needed to support this. While cereal species, such as rice and wheat, are most likely not suitable candidates due to the decline of ascorbate levels in the maturing grain, many other major global crops may be suitable (Table 5.1). In particular, sweet potato (*Ipomoea batatas* L.) and cassava (*Manihot esculenta*), the two vegetables currently targeted for Fe biofortification [181], could be suitable candidates since they both possess meaningful ascorbate levels at harvest to improve upon (Table 5.1). Sweet potato has a relatively low Fe to ascorbate ratio of 1:1.1, therefore increasing ascorbate levels

Table 5.1. Ascorbate and Fe concentrations^a of major global commodity crops^b including the molar ratio of Fe to ascorbate and the fold-increase in ascorbate required to achieve molar ratios of 1:5, 1:10, and 1:20 of Fe to ascorbate.

Crop	Ascorbate (mg/100 g FW)	Fe (mg/100 g FW)	Molar Ratio of Fe to Ascorbate	Fold-increase in ascorbate to achieve 1:5	Fold-increase in ascorbate to achieve 1:10	Fold-increase in ascorbate to achieve 1:20
Wheat	0.0	3.6	1:0.0	n/a	n/a	n/a
Rice	0.0	1.8	1:0.0	n/a	n/a	n/a
Barley	0.0	3.6	1:0.0	n/a	n/a	n/a
Sorghum	0.0	3.4	1:0.0	n/a	n/a	n/a
Eggplants	2.2	0.2	1:2.7	1.9	3.7	7.4
Sweet potatoes	2.4	0.6	1:1.1	4.5	9.0	18.0
Coconuts	3.3	2.4	1:0.4	13.1	26.1	52.2
Grapes	4.0	0.4	1:3.1	1.6	3.2	6.4
Apples	4.6	0.1	1:10.8	✓	✓	1.9
Soybeans	6.0	3.6	1:0.5	10.5	21.0	42.0
Maize	6.8 ^c	0.5	1:3.8	1.3	2.6	5.2
Onions	7.4	0.2	1:9.9	✓	✓	2.0
Watermelons	8.1	0.2	1:9.5	✓	1.1	2.1
Bananas	8.7	0.3	1:9.4	✓	1.1	2.1
Tomatoes	14.0	0.3	1:14.6	✓	✓	1.4
Yams	17.1	0.5	1:8.9	✓	1.1	2.2
Potatoes	19.7	0.8	1:6.9	✓	1.5	2.9
Cassava	20.6	0.3	1:21.5	✓	✓	✓
Oranges	53.2	0.1	1:150.0	✓	✓	✓

^aAccording to the USDA Food Composition Database [52]. ^bAccording to FAOSTAT (>50,000,000 tonnes) [57]. Excluding ‘sugar cane’, ‘oil palm fruit’, ‘vegetable, fresh’, ‘cucumbers and gherkins’, ‘rapeseed’, ‘seed cotton’, ‘cabbages and other brassicas’, and ‘mangoes, mangosteens, guavas’. ^cMaize ascorbate value is based on the sweet corn variety. n/a; not applicable due to negligible concentrations of ascorbate. ✓; molar ratio of Fe to ascorbate present.

could readily improve the ratio of Fe to ascorbate in order to enhance the bioavailability of Fe in this important staple crop (Table 5.1). Moreover, any increases in the density of Fe in sweet potato achieved through Fe biofortification programmes should also be accompanied by increases in ascorbate levels to promote the bioavailability of the Fe. Cassava already has a favourable ratio of Fe to ascorbate of 1:21.5 due to low Fe levels and relatively high ascorbate levels (Table 5.1). However, any increases in the density of Fe in cassava achieved through Fe biofortification programmes should again be accompanied by increases in ascorbate levels to maintain this favourable ratio of Fe to ascorbate. Disrupting the function of the uORF of the sweet potato and cassava *GGP* genes with the CRISPR/Cas9 genome editing system could present an effective transgene-free strategy to increase both ascorbate levels and the bioavailability of Fe in these important staple crops.

Ascorbate has recently been reported to be involved in the reduction and transport of Fe into the developing seeds of *Arabidopsis* and pea (*Pisum sativum* L.) plants [7, 218]. It was therefore hypothesized that the ascorbate-enriched *uorf_{OsGGP}* mutants may accumulate more Fe in the grain. However, characterization of the *uorf_{OsGGP}* mutants presented in Chapter 3 revealed that increased ascorbate levels due to CRISPR/Cas9-induced mutations in the *OsGGP* uORF did not affect Fe levels in *uorf_{OsGGP}-1* brown rice. It is possible that this mechanism may not be present in graminaceous species, such as rice. This could be tested by investigating whether grain Fe levels are altered in ascorbate-deficient mutants of graminaceous species, such as rice, and should be the subject of future research.

Constitutive overexpression of the rice *dehydroascorbate reductase 1* (*OsDHAR1*) and *L-galactono-1,4-lactone dehydrogenase* (*OsL-GalLDH*) genes in rice have been reported to improve agronomical traits such as grain yield and spikelet fertility [79, 138]. Characterization of the *uorf_{OsGGP}* mutants presented in Chapter 3 revealed that the elevated ascorbate levels in the *uorf_{OsGGP}-1* mutant were associated with a significant increase in grain weight per plant. The increase in grain weight per plant in the *uorf_{OsGGP}-1* mutant may be attributed to a significant increase in the number of panicles per plant, as well as a trend towards increased spikelet fertility, that was observed. Since this experiment was largely preliminary and confined to one *uorf_{OsGGP}* mutant line over one generation, additional experiments assessing the agro-morphological performance of all three *uorf_{OsGGP}* mutants over multiple generations, including under both

greenhouse and paddy field conditions, should be the focus of future research to determine whether increased grain weight per plant in the *uorf_{OsGGP}-1* mutant is reproducible.

One concern of the CRISPR/Cas9 genome editing system is the production of unwanted off-target mutations. Characterization of the *uorf_{OsGGP}* mutants presented in Chapter 3 revealed that an off-target mutation was detected in the off-target #4 site (LOC_Os06g02530) in the *uorf_{OsGGP}-1* mutant. The occurrence of off-target mutations are very seldom reported in rice, however they have been detected—particularly when sgRNAs with only a 1-bp mismatch to an off-target site are used, as was the case for the off-target mutation found in the *uorf_{OsGGP}-1* mutant [253, 254]. This result highlights the importance of utilizing high-fidelity sgRNAs when using the CRISPR/Cas9 genome editing system to reduce the occurrence of off-target mutations. Backcrossing the *uorf_{OsGGP}-1* mutant to eliminate the off-target mutation will need to be carried out to determine whether this off-target mutation was responsible for any of the phenotypes observed for the *uorf_{OsGGP}-1* mutant.

Finally, the first high-quality annotated reference sequence for the hexaploid bread wheat genome was recently released by the International Wheat Genome Sequencing Consortium [174], presenting an opportunity for researchers to further understand wheat biology and engage in genomics-assisted molecular breeding. The identification and characterization of the *TaGGP* gene family in the bread wheat genome presented in Chapter 4 revealed the presence of six *TaGGP* genes, each with a highly conserved uORF in the 5' leader sequence. Phylogenetic analyses demonstrated that the *TaGGP* genes separate into two distinct groups and highlighted a duplication event of the *GGP* gene in the *Brachypodium*/Triticeae lineage that would have occurred after the evolutionary divergence from rice approximately 33.5 million years ago but prior to the evolutionary divergence of *Brachypodium* and the Triticeae tribe approximately 24.5 million years ago [274]. A microsynteny analysis of the *TaGGP* genes revealed that the *TaGGP1* and *TaGGP2* subchromosomal regions have no shared synteny, suggesting that *GGP2* may have been duplicated via transposition in the ancestor of the *Brachypodium*/Triticeae lineage since other methods of gene duplication preserve gene synteny [275]. A transcript analysis of the *TaGGP* genes identified distinct expression patterns for the two groups of *TaGGP1* and *TaGGP2* genes with the *TaGGP1* homoeologs broadly expressed across different tissues and developmental stages and the *TaGGP2* homoeologs highly expressed in anthers relative to other tissues, suggesting that the *TaGGP2*

homoeologs and ascorbate may play an important role in anther development in wheat. Transient transformation of the *TaGGP* coding sequences in *Nicotiana benthamiana* significantly increased foliar ascorbate levels, confirming their role toward ascorbate biosynthesis *in planta*. The identification of the six *TaGGP* genes presents an opportunity to fine-tune ascorbate biosynthesis in wheat via CRISPR/Cas9-induced mutations in the *TaGGP* uORFs. As the CRISPR/Cas9 mutation frequency is much lower in wheat, relative to other species such as rice, it may be necessary to cross mutants in particular homoeologs with one another to combine the mutations [306]. These crosses could be accelerated by the use of ‘speed breeding’ which greatly shortens the generation time of wheat [307]. Future experiments on wheat plants with CRISPR/Cas9-induced mutations in the *TaGGP* uORFs should focus on assessing tolerance to stresses that limit global wheat productivity, such as drought, salinity, and heat stress [158], as well as agromorphological performance under greenhouse and field conditions. Disrupting the function of the *TaGGP* uORFs with the CRISPR/Cas9 genome editing system could represent an effective transgene-free strategy to improve the global productivity of wheat—one of the world’s most important staple crops.

Finally, it is worth noting that increasing ascorbate concentrations in crops does not come without its challenges [308]. For example, increased expression of ascorbate biosynthetic and recycling genes has been reported to reduce guard cell responsiveness [309], disturb embryo development [310], and induce parthenocarpic (seedless) fruit in tomato [73]. A fine balance must therefore be struck when increasing ascorbate concentrations in crops to minimize any deleterious effects on plant growth and development that may arise due to perturbed ascorbate concentrations. Generating mutant alleles of varying strength in the *GGP* uORF with the CRISPR/Cas9 genome editing system represents an attractive strategy to fine-tune ascorbate concentrations in crops for enhanced stress tolerance, whilst minimizing any pleiotropic effects on plant growth and development [153, 154, 239].

Bibliography

1. Ashraf, M.; Wu, L., Breeding for salinity tolerance in plants. *Crit Rev Plant Sci* **1994**, 13, (1), 17-42.
2. UN DESA Population Division, World Population Prospects 2019: Highlights. **2019**, ST/ESA/SER.A/423, 1-39.
3. Von Grebmer, K.; Saltzman, A.; Birol, E.; Wiesman, D.; Prasai, N.; Yin, S.; Yohannes, Y.; Menon, P.; Thompson, J.; Sonntag, A., *Synopsis: 2014 global hunger index: The challenge of hidden hunger*. Intl Food Policy Res Inst: 2014; Vol. 83.
4. Bouis, H. E.; Welch, R. M., Biofortification—a sustainable agricultural strategy for reducing micronutrient malnutrition in the global south. *Crop Sci* **2010**, 50, S20-S32.
5. Smirnoff, N., Ascorbic acid metabolism and functions: A comparison of plants and mammals. *Free Radic Biol Med* **2018**, 122, 116-129.
6. Smirnoff, N.; Wheeler, G. L., Ascorbic acid in plants: biosynthesis and function. *Crit Rev Plant Sci* **2000**, 19, (4), 267-290.
7. Grillet, L.; Ouerdane, L.; Flis, P.; Hoang, M. T. T.; Isaure, M.-P.; Lobinski, R.; Curie, C.; Mari, S., Ascorbate efflux as a new strategy for iron reduction and transport in plants. *J Biol Chem* **2014**, 289, (5), 2515-2525.
8. Peterkofsky, B., Ascorbate requirement for hydroxylation and secretion of procollagen: relationship to inhibition of collagen synthesis in scurvy. *Am J Clin Nutr* **1991**, 54, (6), 1135S-1140S.
9. Lane, D. J.; Richardson, D. R., The active role of vitamin C in mammalian iron metabolism: much more than just enhanced iron absorption! *Free Radic Biol Med* **2014**, 75, 69-83.
10. Ishikawa, T.; Maruta, T.; Yoshimura, K.; Smirnoff, N., Biosynthesis and regulation of ascorbic acid in plants. In *Antioxidants and Antioxidant Enzymes in Higher Plants*,

- Gupta, D. K.; Palma, J. M.; Corpas, F. J., Eds. Springer International Publishing: New York City, 2018; pp 163-179.
11. Yoshimura, K.; Ishikawa, T., Chemistry and metabolism of ascorbic acid in plants. In *Ascorbic Acid in Plant Growth, Development and Stress Tolerance*, Hossain, M. A.; Munné-Bosch, S.; Burritt, D. J.; Diaz-Vivancos, P.; Fujita, M.; Lorence, A., Eds. Springer International Publishing: New York City, 2017; pp 1-23.
 12. Dowdle, J.; Ishikawa, T.; Gatzek, S.; Rolinski, S.; Smirnoff, N., Two genes in *Arabidopsis thaliana* encoding GDP-L-galactose phosphorylase are required for ascorbate biosynthesis and seedling viability. *Plant J* **2007**, 52, (4), 673-689.
 13. Lim, B.; Smirnoff, N.; Cobbett, C. S.; Golz, J. F., Ascorbate-deficient *vtc2* mutants in *Arabidopsis* do not exhibit decreased growth. *Front Plant Sci* **2016**, 7, 1025.
 14. Höller, S.; Ueda, Y.; Wu, L.; Wang, Y.; Hajirezaei, M.-R.; Ghaffari, M.-R.; von Wirén, N.; Frei, M., Ascorbate biosynthesis and its involvement in stress tolerance and plant development in rice (*Oryza sativa* L.). *Plant Mol Biol* **2015**, 88, (6), 545-560.
 15. Baldet, P.; Bres, C.; Okabe, Y.; Mauxion, J.-P.; Just, D.; Bournonville, C.; Ferrand, C.; Mori, K.; Ezura, H.; Rothan, C., Investigating the role of vitamin C in tomato through TILLING identification of ascorbate-deficient tomato mutants. *Plant Biotechnol* **2013**, 30, (3), 309-314.
 16. Vidal-Meireles, A.; Neupert, J.; Zsigmond, L.; Rosado-Souza, L.; Kovács, L.; Nagy, V.; Galambos, A.; Fernie, A. R.; Bock, R.; Tóth, S. Z., Regulation of ascorbate biosynthesis in green algae has evolved to enable rapid stress-induced response via the *VTC2* gene encoding GDP-L-galactose phosphorylase. *New Phytol* **2017**, 214, (2), 668-681.
 17. Bulley, S.; Laing, W., The regulation of ascorbate biosynthesis. *Curr Opin Plant Biol* **2016**, 33, 15-22.
 18. Macknight, R. C.; Laing, W. A.; Bulley, S. M.; Broad, R. C.; Johnson, A. A.; Hellens, R. P., Increasing ascorbate levels in crops to enhance human nutrition and plant abiotic stress tolerance. *Curr Opin Biotechnol* **2017**, 44, 153-160.
 19. Zechmann, B., Subcellular distribution of ascorbate in plants. *Plant Signal Behav* **2011**, 6, (3), 360-363.
 20. Bartoli, C. G.; Buet, A.; Grozeff, G. G.; Galatro, A.; Simontacchi, M., Ascorbate-glutathione cycle and abiotic stress tolerance in plants. In *Ascorbic Acid in Plant Growth*,

- Development and Stress Tolerance*, Hossain, M. A.; Munné-Bosch, S.; Burritt, D. J.; Diaz-Vivancos, P.; Fujita, M.; Lorence, A., Eds. Springer: New York City, 2017; pp 177-200.
21. Foyer, C. H.; Noctor, G., Ascorbate and glutathione: the heart of the redox hub. *Plant Physiol* **2011**, 155, (1), 2-18.
 22. Mittler, R., Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci* **2002**, 7, (9), 405-410.
 23. Foyer, C. H.; Shigeoka, S., Understanding oxidative stress and antioxidant functions to enhance photosynthesis. *Plant Physiol* **2011**, 155, (1), 93-100.
 24. Demmig-Adams, B.; Adams III, W. W., The role of xanthophyll cycle carotenoids in the protection of photosynthesis. *Trends Plant Sci* **1996**, 1, (1), 21-26.
 25. Müller-Moulé, P.; Conklin, P. L.; Niyogi, K. K., Ascorbate deficiency can limit violaxanthin de-epoxidase activity in vivo. *Plant Physiol* **2002**, 128, (3), 970-977.
 26. Asada, K., The water-water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons. *Annu Rev Plant Biol* **1999**, 50, (1), 601-639.
 27. Tóth, S. Z.; Nagy, V.; Puthur, J. T.; Kovács, L.; Garab, G., The physiological role of ascorbate as photosystem II electron donor: protection against photoinactivation in heat-stressed leaves. *Plant Physiol* **2011**, 156, (1), 382-392.
 28. Tóth, S. Z.; Puthur, J. T.; Nagy, V.; Garab, G., Experimental evidence for ascorbate-dependent electron transport in leaves with inactive oxygen-evolving complexes. *Plant Physiol* **2009**, 149, (3), 1568-1578.
 29. Carvalho, M. d., Drought stress and reactive oxygen species. *Plant Signal Behav* **2008**, 3, 156-165.
 30. Noctor, G.; Mhamdi, A.; Foyer, C. H., The roles of reactive oxygen metabolism in drought: not so cut and dried. *Plant Physiol* **2014**, 164, (4), 1636-1648.
 31. Abogadallah, G. M., Insights into the significance of antioxidative defense under salt stress. *Plant Signal Behav* **2010**, 5, (4), 369-374.
 32. Suzuki, N.; Mittler, R., Reactive oxygen species and temperature stresses: a delicate balance between signaling and destruction. *Physiol Plant* **2006**, 126, (1), 45-51.
 33. Rocklin, A. M.; Tierney, D. L.; Kofman, V.; Brunhuber, N. M.; Hoffman, B. M.; Christoffersen, R. E.; Reich, N. O.; Lipscomb, J. D.; Que, L., Role of the nonheme Fe (II)

- center in the biosynthesis of the plant hormone ethylene. *Proc Natl Acad Sci U S A* **1999**, 96, (14), 7905-7909.
34. Debolt, S.; Melino, V.; Ford, C. M., Ascorbate as a biosynthetic precursor in plants. *Ann Bot* **2006**, 99, (1), 3-8.
 35. Fryer, M., The antioxidant effects of thylakoid vitamin E (α -tocopherol). *Plant Cell Environ* **1992**, 15, (4), 381-392.
 36. Potters, G.; De Gara, L.; Asard, H.; Horemans, N., Ascorbate and glutathione: guardians of the cell cycle, partners in crime? *Plant Physiol Biochem* **2002**, 40, (6-8), 537-548.
 37. Smirnoff, N., Botanical briefing: the function and metabolism of ascorbic acid in plants. *Ann Bot* **1996**, 78, (6), 661-669.
 38. Barth, C.; De Tullio, M.; Conklin, P. L., The role of ascorbic acid in the control of flowering time and the onset of senescence. *J Exp Bot* **2006**, 57, (8), 1657-1665.
 39. Drouin, G.; Godin, J.-R.; Pagé, B., The genetics of vitamin C loss in vertebrates. *Curr Genomics* **2011**, 12, (5), 371-378.
 40. Combs Jr, G. F., Vitamin C. In *The Vitamins: Fundamental Aspects in Nutrition and Health*, 4th ed.; Elsevier: Amsterdam, Netherlands, 2012; pp 233-259.
 41. Chisnall, M.; Macknight, R., Importance of vitamin C in human health and disease. In *Ascorbic Acid in Plant Growth, Development and Stress Tolerance*, Hossain, M. A.; Munné-Bosch, S.; Burritt, D. J.; Diaz-Vivancos, P.; Fujita, M.; Lorence, A., Eds. Springer International Publishing: New York City, 2017; pp 491-501.
 42. Camarena, V.; Wang, G., The epigenetic role of vitamin C in health and disease. *Cell Mol Life Sci* **2016**, 73, (8), 1645-1658.
 43. Siegenberg, D.; Baynes, R. D.; Bothwell, T. H.; Macfarlane, B. J.; Lamparelli, R. D.; Car, N. G.; MacPhail, P.; Schmidt, U.; Tal, A.; Mayet, F., Ascorbic acid prevents the dose-dependent inhibitory effects of polyphenols and phytates on nonheme-iron absorption. *Am J Clin Nutr* **1991**, 53, (2), 537-541.
 44. Jin, F.; Frohman, C.; Thannhauser, T. W.; Welch, R. M.; Glahn, R. P., Effects of ascorbic acid, phytic acid and tannic acid on iron bioavailability from reconstituted ferritin measured by an in vitro digestion–Caco-2 cell model. *Br J Nutr* **2008**, 101, (7), 972-981.
 45. Lynch, S. R.; Cook, J. D., Interaction of vitamin C and iron. *Ann N Y Acad Sci* **1980**, 355, (3), 33-44.

46. Lane, D. J.; Richardson, D. R., The active role of vitamin C in mammalian iron metabolism: much more than just enhanced iron absorption! *Free Radical Biol Med* **2014**, 75, 69-83.
47. Lane, D. J.; Chikhani, S.; Richardson, V.; Richardson, D. R., Transferrin iron uptake is stimulated by ascorbate *via* an intracellular reductive mechanism. *Biochim Biophys Acta* **2013**, 1833, (6), 1527-1541.
48. Troesch, B.; Hoeft, B.; McBurney, M.; Eggersdorfer, M.; Weber, P., Dietary surveys indicate vitamin intakes below recommendations are common in representative Western countries. *Br J Nutr* **2012**, 108, (4), 692-698.
49. Cahill, L.; Corey, P. N.; El-Sohehy, A., Vitamin C deficiency in a population of young Canadian adults. *Am J Epidemiol* **2009**, 170, (4), 464-471.
50. Hampl, J. S.; Taylor, C. A.; Johnston, C. S., Vitamin C deficiency and depletion in the United States: the third national health and nutrition examination survey, 1988 to 1994. *Am J Public Health* **2004**, 94, (5), 870-875.
51. Christie-David, D.; Gunton, J., Vitamin C deficiency and diabetes mellitus—easily missed? *Diabet Med* **2017**, 34, (2), 294-296.
52. USDA Food Composition Databases. Available online: <https://ndb.nal.usda.gov/ndb/> (accessed on 21 Aug 2019).
53. Lee, S. K.; Kader, A. A., Preharvest and postharvest factors influencing vitamin C content of horticultural crops. *Postharvest Biol Technol* **2000**, 20, (3), 207-220.
54. Every, D.; Griffin, W.; Wilson, P., Ascorbate oxidase, protein disulfide isomerase, ascorbic acid, dehydroascorbic acid and protein levels in developing wheat kernels and their relationship to protein disulfide bond formation. *Cereal Chem* **2003**, 80, (1), 35-39.
55. Paradiso, A.; De Pinto, M.; Locato, V.; De Gara, L., Galactone- γ -lactone-dependent ascorbate biosynthesis alters wheat kernel maturation. *Plant Biol* **2012**, 14, (4), 652-658.
56. World Health Organization, Iron deficiency anaemia: assessment, prevention and control: *A guide for programme managers*. **2001**, 1-114.
57. Food and Agriculture Organization of the United Nations FAOSTAT. Available online: <http://www.fao.org/faostat/en/#home> (accessed on 28 Aug 19).
58. Siendones, E.; González-Reyes, J. A.; Santos-Ocana, C.; Navas, P.; Córdoba, F., Biosynthesis of ascorbic acid in kidney bean. L-Galactono- γ -lactone dehydrogenase is an

- intrinsic protein located at the mitochondrial inner membrane. *Plant Physiol* **1999**, 120, (3), 907-912.
59. Maruta, T.; Yonemitsu, M.; Yabuta, Y.; Tamoi, M.; Ishikawa, T.; Shigeoka, S., *Arabidopsis* phosphomannose isomerase 1, but not phosphomannose isomerase 2, is essential for ascorbic acid biosynthesis. *J Biol Chem* **2008**, 283, (43), 28842-28851.
 60. Badejo, A. A.; Eltelib, H. A.; Fukunaga, K.; Fujikawa, Y.; Esaka, M., Increase in ascorbate content of transgenic tobacco plants overexpressing the acerola (*Malpighia glabra*) phosphomannomutase gene. *Plant Cell Physiol* **2009**, 50, (2), 423-428.
 61. Qian, W.; Yu, C.; Qin, H.; Liu, X.; Zhang, A.; Johansen, I. E.; Wang, D., Molecular and functional analysis of phosphomannomutase (PMM) from higher plants and genetic evidence for the involvement of PMM in ascorbic acid biosynthesis in *Arabidopsis* and *Nicotiana benthamiana*. *Plant J* **2007**, 49, (3), 399-413.
 62. Zhou, Y.; Tao, Q.; Wang, Z.; Fan, R.; Li, Y.; Sun, X.; Tang, K., Engineering ascorbic acid biosynthetic pathway in *Arabidopsis* leaves by single and double gene transformation. *Biol Plantarum* **2012**, 56, (3), 451-457.
 63. Sawake, S.; Tajima, N.; Mortimer, J. C.; Lao, J.; Ishikawa, T.; Yu, X.; Yamanashi, Y.; Yoshimi, Y.; Kawai-Yamada, M.; Dupree, P., KONJAC1 and 2 are key factors for GDP-mannose generation and affect L-ascorbic acid and glucomannan biosynthesis in *Arabidopsis*. *Plant Cell* **2015**, 27, (12), 3397-3409.
 64. Zhang, G.-Y.; Liu, R.-R.; Zhang, C.-Q.; Tang, K.-X.; Sun, M.-F.; Yan, G.-H.; Liu, Q.-Q., Manipulation of the rice L-galactose pathway: evaluation of the effects of transgene overexpression on ascorbate accumulation and abiotic stress tolerance. *PLoS One* **2015**, 10, (5), e0125870.
 65. Imai, T.; Ban, Y.; Yamamoto, T.; Moriguchi, T., Ectopic overexpression of peach *GDP-D-mannose pyrophosphorylase* and *GDP-D-mannose-3', 5'-epimerase* in transgenic tobacco. *Plant Cell Tissue Organ Cult* **2012**, 111, (1), 1-13.
 66. Badejo, A. A.; Tanaka, N.; Esaka, M., Analysis of GDP-D-mannose pyrophosphorylase gene promoter from acerola (*Malpighia glabra*) and increase in ascorbate content of transgenic tobacco expressing the acerola gene. *Plant Cell Physiol* **2008**, 49, (1), 126-132.

67. Li, X.; Ye, J.; Munir, S.; Yang, T.; Chen, W.; Liu, G.; Zheng, W.; Zhang, Y., Biosynthetic gene pyramiding leads to ascorbate accumulation with enhanced oxidative stress tolerance in tomato. *Int J Mol Sci* **2019**, 20, (7), 1558.
68. Ma, L.; Wang, Y.; Liu, W.; Liu, Z., Overexpression of an alfalfa GDP-mannose 3, 5-epimerase gene enhances acid, drought and salt tolerance in transgenic *Arabidopsis* by increasing ascorbate accumulation. *Biotechnol Lett* **2014**, 36, (11), 2331-2341.
69. Huang, M.; Xu, Q.; Deng, X.-X., L-Ascorbic acid metabolism during fruit development in an ascorbate-rich fruit crop chestnut rose (*Rosa roxburghii* Tratt). *J Plant Physiol* **2014**, 171, (14), 1205-1216.
70. Zhang, C.; Liu, J.; Zhang, Y.; Cai, X.; Gong, P.; Zhang, J.; Wang, T.; Li, H.; Ye, Z., Overexpression of *SIGMEs* leads to ascorbate accumulation with enhanced oxidative stress, cold, and salt tolerance in tomato. *Plant Cell Rep* **2011**, 30, (3), 389-398.
71. Bulley, S. M.; Rassam, M.; Hoser, D.; Otto, W.; Schünemann, N.; Wright, M.; MacRae, E.; Gleave, A.; Laing, W., Gene expression studies in kiwifruit and gene over-expression in *Arabidopsis* indicates that GDP-L-galactose guanyltransferase is a major control point of vitamin C biosynthesis. *J Exp Bot* **2009**, 60, (3), 765-778.
72. Ali, B.; Pantha, S.; Acharya, R.; Ueda, Y.; Wu, L.-B.; Ashrafuzzaman, M.; Ishizaki, T.; Wissuwa, M.; Bulley, S.; Frei, M., Enhanced ascorbate level improves multi-stress tolerance in a widely grown indica rice variety without compromising its agronomic characteristics. *J Plant Physiol* **2019**, 240, 152998.
73. Bulley, S.; Wright, M.; Rommens, C.; Yan, H.; Rassam, M.; Lin-Wang, K.; Andre, C.; Brewster, D.; Karunairetnam, S.; Allan, A. C., Enhancing ascorbate in fruits and tubers through over-expression of the L-galactose pathway gene GDP-L-galactose phosphorylase. *Plant Biotechnol J* **2012**, 10, (4), 390-397.
74. Wang, L.; Meng, X.; Yang, D.; Ma, N.; Wang, G.; Meng, Q., Overexpression of tomato *GDP-L-galactose phosphorylase* gene in tobacco improves tolerance to chilling stress. *Plant Cell Rep* **2014**, 33, (9), 1441-1451.
75. Gatzek, S.; Wheeler, G. L.; Smirnoff, N., Antisense suppression of L-galactose dehydrogenase in *Arabidopsis thaliana* provides evidence for its role in ascorbate synthesis and reveals light modulated L-galactose synthesis. *Plant J* **2002**, 30, (5), 541-553.

76. Landi, M.; Fambrini, M.; Basile, A.; Salvini, M.; Guidi, L.; Pugliesi, C., Overexpression of *L-galactono-1, 4-lactone dehydrogenase (L-GalLDH)* gene correlates with increased ascorbate concentration and reduced browning in leaves of *Lactuca sativa* L. after cutting. *Plant Cell Tissue Organ Cult* **2015**, 123, (1), 109-120.
77. Guo, X.; Liu, R. H.; Fu, X.; Sun, X.; Tang, K., Over-expression of l-galactono- γ -lactone dehydrogenase increases vitamin C, total phenolics and antioxidant activity in lettuce through bio-fortification. *Plant Cell Tissue Organ Cult* **2013**, 114, (2), 225-236.
78. Shi, S.; Ma, F.; Li, Y.; Feng, F.; Shang, Z., Overexpression of L-galactono-1, 4-lactone dehydrogenase (GLDH) in Lanzhou lily (*Lilium davidii* var. *unicolor*) via particle bombardment-mediated transformation. *In Vitro Cell Dev Biol Plant* **2012**, 48, (1), 1-6.
79. Liu, Y.; Yu, L.; Wang, R., Level of ascorbic acid in transgenic rice for L-galactono-1, 4-lactone dehydrogenase overexpressing or suppressed is associated with plant growth and seed set. *Acta Physiol Plant* **2011**, 33, (4), 1353-1363.
80. Liu, W.; An, H.-M.; Yang, M., Overexpression of *Rosa roxburghii* L-galactono-1, 4-lactone dehydrogenase in tobacco plant enhances ascorbate accumulation and abiotic stress tolerance. *Acta Physiol Plant* **2013**, 35, (5), 1617-1624.
81. Imai, T.; Niwa, M.; Ban, Y.; Hirai, M.; Ôba, K.; Moriguchi, T., Importance of the L-galactonolactone pool for enhancing the ascorbate content revealed by *L-galactonolactone dehydrogenase*-overexpressing tobacco plants. *Plant Cell Tissue Organ Cult* **2009**, 96, (1), 105-112.
82. Aboobucker, S. I.; Suza, W. P.; Lorence, A., Characterization of two *Arabidopsis* L-gulonono-1, 4-lactone oxidases, AtGulLO3 and AtGulLO5, involved in ascorbate biosynthesis. *React Oxyg Species (Apex)* **2017**, 4, (12), 389.
83. Duan, J.; Zhang, M.; Zhang, H.; Xiong, H.; Liu, P.; Ali, J.; Li, J.; Li, Z., *OsMIOX*, a myo-inositol oxygenase gene, improves drought tolerance through scavenging of reactive oxygen species in rice (*Oryza sativa* L.). *Plant Sci* **2012**, 196, 143-151.
84. Cronje, C.; George, G. M.; Fernie, A. R.; Bekker, J.; Kossmann, J.; Bauer, R., Manipulation of L-ascorbic acid biosynthesis pathways in *Solanum lycopersicum*: elevated GDP-mannose pyrophosphorylase activity enhances L-ascorbate levels in red fruit. *Planta* **2012**, 235, (3), 553-564.

85. Lisko, K. A.; Torres, R.; Harris, R. S.; Belisle, M.; Vaughan, M. M.; Jullian, B.; Chevone, B. I.; Mendes, P.; Nessler, C. L.; Lorence, A., Elevating vitamin C content via overexpression of *myo*-inositol oxygenase and L-gulonolactone oxidase in *Arabidopsis* leads to enhanced biomass and tolerance to abiotic stresses. *In Vitro Cell Dev Biol Plant* **2013**, 49, (6), 643-655.
86. Lorence, A.; Chevone, B. I.; Mendes, P.; Nessler, C. L., *myo*-Inositol oxygenase offers a possible entry point into plant ascorbate biosynthesis. *Plant Physiol* **2004**, 134, (3), 1200-1205.
87. Endres, S.; Tenhaken, R., Myoinositol oxygenase controls the level of myoinositol in *Arabidopsis*, but does not increase ascorbic acid. *Plant Physiol* **2009**, 149, (2), 1042-1049.
88. Agius, F.; González-Lamothe, R.; Caballero, J. L.; Muñoz-Blanco, J.; Botella, M. A.; Valpuesta, V., Engineering increased vitamin C levels in plants by overexpression of a D-galacturonic acid reductase. *Nat Biotechnol* **2003**, 21, (2), 177.
89. Hemavathi; Upadhyaya, C. P.; Young, K. E.; Akula, N.; soon Kim, H.; Heung, J. J.; Oh, O. M.; Aswath, C. R.; Chun, S. C.; Kim, D. H.; Park, S. W., Over-expression of strawberry D-galacturonic acid reductase in potato leads to accumulation of vitamin C with enhanced abiotic stress tolerance. *Plant Sci* **2009**, 177, (6), 659-667.
90. Hemavathi; Upadhyaya, C. P.; Akula, N.; Kim, H. S.; Jeon, J. H.; Ho, O. M.; Chun, S. C.; Kim, D. H.; Park, S. W., Biochemical analysis of enhanced tolerance in transgenic potato plants overexpressing D-galacturonic acid reductase gene in response to various abiotic stresses. *Mol Breeding* **2011**, 28, (1), 105-115.
91. Upadhyaya, C. P.; Venkatesh, J.; Gururani, M. A.; Asnin, L.; Sharma, K.; Ajappala, H.; Park, S. W., Transgenic potato overproducing L-ascorbic acid resisted an increase in methylglyoxal under salinity stress via maintaining higher reduced glutathione level and glyoxalase enzyme activity. *Biotechnol Lett* **2011**, 33, (11), 2297.
92. Lim, M. Y.; Pulla, R. K.; Park, J. M.; Harn, C. H.; Jeong, B. R., Over-expression of L-gulonolactone oxidase (*GLOase*) gene leads to ascorbate accumulation with enhanced abiotic stress tolerance in tomato. *In Vitro Cell Dev Biol Plant* **2012**, 48, (5), 453-461.

93. Cai, X.; Zhang, C.; Ye, J.; Hu, T.; Ye, Z.; Li, H.; Zhang, Y., Ectopic expression of *FaGalUR* leads to ascorbate accumulation with enhanced oxidative stress, cold, and salt tolerance in tomato. *Plant Growth Regul* **2015**, 76, (2), 187-197.
94. Amaya, I.; Osorio, S.; Martinez-Ferri, E.; Lima-Silva, V.; Doblas, V. G.; Fernández-Muñoz, R.; Fernie, A. R.; Botella, M. A.; Valpuesta, V., Increased antioxidant capacity in tomato by ectopic expression of the strawberry D-galacturonate reductase gene. *Biotechnol J* **2015**, 10, (3), 490-500.
95. Bao, G.; Zhuo, C.; Qian, C.; Xiao, T.; Guo, Z.; Lu, S., Co-expression of *NCED* and *ALO* improves vitamin C level and tolerance to drought and chilling in transgenic tobacco and stylo plants. *Plant Biotechnol J* **2016**, 14, (1), 206-214.
96. Chen, Z.; Qin, C.; Lin, L.; Zeng, X.; Zhao, Y.; He, S.; Lu, S.; Guo, Z., Overexpression of yeast arabinono-1, 4-lactone oxidase gene (*ALO*) increases tolerance to oxidative stress and Al toxicity in transgenic tobacco plants. *Plant Mol Biol Rep* **2015**, 33, (4), 806-818.
97. Radzio, J. A.; Lorence, A.; Chevone, B. I.; Nessler, C. L., L-Gulonolactone oxidase expression rescues vitamin C-deficient *Arabidopsis* (*vtc*) mutants. *Plant Mol Biol* **2003**, 53, (6), 837-844.
98. Jain, A. K.; Nessler, C. L., Metabolic engineering of an alternative pathway for ascorbic acid biosynthesis in plants. *Mol Breeding* **2000**, 6, (1), 73-78.
99. Hemavathi; Upadhyaya, C. P.; Akula, N.; Young, K. E.; Chun, S. C.; Kim, D. H.; Park, S. W., Enhanced ascorbic acid accumulation in transgenic potato confers tolerance to various abiotic stresses. *Biotechnol Lett* **2010**, 32, (2), 321-330.
100. Yoshimura, K.; Nakane, T.; Kume, S.; Shiomi, Y.; Maruta, T.; Ishikawa, T.; Shigeoka, S., Transient expression analysis revealed the importance of *VTC2* expression level in light/dark regulation of ascorbate biosynthesis in *Arabidopsis*. *Biosci Biotechnol Biochem* **2014**, 78, (1), 60-66.
101. Hoeberichts, F. A.; Vaeck, E.; Kiddle, G.; Coppens, E.; Van De Cotte, B.; Adamantidis, A.; Ormenese, S.; Foyer, C. H.; Zabeau, M.; Inzé, D., A temperature-sensitive mutation in the *Arabidopsis thaliana* phosphomannomutase gene disrupts protein glycosylation and triggers cell death. *J Biol Chem* **2008**, 283, (9), 5708-5718.

102. Conklin, P. L.; Norris, S. R.; Wheeler, G. L.; Williams, E. H.; Smirnoff, N.; Last, R. L., Genetic evidence for the role of GDP-mannose in plant ascorbic acid (vitamin C) biosynthesis. *Proc Natl Acad Sci U S A* **1999**, 96, (7), 4198-4203.
103. Conklin, P. L.; Pallanca, J. E.; Last, R. L.; Smirnoff, N., L-ascorbic acid metabolism in the ascorbate-deficient *Arabidopsis* mutant *vtc1*. *Plant Physiol* **1997**, 115, (3), 1277-1285.
104. Conklin, P. L.; Williams, E. H.; Last, R. L., Environmental stress sensitivity of an ascorbic acid-deficient *Arabidopsis* mutant. *Proc Natl Acad Sci U S A* **1996**, 93, (18), 9970-9974.
105. Wolucka, B. A.; Van Montagu, M., GDP-mannose 3', 5'-epimerase forms GDP-L-gulose, a putative intermediate for the *de novo* biosynthesis of vitamin C in plants. *J Biol Chem* **2003**, 278, (48), 47483-47490.
106. Gilbert, L.; Alhagdow, M.; Nunes-Nesi, A.; Quemener, B.; Guillon, F.; Bouchet, B.; Faurobert, M.; Gouble, B.; Page, D.; Garcia, V., GDP-D-mannose 3, 5-epimerase (GME) plays a key role at the intersection of ascorbate and non-cellulosic cell-wall biosynthesis in tomato. *Plant J* **2009**, 60, (3), 499-508.
107. Conklin, P. L.; Saracco, S. A.; Norris, S. R.; Last, R. L., Identification of ascorbic acid-deficient *Arabidopsis thaliana* mutants. *Genetics* **2000**, 154, (2), 847-856.
108. Laing, W. A.; Wright, M. A.; Cooney, J.; Bulley, S. M., The missing step of the L-galactose pathway of ascorbate biosynthesis in plants, an L-galactose guanylttransferase, increases leaf ascorbate content. *Proc Natl Acad Sci U S A* **2007**, 104, (22), 9534-9539.
109. Conklin, P. L.; Gatzek, S.; Wheeler, G. L.; Dowdle, J.; Raymond, M. J.; Rolinski, S.; Isupov, M.; Littlechild, J. A.; Smirnoff, N., *Arabidopsis thaliana* *VTC4* encodes L-galactose-1-P phosphatase, a plant ascorbic acid biosynthetic enzyme. *J Biol Chem* **2006**, 281, (23), 15662-15670.
110. Pineau, B.; Layoune, O.; Danon, A.; De Paepe, R., L-galactono-1, 4-lactone dehydrogenase is required for the accumulation of plant respiratory complex I. *J Biol Chem* **2008**, 283, (47), 32500-32505.
111. Linster, C. L.; Adler, L. N.; Webb, K.; Christensen, K. C.; Brenner, C.; Clarke, S. G., A second GDP-L-galactose phosphorylase in *Arabidopsis* en route to vitamin C covalent

- intermediate and substrate requirements for the conserved reaction. *J Biol Chem* **2008**, 283, (27), 18483-18492.
112. Linster, C. L.; Gomez, T. A.; Christensen, K. C.; Adler, L. N.; Young, B. D.; Brenner, C.; Clarke, S. G., *Arabidopsis VTC2* encodes a GDP-L-galactose phosphorylase, the last unknown enzyme in the Smirnoff-Wheeler pathway to ascorbic acid in plants. *J Biol Chem* **2007**, 282, (26), 18879-18885.
 113. Maruta, T.; Ichikawa, Y.; Mieda, T.; Takeda, T.; Tamoi, M.; Yabuta, Y.; Ishikawa, T.; Shigeoka, S., The contribution of *Arabidopsis* homologs of L-gulono-1, 4-lactone oxidase to the biosynthesis of ascorbic acid. *Biosci Biotechnol Biochem* **2010**, 74, (7), 1494-1497.
 114. Zhang, W.; Gruszewski, H. A.; Chevone, B. I.; Nessler, C. L., An *Arabidopsis* purple acid phosphatase with phytase activity increases foliar ascorbate. *Plant Physiol* **2008**, 146, (2), 431-440.
 115. Belgaroui, N.; Lacombe, B.; Rouached, H.; Hanin, M., Phytase overexpression in *Arabidopsis* improves plant growth under osmotic stress and in combination with phosphate deficiency. *Sci Rep* **2018**, 8, (1), 1137.
 116. Kanter, U.; Usadel, B.; Guerineau, F.; Li, Y.; Pauly, M.; Tenhaken, R., The inositol oxygenase gene family of *Arabidopsis* is involved in the biosynthesis of nucleotide sugar precursors for cell-wall matrix polysaccharides. *Planta* **2005**, 221, (2), 243-254.
 117. Ivanov Kavkova, E.; Blöchl, C.; Tenhaken, R., The *myo*-inositol pathway does not contribute to ascorbic acid synthesis. *Plant Biol* **2019**, 21, 95-102.
 118. Lim, M. Y.; Jeong, B. R.; Jung, M.; Harn, C. H., Transgenic tomato plants expressing strawberry *D-galacturonic acid reductase* gene display enhanced tolerance to abiotic stresses. *Plant Biotechnol Rep* **2016**, 10, (2), 105-116.
 119. Gallie, D. R., The role of L-ascorbic acid recycling in responding to environmental stress and in promoting plant growth. *J Exp Bot* **2012**, 64, (2), 433-443.
 120. Eastmond, P. J., *MONODEHYDROASCORBATE REDUCTASE4* is required for seed storage oil hydrolysis and postgerminative growth in *Arabidopsis*. *Plant Cell* **2007**, 19, (4), 1376-1387.
 121. Eltelib, H.; Fujikawa, Y.; Esaka, M., Overexpression of the acerola (*Malpighia glabra*) monodehydroascorbate reductase gene in transgenic tobacco plants results in increased ascorbate levels and enhanced tolerance to salt stress. *S Afr J Bot* **2012**, 78, 295-301.

122. Eltayeb, A. E.; Kawano, N.; Badawi, G. H.; Kaminaka, H.; Sanekata, T.; Shibahara, T.; Inanaga, S.; Tanaka, K., Overexpression of monodehydroascorbate reductase in transgenic tobacco confers enhanced tolerance to ozone, salt and polyethylene glycol stresses. *Planta* **2007**, 225, (5), 1255-1264.
123. Yin, L.; Wang, S.; Eltayeb, A. E.; Uddin, M. I.; Yamamoto, Y.; Tsuji, W.; Takeuchi, Y.; Tanaka, K., Overexpression of dehydroascorbate reductase, but not monodehydroascorbate reductase, confers tolerance to aluminum stress in transgenic tobacco. *Planta* **2010**, 231, (3), 609-621.
124. Li, F.; Wu, Q. Y.; Sun, Y. L.; Wang, L. Y.; Yang, X. H.; Meng, Q. W., Overexpression of chloroplastic monodehydroascorbate reductase enhanced tolerance to temperature and methyl viologen-mediated oxidative stresses. *Physiol Plant* **2010**, 139, (4), 421-434.
125. Kavitha, K.; George, S.; Venkataraman, G.; Parida, A., A salt-inducible chloroplastic monodehydroascorbate reductase from halophyte *Avicennia marina* confers salt stress tolerance on transgenic plants. *Biochimie* **2010**, 92, (10), 1321-1329.
126. Haroldsen, V. M.; Chi-Ham, C. L.; Kulkarni, S.; Lorence, A.; Bennett, A. B., Constitutively expressed DHAR and MDHAR influence fruit, but not foliar ascorbate levels in tomato. *Plant Physiol Biochem* **2011**, 49, (10), 1244-1249.
127. Yoshida, S.; Tamaoki, M.; Shikano, T.; Nakajima, N.; Ogawa, D.; Ioki, M.; Aono, M.; Kubo, A.; Kamada, H.; Inoue, Y., Cytosolic dehydroascorbate reductase is important for ozone tolerance in *Arabidopsis thaliana*. *Plant Cell Physiol* **2006**, 47, (2), 304-308.
128. Hao, Z.; Wang, X.; Zong, Y.; Wen, S.; Cheng, Y.; Li, H., Enzymatic activity and functional analysis under multiple abiotic stress conditions of a dehydroascorbate reductase gene derived from *Liriodendron Chinense*. *Environ Exp Bot* **2019**, 103850.
129. Wang, Z.; Xiao, Y.; Chen, W.; Tang, K.; Zhang, L., Increased vitamin C content accompanied by an enhanced recycling pathway confers oxidative stress tolerance in *Arabidopsis*. *J Integr Plant Biol* **2010**, 52, (4), 400-409.
130. Liu, F.; Guo, X.; Yao, Y.; Tang, W.; Zhang, W.; Cao, S.; Han, Y.; Liu, Y., Cloning and function characterization of two dehydroascorbate reductases from kiwifruit (*Actinidia chinensis* L.). *Plant Mol Biol Rep* **2016**, 34, (4), 815-826.
131. Ushimaru, T.; Nakagawa, T.; Fujioka, Y.; Daicho, K.; Naito, M.; Yamauchi, Y.; Nonaka, H.; Amako, K.; Yamawaki, K.; Murata, N., Transgenic *Arabidopsis* plants expressing the

- rice dehydroascorbate reductase gene are resistant to salt stress. *J Plant Physiol* **2006**, 163, (11), 1179-1184.
132. Wang, W.; Qiu, X.; Kim, H. S.; Yang, Y.; Hou, D.; Liang, X.; Kwak, S.-S., Molecular cloning and functional characterization of a sweetpotato chloroplast *IbDHAR3* gene in response to abiotic stress. *Plant Biotechnol Rep* **2019**, 1-11.
 133. Chen, Z.; Young, T. E.; Ling, J.; Chang, S.-C.; Gallie, D. R., Increasing vitamin C content of plants through enhanced ascorbate recycling. *Proc Natl Acad Sci U S A* **2003**, 100, (6), 3525-3530.
 134. Naqvi, S.; Zhu, C.; Farre, G.; Ramessar, K.; Bassie, L.; Breitenbach, J.; Conesa, D. P.; Ros, G.; Sandmann, G.; Capell, T., Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proc Natl Acad Sci U S A* **2009**, 106, (19), 7762-7767.
 135. Goo, Y.-M.; Chun, H. J.; Kim, T.-W.; Lee, C.-H.; Ahn, M.-J.; Bae, S.-C.; Cho, K.-J.; Chun, J.-A.; Chung, C.-H.; Lee, S.-W., Expressional characterization of dehydroascorbate reductase cDNA in transgenic potato plants. *J Plant Biol* **2008**, 51, (1), 35-41.
 136. Eltayeb, A. E.; Yamamoto, S.; Habora, M. E. E.; Yin, L.; Tsujimoto, H.; Tanaka, K., Transgenic potato overexpressing Arabidopsis cytosolic *AtDHAR1* showed higher tolerance to herbicide, drought and salt stresses. *Breed Sci* **2011**, 61, (1), 3-10.
 137. Qin, A.; Shi, Q.; Yu, X., Ascorbic acid contents in transgenic potato plants overexpressing two dehydroascorbate reductase genes. *Mol Biol Rep* **2011**, 38, (3), 1557-1566.
 138. Kim, Y.-S.; Kim, I.-S.; Bae, M.-J.; Choe, Y.-H.; Kim, Y.-H.; Park, H.-M.; Kang, H.-G.; Yoon, H.-S., Homologous expression of cytosolic dehydroascorbate reductase increases grain yield and biomass under paddy field conditions in transgenic rice (*Oryza sativa* L. *japonica*). *Planta* **2013**, 237, (6), 1613-1625.
 139. Chen, Z.; Gallie, D. R., Increasing tolerance to ozone by elevating foliar ascorbic acid confers greater protection against ozone than increasing avoidance. *Plant Physiol* **2005**, 138, (3), 1673-1689.

140. Kwon, S.-Y.; Choi, S.-M.; Ahn, Y.-O.; Lee, H.-S.; Lee, H.-B.; Park, Y.-M.; Kwak, S.-S., Enhanced stress-tolerance of transgenic tobacco plants expressing a human dehydroascorbate reductase gene. *J Plant Physiol* **2003**, 160, (4), 347-353.
141. Kwon, S.-Y.; Ahn, Y.-O.; Lee, H.-S.; Kwak, S.-S., Biochemical characterization of transgenic tobacco plants expressing a human dehydroascorbate reductase gene. *BMB Rep* **2001**, 34, (4), 316-321.
142. Lee, Y.-P.; Kim, S.-H.; Bang, J.-W.; Lee, H.-S.; Kwak, S.-S.; Kwon, S.-Y., Enhanced tolerance to oxidative stress in transgenic tobacco plants expressing three antioxidant enzymes in chloroplasts. *Plant Cell Rep* **2007**, 26, (5), 591-598.
143. Le Martret, B.; Poage, M.; Shiel, K.; Nugent, G. D.; Dix, P. J., Tobacco chloroplast transformants expressing genes encoding dehydroascorbate reductase, glutathione reductase, and glutathione-S-transferase, exhibit altered anti-oxidant metabolism and improved abiotic stress tolerance. *Plant Biotechnol J* **2011**, 9, (6), 661-673.
144. Li, Q.; Li, Y.; Li, C.; Yu, X., Enhanced ascorbic acid accumulation through overexpression of dehydroascorbate reductase confers tolerance to methyl viologen and salt stresses in tomato. *Czech J Genet Plant Breed* **2012**, 48, (2), 74-86.
145. Liu, F.; Guo, X.; Yao, Y.; Tang, W.; Zhang, W.; Cao, S.; Han, Y.; Liu, Y., Cloning and function characterization of two dehydroascorbate reductases from kiwifruit (*Actinidia chinensis* L.). *Plant Mol Biol Rep* **2015**, 1-12.
146. Eltayeb, A. E.; Kawano, N.; Badawi, G. H.; Kaminaka, H.; Sanekata, T.; Morishima, I.; Shibahara, T.; Inanaga, S.; Tanaka, K., Enhanced tolerance to ozone and drought stresses in transgenic tobacco overexpressing dehydroascorbate reductase in cytosol. *Physiol Plant* **2006**, 127, (1), 57-65.
147. Qin, A.; Huang, X.; Zhang, H.; Wu, J.; Yang, J.; Zhang, S., Overexpression of *PbDHAR2* from *Pyrus sinkiangensis* in transgenic tomato confers enhanced tolerance to salt and chilling stresses. *HortScience* **2015**, 50, (6), 789-796.
148. Zhang, W.; Lorence, A.; Gruszewski, H. A.; Chevone, B. I.; Nessler, C. L., AMR1, an Arabidopsis gene that coordinately and negatively regulates the mannose/l-galactose ascorbic acid biosynthetic pathway. *Plant Physiol* **2009**, 150, (2), 942-950.
149. Ye, J.; Li, W.; Ai, G.; Li, C.; Liu, G.; Chen, W.; Wang, B.; Wang, W.; Lu, Y.; Zhang, J., Genome-wide association analysis identifies a natural variation in basic helix-loop-helix

- transcription factor regulating ascorbate biosynthesis via D-mannose/L-galactose pathway in tomato. *PLoS Genet* **2019**, 15, (5), e1008149.
150. Wang, J.; Yu, Y.; Zhang, Z.; Quan, R.; Zhang, H.; Ma, L.; Deng, X. W.; Huang, R., *Arabidopsis* CSN5B interacts with VTC1 and modulates ascorbic acid synthesis. *Plant Cell* **2013**, 25, (2), 625-636.
 151. Cai, X.; Zhang, C.; Shu, W.; Ye, Z.; Li, H.; Zhang, Y., The transcription factor SlDof22 involved in ascorbate accumulation and salinity stress in tomato. *Biochem Biophys Res Commun* **2016**, 474, (4), 736-741.
 152. Zhang, Z.; Wang, J.; Zhang, R.; Huang, R., The ethylene response factor AtERF98 enhances tolerance to salt through the transcriptional activation of ascorbic acid synthesis in *Arabidopsis*. *Plant J* **2012**, 71, (2), 273-287.
 153. Zhang, H.; Si, X.; Ji, X.; Fan, R.; Liu, J.; Chen, K.; Wang, D.; Gao, C., Genome editing of upstream open reading frames enables translational control in plants. *Nat Biotechnol* **2018**, 36, 894-898.
 154. Li, T.; Yang, X.; Yu, Y.; Si, X.; Zhai, X.; Zhang, H.; Dong, W.; Gao, C.; Xu, C., Domestication of wild tomato is accelerated by genome editing. *Nat Biotechnol* **2018**, 36, 1160-1163.
 155. Hu, T.; Ye, J.; Tao, P.; Li, H.; Zhang, J.; Zhang, Y.; Ye, Z., The tomato HD-Zip I transcription factor SlHZ24 modulates ascorbate accumulation through positive regulation of the d-mannose/l-galactose pathway. *Plant J* **2016**, 85, (1), 16-29.
 156. Xing, C.; Liu, Y.; Zhao, L.; Zhang, S.; Huang, X., A novel MYB transcription factor regulates ascorbic acid synthesis and affects cold tolerance. *Plant Cell Environ* **2019**, 42, (3), 832-845.
 157. Ye, J.; Liu, G.; Chen, W.; Zhang, F.; Li, H.; Ye, Z.; Zhang, Y., Knockdown of *SINL33* accumulates ascorbate, enhances disease and oxidative stress tolerance in tomato (*Solanum lycopersicum*). *Plant Growth Regul* **2019**, 1-10.
 158. Gangadhar, B. H.; Sajeesh, K.; Venkatesh, J.; Baskar, V.; Abhinandan, K.; Yu, J. W.; Prasad, R.; Mishra, R. K., Enhanced tolerance of transgenic potato plants over-expressing non-specific lipid transfer protein-1 (*StnsLTP1*) against multiple abiotic stresses. *Front Plant Sci* **2016**, 7, 1228.

159. Conklin, P. L.; DePaolo, D.; Wintle, B.; Schatz, C.; Buckenmeyer, G., Identification of *Arabidopsis* VTC3 as a putative and unique dual function protein kinase:: protein phosphatase involved in the regulation of the ascorbic acid pool in plants. *J Exp Bot* **2013**, 64, (10), 2793-2804.
160. Zhu, L.; Guo, J.; Zhu, J.; Zhou, C., Enhanced expression of *EsWAXI* improves drought tolerance with increased accumulation of cuticular wax and ascorbic acid in transgenic *Arabidopsis*. *Plant Physiol Biochem* **2014**, 75, 24-35.
161. Li, Y.; Chu, Z.; Luo, J.; Zhou, Y.; Cai, Y.; Lu, Y.; Xia, J.; Kuang, H.; Ye, Z.; Ouyang, B., The C2H2 zinc-finger protein SlZF3 regulates AsA synthesis and salt tolerance by interacting with CSN 5B. *Plant Biotechnol J* **2018**, 16, (6), 1201-1213.
162. Cho, K. M.; Nguyen, H. T. K.; Kim, S. Y.; Shin, J. S.; Cho, D. H.; Hong, S. B.; Shin, J. S.; Ok, S. H., CML 10, a variant of calmodulin, modulates ascorbic acid synthesis. *New Phytol* **2016**, 209, (2), 664-678.
163. Laing, W. A.; Martínez-Sánchez, M.; Wright, M. A.; Bulley, S. M.; Brewster, D.; Dare, A. P.; Rassam, M.; Wang, D.; Storey, R.; Macknight, R. C.; Hellens, R. P., An upstream open reading frame is essential for feedback regulation of ascorbate biosynthesis in *Arabidopsis*. *Plant Cell* **2015**, 27, (3), 772-786.
164. Broad, R. C.; Bonneau, J. P.; Beasley, J. T.; Roden, S.; Philips, J. G.; Baumann, U.; Hellens, R. P.; Johnson, A. A., Genome-wide identification and characterization of the *GDP-L-galactose phosphorylase* gene family in bread wheat. *BMC Plant Biol* **2019**, 19, (1), 1-15.
165. Zhang, Y.; Han, L.; Ye, Z.; Li, H., Ascorbic acid accumulation is transcriptionally modulated in high-pigment-1 tomato fruit. *Plant Mol Biol Rep* **2014**, 32, (1), 52-61.
166. Wang, S.; Liu, J.; Feng, Y.; Niu, X.; Giovannoni, J.; Liu, Y., Altered plastid levels and potential for improved fruit nutrient content by downregulation of the tomato DDB1-interacting protein CUL4. *Plant J* **2008**, 55, (1), 89-103.
167. Gangadhar, B. H.; Yu, J. W.; Sajeesh, K.; Park, S. W., A systematic exploration of high-temperature stress-responsive genes in potato using large-scale yeast functional screening. *Mol Genet Genomics* **2014**, 289, (2), 185-201.

168. Ouyang, B.; Yang, T.; Li, H.; Zhang, L.; Zhang, Y.; Zhang, J.; Fei, Z.; Ye, Z., Identification of early salt stress response genes in tomato root by suppression subtractive hybridization and microarray analysis. *J Exp Bot* **2007**, 58, (3), 507-520.
169. Bonneau, J.; O'Brien, M.; Plett, D. C.; Johnson, A. A., Genetic pathways important for iron nutrition and biofortification of bread wheat. *Ann Plant Rev Onl* **2018**, 1-35.
170. Rensink, W.; Buell, C. R., Arabidopsis to rice. Applying knowledge from a weed to enhance our understanding of a crop species. *Plant Physiol* **2004**, 135, (2), 622-629.
171. Project, I. R. G. S., The map-based sequence of the rice genome. *Nature* **2005**, 436, (7052), 793-800.
172. Hiei, Y.; Ishida, Y.; Komari, T., Progress of cereal transformation technology mediated by *Agrobacterium tumefaciens*. *Front Plant Sci* **2014**, 5, 628.
173. Mishra, R.; Joshi, R. K.; Zhao, K., Genome editing in rice: recent advances, challenges, and future implications. *Front Plant Sci* **2018**, 9, (1361).
174. IWGSC, Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. *Science* **2018**, 361, 661.
175. Lafitte, H.; Ismail, A.; Bennett, J., Abiotic stress tolerance in rice for Asia: progress and the future. In *Proceeding of 4th International Crop Science Congress, Brisbane, Australia*, 2004.
176. Abhinandan, K.; Skori, L.; Stanic, M.; Hickerson, N. M.; Jamshed, M.; Samuel, M. A., Abiotic stress signaling in wheat—an inclusive overview of hormonal interactions during abiotic stress responses in wheat. *Front Plant Sci* **2018**, 9, 734.
177. Zhang, J.; Klueva, N. Y.; Wang, Z.; Wu, R.; Ho, T.-H. D.; Nguyen, H. T., Genetic engineering for abiotic stress resistance in crop plants. *In Vitro Cell Dev Biol Plant* **2000**, 36, (2), 108-114.
178. Pandey, P.; Irulappan, V.; Bagavathiannan, M. V.; Senthil-Kumar, M., Impact of combined abiotic and biotic stresses on plant growth and avenues for crop improvement by exploiting physio-morphological traits. *Front Plant Sci* **2017**, 8, 537.
179. Koyro, H.-W.; Ahmad, P.; Geissler, N., Abiotic stress responses in plants: an overview. In *Environmental adaptations and stress tolerance of plants in the era of climate change*, Ahmad, P.; Prasad, M. N. V., Eds. Springer International Publishing: New York City, United States, 2012; pp 1-28.

180. von Grebmer, K.; Saltzman, A.; Birol, E.; Wiesman, D.; Prasai, N.; Yin, S.; Yohannes, Y.; Menon, P.; Thompson, J.; Sonntag, A., *2014 Global Hunger Index: The Challenge of Hidden Hunger*. International Food Policy Research Institute: Washington, D.C., United States, 2014.
181. Bouis, H. E.; Hotz, C.; McClafferty, B.; Meenakshi, J.; Pfeiffer, W. H., Biofortification: a new tool to reduce micronutrient malnutrition. *Food Nutr Bull* **2011**, 32, (1 supplement), S31-S40.
182. Siegenberg, D.; Baynes, R. D.; Bothwell, T. H.; Macfarlane, B. J.; Lamparelli, R. D.; Car, N.; MacPhail, P.; Schmidt, U.; Tal, A.; Mayet, F., Ascorbic acid prevents the dose-dependent inhibitory effects of polyphenols and phytates on nonheme-iron absorption. *Am J Clin Nutr* **1991**, 53, (2), 537-541.
183. Kim, E.-Y.; Ham, S.-K.; Bradke, D.; Ma, Q.; Han, O., Ascorbic acid offsets the inhibitory effect of bioactive dietary polyphenolic compounds on transepithelial iron transport in Caco-2 intestinal cells. *J Nutr* **2011**, 141, (5), 828-834.
184. Engle-Stone, R.; Yeung, A.; Welch, R.; Glahn, R., Meat and ascorbic acid can promote Fe availability from Fe- phytate but not from Fe- tannic acid complexes. *J Agric Food Chem* **2005**, 53, (26), 10276-10284.
185. Hallberg, L.; Brune, M.; Rossander, L., Iron absorption in man: ascorbic acid and dose-dependent inhibition by phytate. *Am J Clin Nutr* **1989**, 49, (1), 140-144.
186. Fairweather-Tait, S. J., The concept of bioavailability as it relates to iron nutrition. *Nutr Res* **1987**, 7, (3), 319-325.
187. Ali, B.; Pantha, S.; Acharya, R.; Ueda, Y.; Wu, L.-B.; Ashrafuzzaman, M.; Ishizaki, T.; Wissuwa, M.; Bulley, S.; Frei, M., Enhanced ascorbate level improves multi-stress tolerance in a widely grown *indica* rice variety without compromising its agronomic characteristics. *J Plant Physiol* **2019**, 152998.
188. Haas, J. D.; Beard, J. L.; Murray-Kolb, L. E.; Del Mundo, A. M.; Felix, A.; Gregorio, G. B., Iron-biofortified rice improves the iron stores of nonanemic Filipino women. *J Nutr* **2005**, 135, (12), 2823-2830.
189. Cho, D.-H.; Lim, S.-T., Germinated brown rice and its bio-functional compounds. *Food Chem* **2016**, 196, 259-271.

190. Patil, S. B.; Khan, M. K., Germinated brown rice as a value added rice product: A review. *J Food Sci Tech* **2011**, 48, (6), 661-667.
191. James, V.; Avart, C.; Worland, B.; Snape, J.; Vain, P., The relationship between homozygous and hemizygous transgene expression levels over generations in populations of transgenic rice plants. *Theor Appl Genet* **2002**, 104, (4), 553-561.
192. Macknight, R. C.; Laing, W. A.; Bulley, S. M.; Broad, R. C.; Johnson, A. A.; Hellens, R. P., Increasing ascorbate levels in crops to enhance human nutrition and plant abiotic stress tolerance. *Curr Opin Biotech* **2017**, 44, 153-160.
193. Si, X.; Zhang, H.; Wang, Y.; Chen, K.; Gao, C., Manipulating gene translation in plants by CRISPR–Cas9-mediated genome editing of upstream open reading frames. *Nature Protoc* **2020**, 1-26.
194. Al-Tamimi, N.; Brien, C.; Oakey, H.; Berger, B.; Saade, S.; Ho, Y. S.; Schmöckel, S. M.; Tester, M.; Negrao, S., Salinity tolerance loci revealed in rice using high-throughput non-invasive phenotyping. *Nat Commun* **2016**, 7, 13342.
195. Yichie, Y.; Brien, C.; Berger, B.; Roberts, T. H.; Atwell, B. J., Salinity tolerance in Australian wild *Oryza* species varies widely and matches that observed in *O. sativa*. *Rice* **2018**, 11, (1), 66.
196. Campbell, M. T.; Knecht, A. C.; Berger, B.; Brien, C. J.; Wang, D.; Walia, H., Integrating image-based phenomics and association analysis to dissect the genetic architecture of temporal salinity responses in rice. *Plant Physiol* **2015**, 168, (4), 1476-1489.
197. Hairmansis, A.; Berger, B.; Tester, M.; Roy, S. J., Image-based phenotyping for non-destructive screening of different salinity tolerance traits in rice. *Rice* **2014**, 7, (1), 16.
198. Lisko, K. A.; Torres, R.; Harris, R. S.; Belisle, M.; Vaughan, M. M.; Jullian, B.; Chevone, B. I.; Mendes, P.; Nessler, C. L.; Lorence, A., Elevating vitamin C content via overexpression of *myo*-inositol oxygenase and L-gulonolactone oxidase in *Arabidopsis* leads to enhanced biomass and tolerance to abiotic stresses. *In Vitro Cell Dev Biol Plant* **2013**, 49, (6), 643-655.
199. Glahn, R. P.; Lee, O. A.; Yeung, A.; Goldman, M. I.; Miller, D. D., Caco-2 cell ferritin formation predicts nonradiolabeled food iron availability in an in vitro digestion/Caco-2 cell culture model. *J Nutr* **1998**, 128, (9), 1555-1561.

200. Glahn, R. P.; Wortley, G. M.; South, P. K.; Miller, D. D., Inhibition of iron uptake by phytic acid, tannic acid, and ZnCl₂: studies using an in vitro digestion/Caco-2 cell model. *J Agric Food Chem* **2002**, 50, (2), 390-395.
201. Andre, C. M.; Evers, D. I.; Ziebel, J.; Guignard, C. d.; Hausman, J.-F.; Bonierbale, M.; zum Felde, T.; Burgos, G., In vitro bioaccessibility and bioavailability of iron from potatoes with varying vitamin C, carotenoid, and phenolic concentrations. *J Agric Food Chem* **2015**, 63, (41), 9012-9021.
202. Wei, Y.; Shohag, M.; Ying, F.; Yang, X.; Wu, C.; Wang, Y., Effect of ferrous sulfate fortification in germinated brown rice on seed iron concentration and bioavailability. *Food Chem* **2013**, 138, (2-3), 1952-1958.
203. Curtis, M. D.; Grossniklaus, U., A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* **2003**, 133, (2), 462-469.
204. Sallaud, C.; Meynard, D.; Van Boxtel, J.; Gay, C.; Bes, M.; Brizard, J.-P.; Larmande, P.; Ortega, D.; Raynal, M.; Portefaix, M., Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theor Appl Genet* **2003**, 106, (8), 1396-1408.
205. Rassam, M.; Laing, W., Variation in ascorbic acid and oxalate levels in the fruit of *Actinidia chinensis* tissues and genotypes. *J Agric Food Chem* **2005**, 53, (6), 2322-2326.
206. Nagy, T.; Kis, A.; Poliska, S.; Barta, E.; Havelda, Z.; Marincs, F., Comparison of small RNA next-generation sequencing with and without isolation of small RNA fraction. *BioTechniques* **2016**, 60, (6), 273-278.
207. Koressaar, T.; Remm, M., Enhancements and modifications of primer design program Primer3. *Bioinformatics* **2007**, 23, (10), 1289-1291.
208. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B. C.; Remm, M.; Rozen, S. G., Primer3—new capabilities and interfaces. *Nucleic Acids Res* **2012**, 40, (15), e115.
209. Andersen, C. L.; Jensen, J. L.; Ørntoft, T. F., Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* **2004**, 64, (15), 5245-5250.

210. Pham, A.-T.; Maurer, A.; Pillen, K.; Brien, C.; Dowling, K.; Berger, B.; Eglinton, J. K.; March, T. J., Genome-wide association of barley plant growth under drought stress using a nested association mapping population. *BMC Plant Biol* **2019**, 19, (1), 134.
211. Brien, C. J. growthPheno: plotting, smoothing and growth trait extraction for longitudinal data. R package version 1.0-13. <https://cran.at.r-project.org/package=growthPheno> (June 17, 2019),
212. Team, R. C. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing: Vienna, Austria, 2019.
213. Dias, D.; Kolba, N.; Binyamin, D.; Ziv, O.; Regini Nutti, M.; Martino, H.; Glahn, R.; Koren, O.; Tako, E., Iron biofortified carioca bean (*Phaseolus vulgaris* L.)—based brazilian diet delivers more absorbable iron and affects the gut microbiota in vivo (*Gallus gallus*). *Nutrients* **2018**, 10, (12), 1970.
214. Trijatmiko, K. R.; Dueñas, C.; Tsakirpaloglou, N.; Torrizo, L.; Arines, F. M.; Adeva, C.; Balindong, J.; Oliva, N.; Sapasap, M. V.; Borrero, J., Biofortified indica rice attains iron and zinc nutrition dietary targets in the field. *Sci Rep* **2016**, 6, 19792.
215. Beasley, J. T.; Bonneau, J. P.; Sánchez-Palacios, J. T.; Moreno-Moyano, L. T.; Callahan, D. L.; Tako, E.; Glahn, R. P.; Lombi, E.; Johnson, A. A., Metabolic engineering of bread wheat improves grain iron concentration and bioavailability. *Plant Biotechnol J* **2019**.
216. Butler, D. G.; Cullis, B. R.; Gilmour, A. R.; Gogel, B. J.; Thompson, R. ASReml-R reference manual. Version 4. <http://asreml.org>
217. Brien, C. J. asremlPlus: augments ASReml-R in fitting mixed models and packages generally in exploring prediction differences. R package version 4.1-28. <https://cran.at.r-project.org/package=asremlPlus> (June 17, 2019),
218. Curie, C.; Mari, S., New routes for plant iron mining. *New Phytol* **2017**, 214, (2), 521-525.
219. Hellens, R. P.; Brown, C. M.; Chisnall, M. A.; Waterhouse, P. M.; Macknight, R. C., The emerging world of small ORFs. *Trends Plant Sci* **2016**, 21, (4), 317-328.
220. van der Horst, S.; Filipovska, T.; Hanson, J.; Smeekens, S., Metabolite control of translation by conserved peptide uORFs: the ribosome as a metabolite multisensor. *Plant Physiol* **2020**, 182, (1), 110.

221. Guerrero-González, M. L.; Rodríguez-Kessler, M.; Jiménez-Bremont, J. F., uORF, a regulatory mechanism of the Arabidopsis polyamine oxidase 2. *Mol Biol Rep* **2014**, 41, (4), 2427-2443.
222. Hanfrey, C.; Elliott, K. A.; Franceschetti, M.; Mayer, M. J.; Illingworth, C.; Michael, A. J., A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. *J Biol Chem* **2005**.
223. Hanfrey, C.; Franceschetti, M.; Mayer, M. J.; Illingworth, C.; Michael, A. J., Abrogation of upstream open reading frame-mediated translational control of a plant S-adenosylmethionine decarboxylase results in polyamine disruption and growth perturbations. *J Biol Chem* **2002**.
224. Hu, W.-W.; Gong, H.; Pua, E. C., The pivotal roles of the plant S-adenosylmethionine decarboxylase 5' untranslated leader sequence in regulation of gene expression at the transcriptional and posttranscriptional levels. *Plant Physiol* **2005**, 138, (1), 276-286.
225. Imai, A.; Hanzawa, Y.; Komura, M.; Yamamoto, K. T.; Komeda, Y.; Takahashi, T., The dwarf phenotype of the Arabidopsis *acl5* mutant is suppressed by a mutation in an upstream ORF of a bHLH gene. *Development* **2006**, 133, (18), 3575-3585.
226. Imai, A.; Komura, M.; Kawano, E.; Kuwashiro, Y.; Takahashi, T., A semi-dominant mutation in the ribosomal protein L10 gene suppresses the dwarf phenotype of the *acl5* mutant in *Arabidopsis thaliana*. *Plant J* **2008**, 56, (6), 881-890.
227. Cai, Q.; Fukushima, H.; Yamamoto, M.; Ishii, N.; Sakamoto, T.; Kurata, T.; Motose, H.; Takahashi, T., The SAC51 family plays a central role in thermospermine responses in Arabidopsis. *Plant and Cell Physiology* **2016**, 57, (8), 1583-1592.
228. Ishitsuka, S.; Yamamoto, M.; Miyamoto, M.; Kuwashiro, Y.; Imai, A.; Motose, H.; Takahashi, T., Complexity and conservation of thermospermine-responsive uORFs of *SAC51* family genes in angiosperms. *Frontiers in plant science* **2019**, 10, 564.
229. Uchiyama-Kadokura, N.; Murakami, K.; Takemoto, M.; Koyanagi, N.; Murota, K.; Naito, S.; Onouchi, H., Polyamine-responsive ribosomal arrest at the stop codon of an upstream open reading frame of the *AdoMetDC1* gene triggers nonsense-mediated mRNA decay in *Arabidopsis thaliana*. *Plant Cell Physiol* **2014**, 55, (9), 1556-1567.

230. Guerrero-González, M. L.; Ortega-Amaro, M. A.; Juárez-Montiel, M.; Jiménez-Bremont, J. F., *Arabidopsis polyamine oxidase-2* uORF is required for downstream translational regulation. *Plant Physiol Biochem* **2016**, 108, 381-390.
231. Kwak, S.-H.; Lee, S. H., The regulation of ornithine decarboxylase gene expression by sucrose and small upstream open reading frame in tomato (*Lycopersicon esculentum* Mill). *Plant Cell Physiol* **2001**, 42, (3), 314-323.
232. Rahmani, F.; Hummel, M.; Schuurmans, J.; Wiese-Klinkenberg, A.; Smeekens, S.; Hanson, J., Sucrose control of translation mediated by an upstream open reading frame-encoded peptide. *Plant Physiol* **2009**, 150, (3), 1356-1367.
233. Thalor, S. K.; Berberich, T.; Lee, S. S.; Yang, S. H.; Zhu, X.; Imai, R.; Takahashi, Y.; Kusano, T., Dereglulation of sucrose-controlled translation of a bZIP-type transcription factor results in sucrose accumulation in leaves. *PLoS One* **2012**, 7, (3), e33111.
234. Wiese, A.; Elzinga, N.; Wobbes, B.; Smeekens, S., A conserved upstream open reading frame mediates sucrose-induced repression of translation. *Plant Cell* **2004**, 16, (7), 1717-1729.
235. Yamashita, Y.; Takamatsu, S.; Glasbrenner, M.; Becker, T.; Naito, S.; Beckmann, R., Sucrose sensing through nascent peptide-mediated ribosome stalling at the stop codon of *Arabidopsis bZIP 11* uORF 2. *FEBS Lett* **2017**, 591, (9), 1266-1277.
236. Alatorre-Cobos, F.; Cruz-Ramírez, A.; Hayden, C. A.; Pérez-Torres, C.-A.; Chauvin, A.-L.; Ibarra-Laclette, E.; Alva-Cortés, E.; Jorgensen, R. A.; Herrera-Estrella, L., Translational regulation of *Arabidopsis XIPOTL1* is modulated by phosphocholine levels via the phylogenetically conserved upstream open reading frame 30. *J Exp Bot* **2012**, 63, (14), 5203-5221.
237. Tabuchi, T.; Okada, T.; Azuma, T.; Nanmori, T.; Yasuda, T., Posttranscriptional regulation by the upstream open reading frame of the phosphoethanolamine N-methyltransferase gene. *Biosci Biotechnol Biochem* **2006**, 70, (60309), 1-5.
238. Zhu, X.; Li, Y.; Fang, W.; Kusano, T., Galactinol is involved in sequence-conserved upstream open reading frame-mediated repression of *Arabidopsis HsfB1* translation. *Environ Exp Bot* **2018**, 156, 120-129.

239. Si, X.; Zhang, H.; Wang, Y.; Chen, K.; Gao, C., Manipulating gene translation in plants by CRISPR–Cas9-mediated genome editing of upstream open reading frames. *Nat Protoc* **2020**, 15, (2), 338-363.
240. Lucht, J. M., Public acceptance of plant biotechnology and GM crops. *Viruses* **2015**, 7, (8), 4254-4281.
241. Naim, F.; Dugdale, B.; Kleidon, J.; Brinin, A.; Shand, K.; Waterhouse, P.; Dale, J., Gene editing the phytoene desaturase alleles of Cavendish banana using CRISPR/Cas9. *Transgenic Res* **2018**, 27, (5), 451-460.
242. Xie, K.; Minkenberg, B.; Yang, Y., Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc Natl Acad Sci U S A* **2015**, 112, (11), 3570-3575.
243. Zhang, H.; Zhang, J.; Wei, P.; Zhang, B.; Gou, F.; Feng, Z.; Mao, Y.; Yang, L.; Zhang, H.; Xu, N., The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol J* **2014**, 12, (6), 797-807.
244. Ishizaki, T., CRISPR/Cas9 in rice can induce new mutations in later generations, leading to chimerism and unpredicted segregation of the targeted mutation. *Mol Breeding* **2016**, 36, (12), 165.
245. Charneski, C. A.; Hurst, L. D., Positively charged residues are the major determinants of ribosomal velocity. *PLoS Biol* **2013**, 11, (3), e1001508.
246. Wilson, D. N.; Arenz, S.; Beckmann, R., Translation regulation via nascent polypeptide-mediated ribosome stalling. *Curr Opin Struct Biol* **2016**, 37, 123-133.
247. Seidelt, B.; Innis, C. A.; Wilson, D. N.; Gartmann, M.; Armache, J.-P.; Villa, E.; Trabuco, L. G.; Becker, T.; Mielke, T.; Schulten, K., Structural insight into nascent polypeptide chain-mediated translational stalling. *Science* **2009**, 326, (5958), 1412-1415.
248. Bischoff, L.; Berninghausen, O.; Beckmann, R., Molecular basis for the ribosome functioning as an L-tryptophan sensor. *Cell Rep* **2014**, 9, (2), 469-475.
249. Arenz, S.; Meydan, S.; Starosta, A. L.; Berninghausen, O.; Beckmann, R.; Vázquez-Laslop, N.; Wilson, D. N., Drug sensing by the ribosome induces translational arrest via active site perturbation. *Mol Cell* **2014**, 56, (3), 446-452.

250. Young, S. K.; Wek, R. C., Upstream open reading frames differentially regulate gene-specific translation in the integrated stress response. *J Biol Chem* **2016**, 291, (33), 16927-16935.
251. Hellens, R. P.; Allan, A. C.; Friel, E. N.; Bolitho, K.; Grafton, K.; Templeton, M. D.; Karunairetnam, S.; Gleave, A. P.; Laing, W. A., Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* **2005**, 1, (1), 13.
252. Lim, M. Y.; Pulla, R. K.; Park, J. M.; Harn, C. H.; Jeong, B. R., Over-expression of L-gulonolactone oxidase (GLOase) gene leads to ascorbate accumulation with enhanced abiotic stress tolerance in tomato. *In Vitro Cell Dev Biol Plant* **2012**, 48, (5), 453-461.
253. Tang, X.; Liu, G.; Zhou, J.; Ren, Q.; You, Q.; Tian, L.; Xin, X.; Zhong, Z.; Liu, B.; Zheng, X., A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol* **2018**, 19, (1), 84.
254. Endo, M.; Mikami, M.; Toki, S., Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell Physiol* **2015**, 56, (1), 41-47.
255. Mariño-Ramírez, L.; Kann, M. G.; Shoemaker, B. A.; Landsman, D., Histone structure and nucleosome stability. *Expert Rev Proteomic* **2005**, 2, (5), 719-729.
256. Duc, C.; Benoit, M.; Détourné, G.; Simon, L.; Poulet, A.; Jung, M.; Veluchamy, A.; Latrasse, D.; Le Goff, S.; Cotterell, S., Arabidopsis ATRX modulates H3.3 occupancy and fine-tunes gene expression. *Plant Cell* **2017**, 29, (7), 1773-1793.
257. Sanders, D.; Qian, S.; Fieweger, R.; Lu, L.; Dowell, J. A.; Denu, J. M.; Zhong, X., Histone lysine-to-methionine mutations reduce histone methylation and cause developmental pleiotropy. *Plant Physiol* **2017**, 173, (4), 2243-2252.
258. Zhao, W.; Neyt, P.; Van Lijsebettens, M.; Shen, W. H.; Berr, A., Interactive and noninteractive roles of histone H2B monoubiquitination and H3K36 methylation in the regulation of active gene transcription and control of plant growth and development. *New Phytol* **2019**, 221, (2), 1101-1116.
259. Foyer, C. H.; Noctor, G., Managing the cellular redox hub in photosynthetic organisms. *Plant Cell Environ* **2012**, 35, (2), 199-201.

260. Lisko, K. A.; Aboobucker, S. I.; Torres, R.; Lorence, A., Engineering elevated vitamin C in plants to improve their nutritional content, growth, and tolerance to abiotic stress. In *In: Jetter R, editor. Phytochemicals–Biosynthesis, Function and Application*, Springer International Publishing: New York City, 2014; pp 109-128.
261. De Tullio, M.; Arrigoni, O., Hopes, disillusion and more hopes from vitamin C. *Cell Mol Life Sci* **2004**, 61, (2), 209-219.
262. Horemans, N.; Foyer, C. H.; Potters, G.; Asard, H., Ascorbate function and associated transport systems in plants. *Plant Physiol Biochem* **2000**, 38, (7-8), 531-540.
263. Hayden, C. A.; Jorgensen, R. A., Identification of novel conserved peptide uORF homology groups in Arabidopsis and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. *BMC Biol* **2007**, 5, (1), 32.
264. Tran, M. K.; Schultz, C. J.; Baumann, U., Conserved upstream open reading frames in higher plants. *BMC Genomics* **2008**, 9, (1), 361.
265. Sanahuja, G.; Farré, G.; Bassie, L.; Zhu, C.; Christou, P.; Capell, T., Ascorbic acid synthesis and metabolism in maize are subject to complex and genotype-dependent feedback regulation during endosperm development. *Biotechnol J* **2013**, 8, (10), 1221-1230.
266. IWGSC, Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **2018**, 361, (6403), eaar7191.
267. Abhinandan, K.; Skori, L.; Stanic, M.; Hickerson, N. M.; Jamshed, M.; Samuel, M. A., Abiotic Stress Signaling in Wheat–An Inclusive Overview of Hormonal Interactions During Abiotic Stress Responses in Wheat. *Front Plant Sci* **2018**, 9.
268. Liu, C.; Atkinson, M.; Chinoy, C.; Devos, K.; Gale, M., Nonhomoeologous translocations between group 4, 5 and 7 chromosomes within wheat and rye. *Theor Appl Genet* **1992**, 83, (3), 305-312.
269. Akhunov, E. D.; Sehgal, S.; Liang, H.; Wang, S.; Akhunova, A. R.; Kaur, G.; Li, W.; Forrest, K. L.; See, D.; Šimková, H., Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. *Plant Physiol* **2013**, 161, (1), 252-265.

270. Brenner, C., Hint, Fhit, and GalT: function, structure, evolution, and mechanism of three branches of the histidine triad superfamily of nucleotide hydrolases and transferases. *Biochemistry* **2002**, 41, (29), 9003-9014.
271. Müller-Moulé, P., An expression analysis of the ascorbate biosynthesis enzyme VTC2. *Plant Mol Biol* **2008**, 68, 31-41.
272. Wang, L.-Y.; Li, D.; Deng, Y.-S.; Lv, W.; Meng, Q.-W., Antisense-mediated depletion of tomato GDP-L-galactose phosphorylase increases susceptibility to chilling stress. *J Plant Physiol* **2013**, 170, (3), 303-314.
273. Yang, D.-Y.; Ma, N.-N.; Zhuang, K.-Y.; Zhu, S.-B.; Liu, Z.-M.; Yang, X.-H., Overexpression of tomato *SIGGP-LIKE* gene improves tobacco tolerance to methyl viologen-mediated oxidative stress. *J Plant Physiol* **2017**, 209, 31-41.
274. Wicker, T.; Wing, R. A.; Schubert, I., Recurrent sequence exchange between homeologous grass chromosomes. *Plant J* **2015**, 84, (4), 747-759.
275. Cerbin, S.; Jiang, N., Duplication of host genes by transposable elements. *Curr Opin Genetics Dev* **2018**, 49, 63-69.
276. Jiang, N.; Bao, Z.; Zhang, X.; Eddy, S. R.; Wessler, S. R., Pack-MULE transposable elements mediate gene evolution in plants. *Nature* **2004**, 431, (7008), 569.
277. Hegde, R. R., Differential distribution of ascorbic acid and RNA in the developing anthers of *Datura stramonium* L. *Bot Mag Tokyo* **1985**, 98, (3), 219-223.
278. Kumar, R. R.; Goswami, S.; Gadpayle, K. A.; Singh, K.; Sharma, S. K.; Singh, G.; Pathak, H.; Rai, R. D., Ascorbic acid at pre-anthesis modulate the thermotolerance level of wheat (*Triticum aestivum*) pollen under heat stress. *J Plant Biochem Biot* **2014**, 23, (3), 293-306.
279. Zhao, D.; Hamilton, J. P.; Vaillancourt, B.; Zhang, W.; Eizenga, G. C.; Cui, Y.; Jiang, J.; Buell, C. R.; Jiang, N., The unique epigenetic features of Pack-MULEs and their impact on chromosomal base composition and expression spectrum. *Nucleic Acids Res* **2018**, 46, (5), 2380-2397.
280. Asseng, S.; Ewert, F.; Martre, P.; Rötter, R. P.; Lobell, D.; Cammarano, D.; Kimball, B.; Ottman, M.; Wall, G.; White, J. W., Rising temperatures reduce global wheat production. *Nat Clim Change* **2015**, 5, 143-147.

281. Afzal, I.; Basra, S. M.; Farooq, M.; Nawaz, A., Alleviation of salinity stress in spring wheat by hormonal priming with ABA, salicylic acid and ascorbic acid. *Int J Agric Biol* **2006**, 8, (1), 23-28.
282. Al-Hakimi, A.; Hamada, A., Counteraction of salinity stress on wheat plants by grain soaking in ascorbic acid, thiamin or sodium salicylate. *Biol Plantarum* **2001**, 44, (2), 253-261.
283. Athar, H.-u.-R.; Khan, A.; Ashraf, M., Inducing salt tolerance in wheat by exogenously applied ascorbic acid through different modes. *J Plant Nutr* **2009**, 32, (11), 1799-1817.
284. Hafez, E.; Gharib, H., Effect of exogenous application of ascorbic acid on physiological and biochemical characteristics of wheat under water stress. *Int J Plant Prod* **2016**, 10, (4), 579-596.
285. Khan, A.; Ashraf, M., Exogenously applied ascorbic acid alleviates salt-induced oxidative stress in wheat. *Environ Exp Bot* **2008**, 63, (1-3), 224-231.
286. Malik, S.; Ashraf, M., Exogenous application of ascorbic acid stimulates growth and photosynthesis of wheat (*Triticum aestivum* L.) under drought. *Soil & Environ* **2012**, 31, (1).
287. De Gara, L.; De Pinto, M.; Arrigoni, O., Ascorbate synthesis and ascorbate peroxidase activity during the early stage of wheat germination. *Physiol Plant* **1997**, 100, (4), 894-900.
288. Yang, T. B., B. Ooraikul, F, Studies on germination conditions and antioxidant contents of wheat grain. *Int J Food Sci Nutr* **2001**, 52, (4), 319-330.
289. Kapustin, Y.; Souvorov, A.; Tatusova, T.; Lipman, D., Splign: algorithms for computing spliced alignments with identification of paralogs. *Biol Direct* **2008**, 3, (1), 20.
290. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O., New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **2010**, 59, (3), 307-321.
291. Schreiber, A. W.; Sutton, T.; Caldo, R. A.; Kalashyan, E.; Lovell, B.; Mayo, G.; Muehlbauer, G. J.; Druka, A.; Waugh, R.; Wise, R. P., Comparative transcriptomics in the Triticeae. *BMC Genomics* **2009**, 10, (1), 285.

292. Beasley, J. T.; Bonneau, J. P.; Johnson, A. A., Characterisation of the nicotianamine aminotransferase and deoxymugineic acid synthase genes essential to Strategy II iron uptake in bread wheat (*Triticum aestivum* L.). *PLoS One* **2017**, 12, (5), e0177061.
293. Sears, E. R., The aneuploids of common wheat. *Mo Agr Exp Sta Res Bull* **1954**, 572, 1-58.
294. Vandesompele, J.; De Preter, K.; Pattyn, F.; Poppe, B.; Van Roy, N.; De Paepe, A.; Speleman, F., Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **2002**, 3, (7), research0034. 1.
295. Ramírez-González, R.; Borrill, P.; Lang, D.; Harrington, S.; Brinton, J.; Venturini, L.; Davey, M.; Jacobs, J.; Van Ex, F.; Pasha, A., The transcriptional landscape of polyploid wheat. *Science* **2018**, 361, (6403), eaar6089.
296. Druka, A.; Muehlbauer, G.; Druka, I.; Caldo, R.; Baumann, U.; Rostoks, N.; Schreiber, A.; Wise, R.; Close, T.; Kleinhofs, A., An atlas of gene expression from seed to seed through barley development. *Funct Integr Genomics* **2006**, 6, (3), 202-211.
297. Sekhon, R. S.; Lin, H.; Childs, K. L.; Hansey, C. N.; Buell, C. R.; de Leon, N.; Kaeppler, S. M., Genome-wide atlas of transcription during maize development. *Plant J* **2011**, 66, (4), 553-563.
298. Davidson, R. M.; Gowda, M.; Moghe, G.; Lin, H.; Vaillancourt, B.; Shiu, S. H.; Jiang, N.; Robin Buell, C., Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. *Plant J* **2012**, 71, (3), 492-502.
299. Katoh, K.; Standley, D. M., MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **2013**, 30, (4), 772-780.
300. Lindbo, J. A., High-efficiency protein expression in plants from agroinfection-compatible Tobacco mosaic virus expression vectors. *BMC Biotechnol* **2007**, 7, (1), 52.
301. Bally, J.; Jung, H.; Mortimer, C.; Naim, F.; Philips, J. G.; Hellens, R.; Bombarely, A.; Goodin, M. M.; Waterhouse, P. M., The rise and rise of *Nicotiana benthamiana*: a plant for all reasons. *Annu Rev Phytopathol* **2018**, 56, 405-426.
302. Philips, J. G.; Dudley, K. J.; Waterhouse, P. M.; Hellens, R. P., The rapid methylation of T-DNAs upon *Agrobacterium* inoculation in plant leaves. *Front Plant Sci* **2019**, 10, 312.

303. Hellens, R. P.; Edwards, E. A.; Leyland, N. R.; Bean, S.; Mullineaux, P. M., pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol Biol* **2000**, 42, (6), 819-832.
304. Amaya, I.; Osorio, S.; Martinez-Ferri, E.; Lima-Silva, V.; Doblas, V. G.; Fernández-Muñoz, R.; Fernie, A. R.; Botella, M. A.; Valpuesta, V., Increased antioxidant capacity in tomato by ectopic expression of the strawberry *D-galacturonate reductase* gene. *Plant Biotechnol* **2015**, 10, (3), 490-500.
305. Nunes-Nesi, A.; Carrari, F.; Lytovchenko, A.; Smith, A. M.; Loureiro, M. E.; Ratcliffe, R. G.; Sweetlove, L. J.; Fernie, A. R., Enhanced photosynthetic performance and growth as a consequence of decreasing mitochondrial malate dehydrogenase activity in transgenic tomato plants. *Plant Physiol* **2005**, 137, (2), 611-622.
306. Zhang, S.; Zhang, R.; Gao, J.; Gu, T.; Song, G.; Li, W.; Li, D.; Li, Y.; Li, G., Highly efficient and heritable targeted mutagenesis in wheat via the *Agrobacterium tumefaciens*-mediated CRISPR/Cas9 system. *International journal of molecular sciences* **2019**, 20, (17), 4257.
307. Watson, A.; Ghosh, S.; Williams, M. J.; Cuddy, W. S.; Simmonds, J.; Rey, M.-D.; Hatta, M. A. M.; Hinchliffe, A.; Steed, A.; Reynolds, D., Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat Plants* **2018**, 4, (1), 23-29.
308. Gallie, D. R., Increasing vitamin C content in plant foods to improve their nutritional value—successes and challenges. *Nutrients* **2013**, 5, (9), 3424-3446.
309. Chen, Z.; Gallie, D. R., The ascorbic acid redox state controls guard cell signaling and stomatal movement. *Plant Cell* **2004**, 16, (5), 1143-1162.
310. Chen, Z.; Gallie, D. R., Induction of monozygotic twinning by ascorbic acid in tobacco. *PLoS One* **2012**, 7, (6), e39147.

Appendices

Appendix 4. Eprint of Chapter 4 ‘Genome-Wide Identification and Characterization of the *GDP-L-Galactose Phosphorylase* Gene Family in Bread Wheat’ published by *BMC Plant Biology* on 26.11.19

RESEARCH ARTICLE

Open Access



Genome-wide identification and characterization of the *GDP-L-galactose phosphorylase* gene family in bread wheat

Ronan C. Broad^{1†}, Julien P. Bonneau^{1†}, Jesse T. Beasley¹, Sally Roden², Joshua G. Philips², Ute Baumann³, Roger P. Hellens² and Alexander A. T. Johnson^{1*} 

Abstract

Background: Ascorbate is a powerful antioxidant in plants and an essential micronutrient for humans. The *GDP-L-galactose phosphorylase* (*GGP*) gene encodes the rate-limiting enzyme of the L-galactose pathway—the dominant ascorbate biosynthetic pathway in plants—and is a promising gene candidate for increasing ascorbate in crops. In addition to transcriptional regulation, *GGP* production is regulated at the translational level through an upstream open reading frame (uORF) in the long 5′-untranslated region (5′UTR). The *GGP* genes have yet to be identified in bread wheat (*Triticum aestivum* L.), one of the most important food grain sources for humans.

Results: Bread wheat chromosomal groups 4 and 5 were found to each contain three homoeologous *TaGGP* genes on the A, B, and D subgenomes (*TaGGP2-A/B/D* and *TaGGP1-A/B/D*, respectively) and a highly conserved uORF was present in the long 5′UTR of all six genes. Phylogenetic analyses demonstrated that the *TaGGP* genes separate into two distinct groups and identified a duplication event of the *GGP* gene in the ancestor of the *Brachypodium*/*Triticeae* lineage. A microsynteny analysis revealed that the *TaGGP1* and *TaGGP2* subchromosomal regions have no shared synteny suggesting that *TaGGP2* may have been duplicated via a transposable element. The two groups of *TaGGP* genes have distinct expression patterns with the *TaGGP1* homoeologs broadly expressed across different tissues and developmental stages and the *TaGGP2* homoeologs highly expressed in anthers. Transient transformation of the *TaGGP* coding sequences in *Nicotiana benthamiana* leaf tissue increased ascorbate concentrations more than five-fold, confirming their functional role in ascorbate biosynthesis *in planta*.

Conclusions: We have identified six *TaGGP* genes in the bread wheat genome, each with a highly conserved uORF. Phylogenetic and microsynteny analyses highlight that a transposable element may have been responsible for the duplication and specialized expression of *GGP2* in anthers in the *Brachypodium*/*Triticeae* lineage. Transient transformation of the *TaGGP* coding sequences in *N. benthamiana* demonstrated their activity *in planta*. The six *TaGGP* genes and uORFs identified in this study provide a valuable genetic resource for increasing ascorbate concentrations in bread wheat.

Keywords: Ascorbic acid, Vitamin C, Upstream open reading frame, Phylogeny, Synteny, Gene expression, Transient expression

* Correspondence: johnsa@unimelb.edu.au

[†]Ronan C. Broad and Julien P. Bonneau should be considered joint first author

¹School of BioSciences, The University of Melbourne, Melbourne, Victoria 3010, Australia

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Background

Ascorbate, also known as vitamin C, is the most abundant water-soluble antioxidant in plant cells and plays important roles in photosynthetic function and stress tolerance [1, 2]. Ascorbate can mitigate the harmful effects of reactive oxygen species (ROS) produced by normal or stressed cellular metabolism, either directly as an ROS scavenger or indirectly as a substrate for the enzyme ascorbate peroxidase [3]. Increased biosynthesis of ascorbate in model and crop species has demonstrated that even minor increases in ascorbate can confer enhanced tolerance to a broad range of abiotic stresses [4, 5]. Additional roles for ascorbate within plant cells include: serving as an enzymatic cofactor [6] and precursor for organic acids [7], as well as influencing the cell cycle [8], cell wall expansion [9], and flowering time and senescence [10].

In humans, ascorbate is an essential micronutrient due to loss-of-function mutations within the *L-gulonono-1,4-lactone oxidase* gene which encodes the enzyme responsible for catalysing the last step in vertebrate ascorbate biosynthesis [11]. Beyond its well-established role in preventing the disease scurvy, ascorbate plays key roles in many physiological processes important for human health, including promoting and regulating the uptake of dietary iron in the human digestive process [12] and serving as a cofactor for enzymes involved in epigenetic programming [5, 13]. Although ascorbate deficiency is considered a historical disease, suboptimal intake of ascorbate is currently present in both developing and developed countries [14], and the re-emergence of scurvy has been documented in Australia due to poor dietary habits [15].

Four pathways for ascorbate biosynthesis have been proposed in plants: the L-galactose, L-gulose, *myo*-inositol, and D-galacturonate pathways. Despite evidence for each of these pathways in plants, the L-galactose pathway, which converts D-fructose-6-P to ascorbate via eight enzymatic steps, is the dominant pathway leading to ascorbate biosynthesis in *Arabidopsis thaliana* [16–18], tomato (*Solanum lycopersicum* L.) [19], rice (*Oryza sativa* L.) [20], and green algae (*Chlamydomonas reinhardtii*) [21]. The final gene to be identified in the L-galactose pathway, *GDP-L-galactose phosphorylase* (*GGP*, also known as *VTC2*), was discovered and functionally characterized in *Arabidopsis* in 2007 [16, 22, 23]. The *GGP* enzyme catalyzes the conversion of GDP-L-galactose to L-galactose-1-P and represents the first committed step toward ascorbate biosynthesis. Several studies have now identified the *GGP* enzyme as the rate-limiting step in the L-galactose pathway [22, 24–26] and as the key regulatory enzyme in ascorbate biosynthesis [27]. Overexpression of *GGP* genes has increased ascorbate concentrations by up to 4.2-fold in *Arabidopsis* [24, 26], 3.1-fold in potato (*Solanum tuberosum* L.) [28], 2.1-fold in strawberry (*Fragaria x ananassa*) [28], 6.2-fold in tomato

[28, 29], 1.4-fold in tobacco (*Nicotiana tabacum* L.) [30], and 2.6-fold in rice [25, 31].

A highly conserved upstream open reading frame (uORF)—a class of small ORFs located upstream of protein-coding major ORFs (mORFs) in the 5′-untranslated region (5′UTR) of mRNAs—of the *GGP* gene has been shown to regulate *GGP* translation and ascorbate concentrations [32–35]. Although uORFs are abundant in plant genomes with up to 60% of genes estimated to contain an uORF, less than 1% are conserved between plant species [36, 37]. The translational regulation of mORFs by highly conserved uORFs in response to cellular metabolite levels has been documented in several plant studies. For example, uORFs have been shown to regulate translation of the *Arabidopsis S-adenosylmethionine decarboxylase* and *polyamine oxidase 2* genes in response to polyamines [38–41], the *Arabidopsis* S1 class *basic-leucine zipper* and tomato *ornithine decarboxylase* genes in response to sucrose [42–44], the *Arabidopsis phosphoethanolamine N-methyltransferase* gene in response to phosphocholine [45], and the *Arabidopsis heat shock transcription factor B1* gene in response to galactinol [46]. In the case of the *GGP* uORF, the encoded peptide which is predicted to initiate from a non-canonical AUC or ACG start-codon is proposed to cause ribosomal stalling, thereby preventing translation of the downstream *GGP* mORF under high ascorbate concentrations, whereas under low ascorbate concentrations the uORF is skipped and the *GGP* mORF translated [33]. The precise mechanism of how ascorbate may influence ribosomal initiation at the uORF or mORF has yet to be elucidated. Disruption of the *GGP* uORF through mutation of key residues was shown to interfere with ascorbate feedback regulation of translation and increase ascorbate concentration when transiently transformed in *Nicotiana benthamiana* [33]. Recently, CRISPR/Cas9 genome editing has been used to disrupt the *GGP* uORF in *Arabidopsis*, lettuce (*Lactuca sativa*), and tomato to increase ascorbate concentrations and enhance oxidative stress tolerance [34, 35].

Despite the importance of the *GGP* gene in ascorbate biosynthesis and regulation, characterisation of *GGP* genes within graminaceous species has been limited to rice and maize (*Zea mays* L.). Both rice and maize have one copy of the *GGP* gene present in their genomes, located on chromosomes Os11 and Zm6, respectively. Complete knock-out of the *OsGGP* gene in rice severely reduced foliar ascorbate concentrations by 80% and was correlated with reduced photosynthetic efficiency, lower biomass production, and reduced tolerance to ozone stress and zinc deficiency [20]. In maize, the *ZmGGP* gene is expressed during endosperm development and expression levels have been shown to be genotype-dependent [47].

Wheat is the world's most cultivated crop in terms of land area and accounts for one fifth of the calories

consumed by humans [48]. Abiotic stresses such as drought, salinity, and high temperature represent the major limiting factors in global wheat productivity [49]. Increased ascorbate concentrations in wheat could help to mitigate yield losses associated with these abiotic stresses. Like many other cereals, wheat contains negligible levels of ascorbate in the mature grain [50], therefore, ascorbate enrichment of wheat grain could also provide a novel means of improving human dietary intakes of this essential vitamin.

Here we expand upon our knowledge of *GGP* genes in graminaceous species by describing the *TaGGP* gene family in bread wheat. We describe the chromosomal location, confirm the presence of the highly conserved uORFs, present phylogenetic relationship, gene synteny, and tissue and developmental expression patterns, and demonstrate activity of the *TaGGP* coding sequences by transient transformation in *N. benthamiana*.

Results

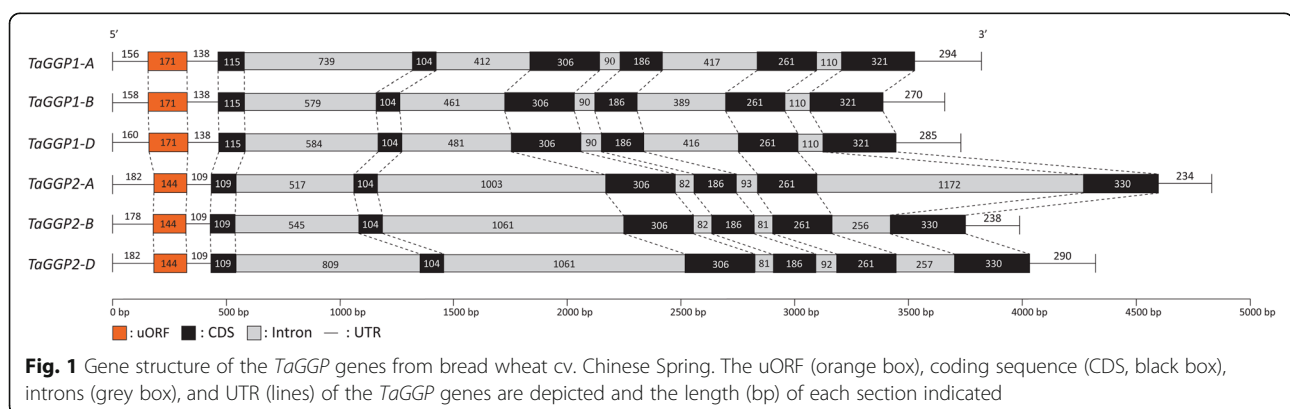
Six *TaGGP* genes were identified on bread wheat chromosomal groups 4 and 5

A total of six *TaGGP* genes were identified within the bread wheat genome and their genomic locations were determined. Chromosomal groups 4 and 5 each contain three homoeologous *TaGGP* genes located on the A, B, and D subgenomes (*TaGGP2-A*/*TaGGP2-B*/*TaGGP2-D* and *TaGGP1-A*/*TaGGP1-B*/*TaGGP1-D*, respectively). All six genes are located on the short arm of their respective chromosomal group except *TaGGP2-A* which is located on the long arm. In addition to the six *TaGGP* genes, we identified two *HvGGP* genes within the barley (*Hordeum vulgare*) genome on chromosomes 4 and 5 (*HvGGP2* and *HvGGP1*, respectively), two *BdGGP* genes within the *Brachypodium distachyon* (*Brachypodium* hereafter) genome on chromosome 4 (*BdGGP1* and *BdGGP2*), two *AetGGP* genes within the *Aegilops tauschii* genome on chromosome 4 and 5 (*AetGGP2* and *AetGGP1*, respectively), and one *SbGGP* gene within the *Sorghum bicolor* L. genome on chromosome 8.

The six *TaGGP* genomic sequences range in length from 2922 bp to 4163 bp mainly due to differences in intron sequence length (Fig. 1) and are similar in size and structure to other graminaceous *GGP* genes (Additional file 1: Figure S1). Genomic sequence comparisons of the six *TaGGP* genes revealed that they share between 36.0 to 94.9% identity. The *TaGGP1* and *TaGGP2* proteins are 430 and 431 amino acids in length, respectively, and sequence comparisons of the six *TaGGP* proteins revealed that they share between 78.1 to 99.3% identity. Alignment of the amino acid sequence of the wheat, barley, rice, maize, sorghum, *Brachypodium*, and *A. tauschii* *GGP* proteins identified highly conserved regions within the proteins, as well as the conserved histidine triad (HIT) motif of the HIT protein superfamily, the conserved KKRP nuclear localization signal (NLS), and 13 residues that distinguish *GGP2* proteins from *GGP1* proteins within these species (Fig. 2).

Each *TaGGP* gene contains a highly conserved uORF

For all six *TaGGP* genes, a highly conserved uORF was identified in the long 5'UTR of the mRNA initiating either from the non-canonical AUC (isoleucine) or ACG (threonine) start-codon. The uORF peptides of the *TaGGP1* homoeologs and *TaGGP2* homoeologs are 56 and 47 amino acids in length, respectively, and sequence comparisons of the six *TaGGP* uORF peptides revealed that they share between 67.9 to 100% identity. Alignment of the amino acid sequence of the *GGP* uORF peptides from a range of graminaceous and non-graminaceous species identified highly conserved regions within the peptides, as well as a lysine at the first residue in rice instead of a highly conserved isoleucine in all other species, a truncation of 11 residues in the *TaGGP2*, *HvGGP2*, *BdGGP2*, and *AetGGP2* uORF peptides relative to the *TaGGP1*, *HvGGP1*, *BdGGP1*, and *AetGGP1* uORF peptides, and a single amino acid change that distinguishes graminaceous species (alanine) from non-graminaceous species (glutamic acid) (Fig. 3).



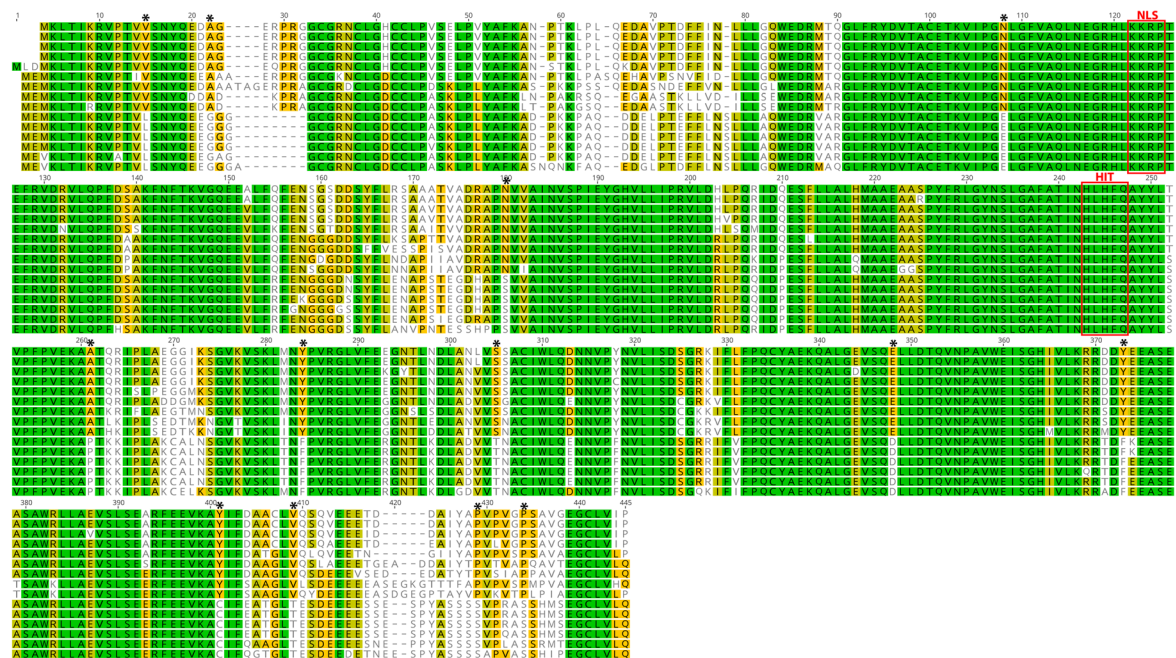


Fig. 2 Amino acid sequence alignment of GGP proteins from a range of graminaceous species. Green, olive green, yellow and white background colour represents 100%, 80 to 100%, 60 to 80%, and less than 60% conservation of amino acids between species, respectively. The HIT motif (HφHφQ, where φ is a hydrophobic amino acid) of the HIT protein superfamily and the KKR NLS are outlined in red. The 13 residues across the protein that distinguish GGP1 proteins from GGP2 proteins within the graminaceous species are indicated by an asterisk. The prefixes for the graminaceous species are as follows: Aet is *Aegilops tauschii*; Bd is *Brachypodium distachyon*; Hv is *Hordeum vulgare*; Os is *Oryza sativa*; Sb is *Sorghum bicolor*; Ta is *Triticum aestivum*; and Zm is *Zea mays*

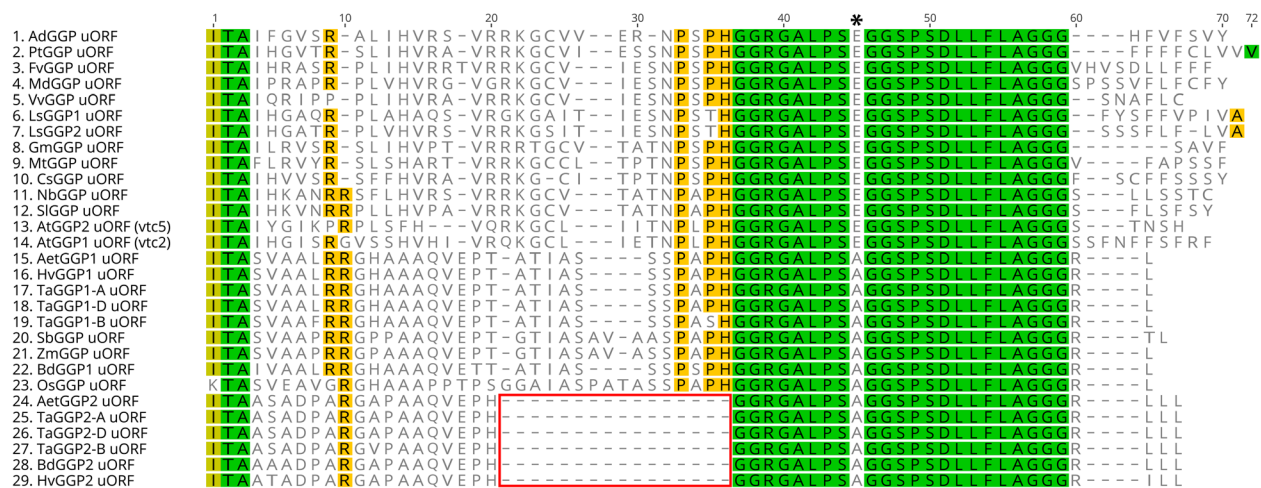


Fig. 3 Amino acid sequence alignment of GGP uORFs from a range of graminaceous and non-graminaceous species. Green, olive green, yellow, and white background colour represents 100%, 80 to 100%, 60 to 80%, and less than 60% conservation of amino acids between species, respectively. The 45th residue of the consensus sequence distinguishing the graminaceous (A; alanine) from the non-graminaceous (E; glutamic acid) species is indicated with an asterisk. The truncated 11 residues from the *TaGGP2*, *HvGGP2*, and *AetGGP2* uORF peptide sequences relative to the *TaGGP1*, *HvGGP1*, *BdGGP1*, and *AetGGP1* uORF peptide sequences are outlined in red. The prefixes for the graminaceous species are the same as those presented in Fig. 1. The prefixes for the non-graminaceous species are as follows: Ad is *Actinidia deliciosa*; At is *Arabidopsis thaliana*; Cs is *Cucumis sativus*; Fv is *Fragaria vesca*; Gm is *Glycine max*; Ls is *Lactuca sativa*; Md is *Malus x domestica*; Mt is *Medicago truncatula*; Nb is *Nicotiana benthamiana*; Pt is *Populus trichocarpa*; Sl is *Solanum lycopersicum*; and Vv is *Vitis vinifera*

Graminaceous GGP proteins separate into two groups

Phylogenetic analysis of the GGP protein sequences from a range of graminaceous and non-graminaceous plant species identified two clades of GGP proteins, belonging to either graminaceous or non-graminaceous species (Fig. 4). The graminaceous clade of GGP proteins further separates into two groups, with the *TaGGP1* proteins forming one group with *HvGGP1*, *BdGGP1*, *AetGGP1*, *OsGGP*, *ZmGGP*, and *SbGGP*, and the *TaGGP2* proteins forming a second group with *HvGGP2*, *BdGGP2*, and *AetGGP2*. This classification of group 1 and group 2 graminaceous GGP proteins is further supported by a phylogenetic analysis of the graminaceous *GGP* coding sequences which clearly demonstrates that the *TaGGP1* homoeologs, *HvGGP1*, *BdGGP1*, and *AetGGP1* are more closely related to the *OsGGP*, *ZmGGP*, and *SbGGP* in terms of evolutionary distance than the *TaGGP2* homoeologs, *HvGGP2*, *BdGGP2*, and *AetGGP2* (Additional file 1: Figure S2). Phylogenetic analysis of the *GGP* uORF peptide sequences revealed similar relationships to those observed for the GGP protein sequences (Additional file 1: Figure S3).

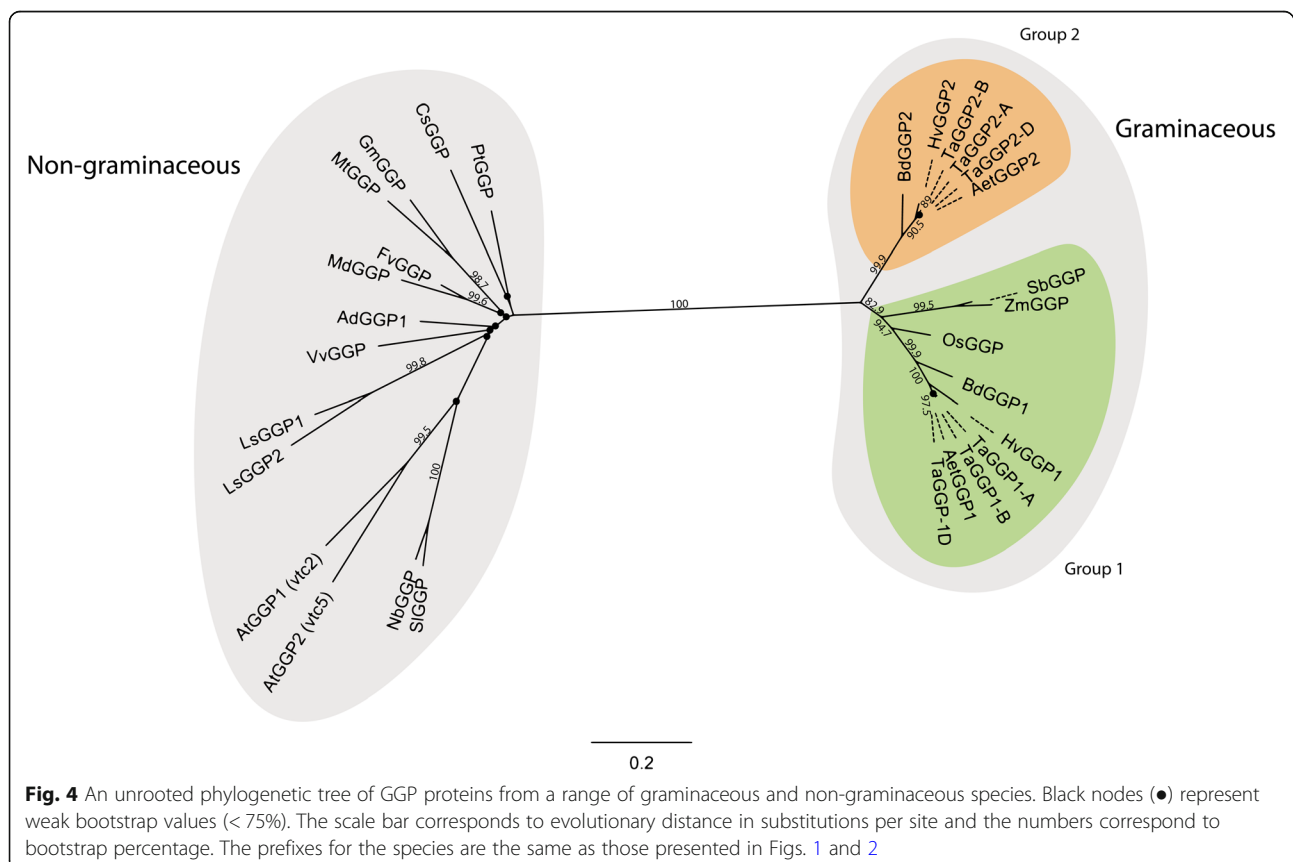
The *TaGGP1* and *TaGGP2* subchromosomal regions do not share synteny

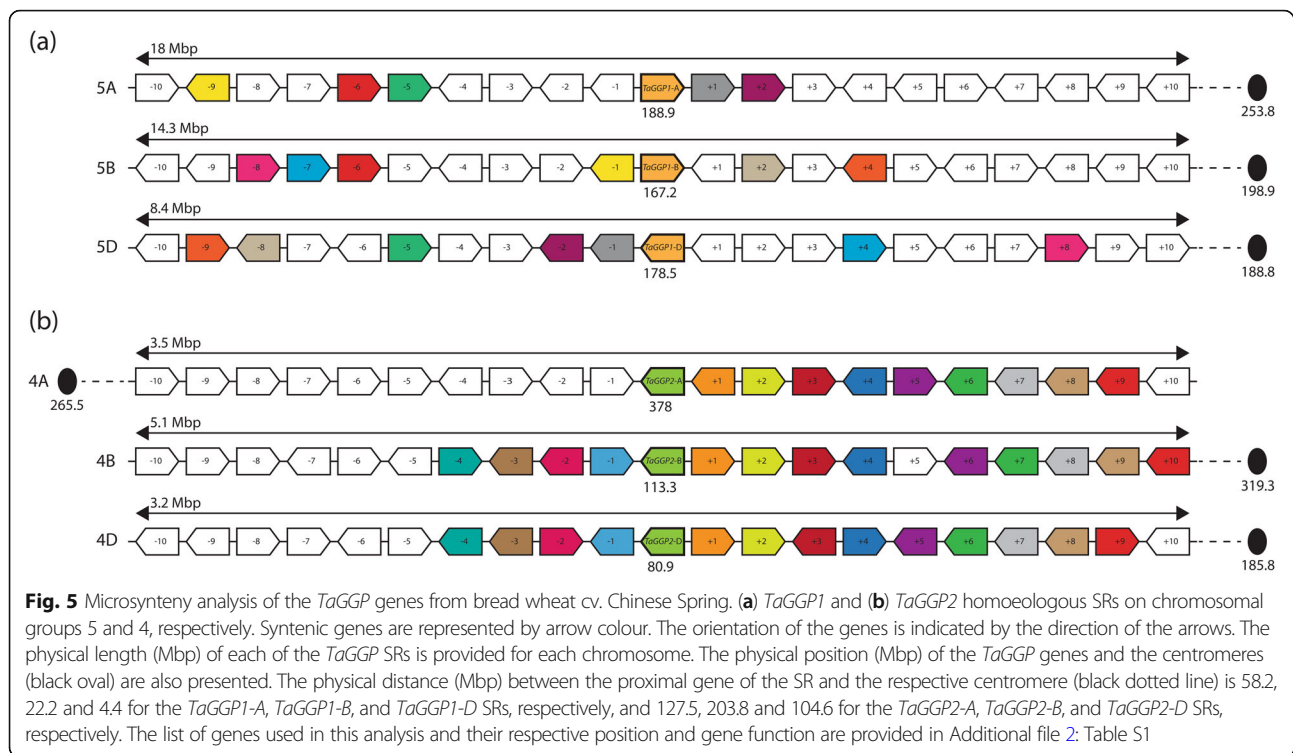
Selection of ten genes upstream and downstream of the respective *TaGGP* genes on each of the six chromosomes,

referred to as the *TaGGP* subchromosomal regions (SR), were analysed for gene density and synteny (Fig. 5 and Additional file 2: Table S1). The physical length of the *TaGGP1-A*, *TaGGP1-B*, and *TaGGP1-D* SRs were 18, 14.8, and 8.7 Mbp, respectively, and the *TaGGP2-A*, *TaGGP2-B*, and *TaGGP2-D* SRs were 3.4, 4.8, and 3 Mbp, respectively. No shared synteny was found between the *TaGGP1* and *TaGGP2* SRs on chromosomal groups 4 and 5 except for the *TaGGP* genes themselves. Synteny between the *TaGGP1* homoeologous SRs was much lower than that of the *TaGGP2* homoeologous SRs with only 19 genes having shared synteny—the *TaGGP1* homoeologs were the only genes to share synteny on all three homoeologous SRs—whereas the *TaGGP2* homoeologous SRs have 37 genes with shared synteny. The *TaGGP1* homoeologs are found in the centromeric region of their respective chromosomes, whereas the *TaGGP2* homoeologs are found in the interstitial region of their respective chromosomes between the pericentromeric and telomeric regions.

The two groups of *TaGGP* genes have distinct expression patterns in bread wheat tissues

Quantitative reverse transcription-PCR analysis of a range of bread wheat tissues and developmental stages revealed distinct patterns of expression for the *TaGGP1* and *TaGGP2* genes (Fig. 6). The *TaGGP1* homoeologs were





broadly expressed across different tissues and developmental stages, with the highest levels of expression detected in the seedling leaf, bracts, and anthers (Fig. 6a). The *TaGGP1* homoeologs displayed a trend of nonbalanced expression with the *TaGGP1-D* homoeolog having lower expression levels across all tissues relative to *TaGGP1-A* and *TaGGP1-B*. In contrast to the *TaGGP1* homoeologs, the *TaGGP2* homoeologs were highly expressed in the anthers, with approximately 50- to 580-fold higher expression in anthers relative to other tissues (Fig. 6b). The *TaGGP2* homoeologs also displayed a trend of nonbalanced expression with the *TaGGP2-B* homoeolog having lower expression levels across all tissues relative to *TaGGP2-A* and *TaGGP2-D*. Taken together, the *TaGGP* genes were most highly expressed in anthers, seedling leaf, and bracts and most lowly expressed in the crown, seedling root, and caryopses 3–5 DAP.

We searched for anther/pollen specific cis-regulatory elements in the 1-kb promoter region of the *TaGGP* genes to identify whether the high expression of the *TaGGP2* homoeologs in anthers could be explained due to differences in these regulatory elements (Additional file 1: Figure S4). However, the high level of expression of the *TaGGP2* homoeologs in anthers does not appear to correlate with the density of anther/pollen cis-regulatory elements surrounding the transcription start site (TSS), with a higher density of anther/pollen cis-regulatory elements instead being present in regions surrounding the TSS of the *TaGGP1* homoeologs (Additional file 1: Figure S4).

Transient transformation of *TaGGP* genes in *N. benthamiana* leaf increases ascorbate concentration

Coding sequences of the *TaGGP1-A*, *TaGGP1-B*, *TaGGP1-D*, *TaGGP2-A*, and *TaGGP2-B* genes increased ascorbate concentrations in transiently transformed *N. benthamiana* leaf relative to the control (Fig. 7). Transformation with *TaGGP1-D* produced the greatest increase in ascorbate concentration (5.3-fold) relative to the control. Similar fold increases in ascorbate concentrations were also observed with transformation of *TaGGP1-B* (4.1-fold), *TaGGP2-A* (4.4-fold), and *TaGGP2-B* (4.5-fold). Transformation with *TaGGP1-A* produced a 2.4-fold increase relative to the control, which was the lowest increase in ascorbate concentration observed. *Agrobacterium tumefaciens* containing the *TaGGP2-D* construct were unable to be recovered. The higher fold increase in ascorbate observed for some of the *TaGGP* genes relative to the positive control *AcGGP* gene from kiwifruit may be explained by the strength of the promoters used to drive transgene expression: the dual 35S promoter used to drive the *TaGGP* genes is a stronger promoter than the single 35S promoter used to drive *AcGGP*.

Discussion

The *TaGGP* genes are highly conserved, functional, and have an uORF amenable to genome editing

We identified six *TaGGP* genes in the bread wheat genome, comprising of three *TaGGP1* homoeologs on chromosomal group 5 and three *TaGGP2* homoeologs on chromosomal

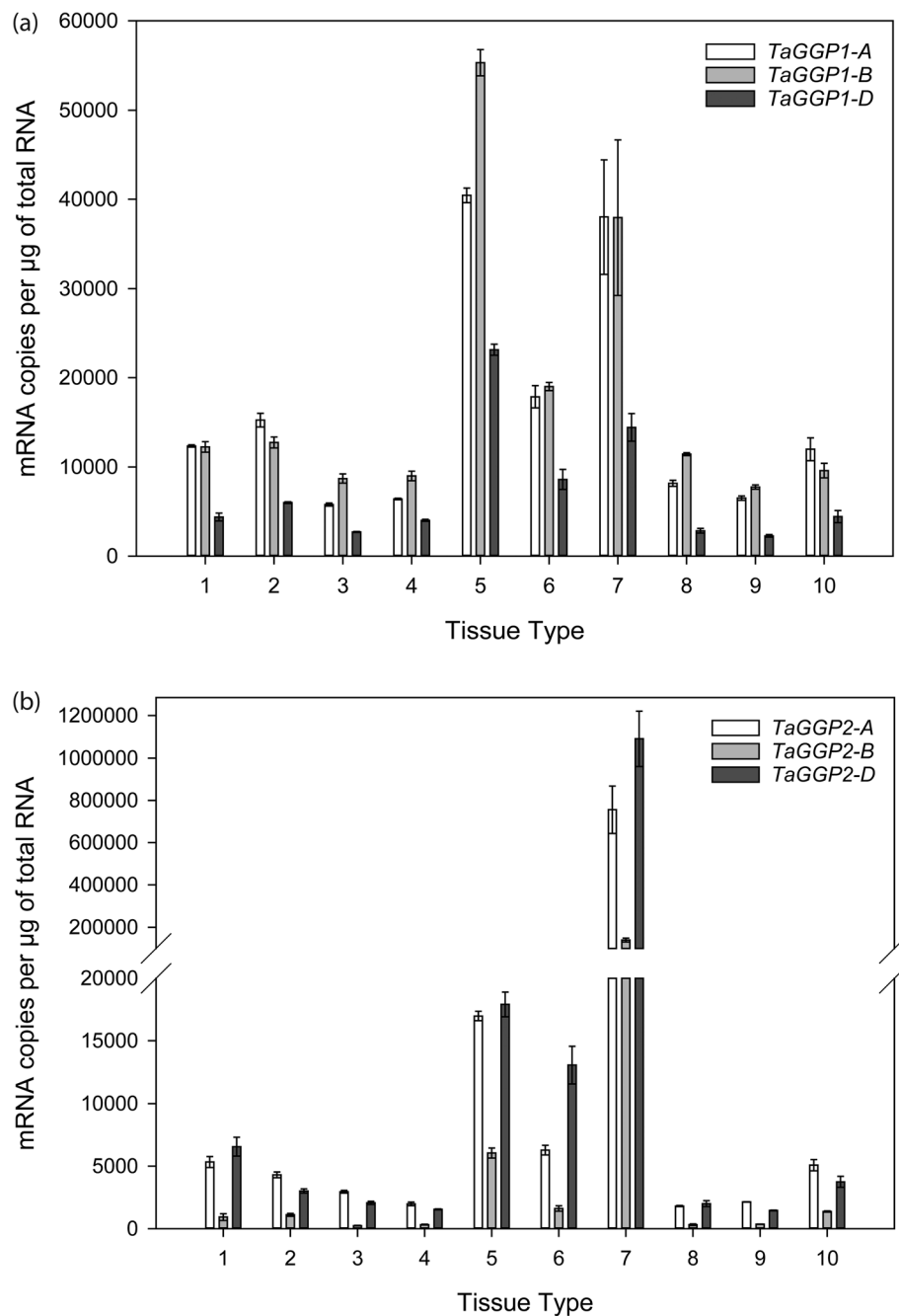


Fig. 6 Quantitative reverse transcription-PCR analysis of the *TaGGP* genes from bread wheat cv. Chinese Spring. Relative expression of the (a) *TaGGP1* and (b) *TaGGP2* homoeologs is provided in: (1) embryonic root; (2) mesocotyl; (3) seedling root; (4) crown; (5) seedling leaf; (6) bracts; (7) anthers; (8) pistil; (9) caryopsis 3–5 DAP; and (10) embryo 22 DAP. The geometric mean of *TaActin*, *TaCyclophilin*, and *TaELF* were used as normalisation factors. Error bars indicate SEM of three technical replicates derived from a bulk of three independent biological samples

group 4. Five of the genes are located on the short arm of their respective chromosomal group; the *TaGGP2-A* gene is located on the long arm of chromosome 4A due to a known pericentric inversion between the long and short chromosome arms [51]. We have also identified two copies of the *GGP* gene in the barley, *Brachypodium*, and *A.*

tauschii genomes, as well as one copy in the sorghum genome. To our knowledge, this is the first time that the *HvGGP1*, *HvGGP2*, *BdGGP1*, *BdGGP2*, *AetGGP1*, *AetGGP2*, and *SbGGP* gene have been reported on.

The *TaGGP* protein sequences are highly conserved, particularly between homoeologs, with up to 99.3% shared

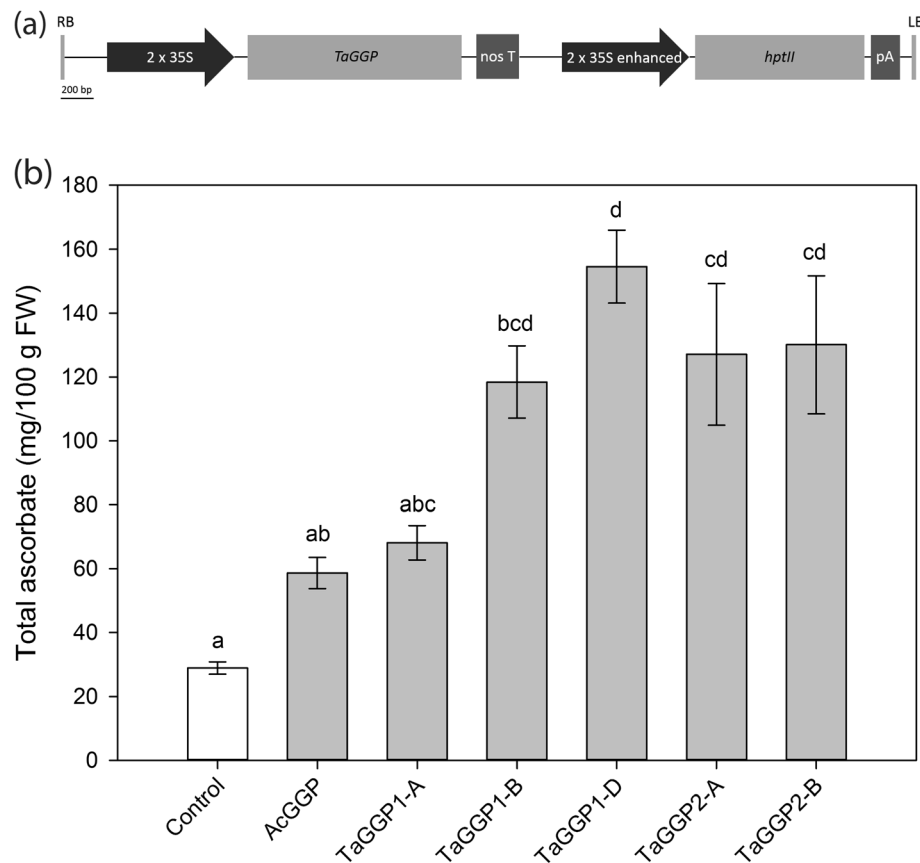


Fig. 7 Transient transformation of the *TaGGP* coding sequences in *N. benthamiana*. **a** Schematic representation of the T-DNAs used for constitutive overexpression of the *TaGGP* genes in *N. benthamiana*. RB, right border; 2 x 35S, dual CaMV 35S promoter; *TaGGP*, coding sequence of *TaGGP1-A* (1,293 bp), *TaGGP1-B* (1,293 bp), *TaGGP1-D* (1,293 bp), *TaGGP2-A* (1,296 bp), and *TaGGP2-B* (1,296 bp); nos T, nopaline synthase terminator; 2 x 35S enhanced, dual CaMV 35S promoter enhanced; *hptII*, hygromycin phosphotransferase II; pA, CaMV poly(A) signal; LB, left border. **b** Ascorbate concentrations of *N. benthamiana* leaves co-infiltrated with *A. tumefaciens* (GV3101 MP90) containing the constructs of interest and a P19 construct to prevent post-transcriptional gene silencing. The control was infiltrated with *A. tumefaciens* containing the P19 construct alone. The *AcGGP* gene from kiwifruit driven by a single 35S promoter in the pGreen vector system was used as a positive control. Bars represent mean $\pm$ SEM of three infiltrated young leaves. Means that do not share a letter are significantly different (one-way ANOVA followed by Tukey post-hoc test with 95% confidence level)

identity (Fig. 2). This high level of conservation is striking given the genetic redundancy of the hexaploid wheat genome, which generally displays accelerated rates of coding sequence evolution relative to diploid organisms such as *Brachypodium* and suggests that all six *TaGGP* genes play important functional roles during wheat development and as such may be resistant to genetic drift [52]. This hypothesis is consistent with our finding that all five *TaGGP* genes transiently transformed into *N. benthamiana* leaf increased leaf ascorbate concentrations relative to the control, confirming their functional role in ascorbate biosynthesis *in planta* (Fig. 7). We can also conclude from the transient transformation that the 13 residues that distinguish the graminaceous GGP2 proteins from GGP1 proteins do not appear to affect the function of the GGP2 proteins. The substantial increase in ascorbate concentrations in *N. benthamiana* leaf observed for the *TaGGP* genes is consistent with GGP as the rate-limiting step towards ascorbate

biosynthesis in plants [22, 24–26], and demonstrates the potential of utilising these genes, particularly *TaGGP1-D*, to increase ascorbate concentrations not only in wheat but other crops.

All six *TaGGP* proteins contained the highly conserved H ϕ H ϕ Q motif (where ϕ is a hydrophobic amino acid) that is characteristic of the galactose-1-phosphate uridylyltransferase branch of the HIT superfamily (Fig. 2.) [22, 23, 53]. They also all possess the highly conserved KKRP NLS, which may be responsible for localizing the *TaGGP* proteins to the nucleus, in addition to the cytoplasm where most of the ascorbate biosynthesis occurs (Fig. 2.) [54]. Fusion proteins of *AtGGP1*:YFP, *SIGGP1*:GFP, and *SIGGP2*:GFP have been found to localize to the cytoplasm and nucleus suggesting that GGP may be a dual-function protein serving both enzymatic and regulatory roles [54–56]. Determining the extent to which this NLS influences subcellular localization of

GGP and what role GGP may be playing in the nucleus warrants further investigation.

We also identified a highly conserved uORF in the long 5'UTR of the mRNA for each of the six *TaGGP* genes (Figs. 1 and 3). Notably, the uORF of the *TaGGP2* homoeologs had a truncation of 11 residues relative to the *TaGGP1* homoeologs and the same truncation of 11 residues was also observed for the *BdGGP2* and *HvGGP2* uORFs relative to the *BdGGP1* and *HvGGP2* uORFs, respectively, indicating that this truncation occurred prior to the evolutionary divergence of *Brachypodium* and the Triticeae tribe (Fig. 3). If these graminaceous *GGP2* uORFs are still functional this would suggest that these 11 residues are not critical to the functionality of the peptide. Disruption of the *GGP* uORF through genome editing has recently been shown as a viable strategy to increase ascorbate concentrations in a variety of plant species [34, 35]. For example, CRISPR/Cas9-targeted mutagenesis of the *AtGGP1* uORF in *Arabidopsis* and the *LsGGP1* and *LsGGP2* uORFs in lettuce to disrupt their function increased foliar ascorbate concentrations 1.7-, 1.4-, and 2.6-fold, respectively. Additionally, the increased ascorbate concentrations in lettuce conferred enhanced tolerance to methyl viologen-induced oxidative stress [35]. Similarly, CRISPR/Cas9-targeted mutagenesis of the *SlGGP2* uORF in tomato to disrupt its function increased foliar ascorbate concentrations approximately 1.4-fold [34]. Due to the plasticity of the hexaploid wheat genome, manipulating individual or multiple *TaGGP* uORFs with genome editing tools such as CRISPR/Cas9 represents a promising strategy to produce ascorbate-enriched wheat plants whilst also minimizing any pleiotropic effects on plant development that may arise due to disruption of the *GGP* uORF.

A *Mutator*-like transposable element may be responsible for the duplication and specialized gene expression of *GGP2* in the *Brachypodium*/Triticeae lineage

We have categorized the graminaceous *GGP* genes described in this manuscript as group 1 or group 2 based on phylogeny. The protein and coding sequences of the *TaGGP1* homoeologs, *AetGGP1*, *HvGGP1*, and *BdGGP1* are more closely related to *OsGGP*, *ZmGGP*, and *SbGGP* in terms of evolutionary distance than that of the *TaGGP2* homoeologs, *AetGGP2*, *HvGGP2*, and *BdGGP2* (Fig. 4). This grouping of graminaceous *GGP1* and *GGP2* genes is also consistent with the chromosomal location of the *GGP* genes and the known evolutionary history of grass genomes. For example, the wheat chromosomal group Ta5, barley chromosome Hv5, *Brachypodium* chromosome Bd4, rice chromosome Os12, and sorghum chromosome Sb8, which each harbour *GGP1* genes, are all descended from the ancestral grass chromosome A12 [57]. Whereas, the wheat chromosomal group Ta4, barley chromosome Hv4, and *Brachypodium* chromosome Bd4,

which each harbour *GGP2* genes, are all descended from the ancestral grass chromosome A11 (in *Brachypodium* A11 and A12 were fused to form the single chromosome Bd4) [57]. Given that rice, maize, and sorghum each have a single copy of *GGP*, while wheat, barley, *Brachypodium*, and *A. tauschii* each have two copies of *GGP*, it appears that *GGP* may have been duplicated from A12 to A11 in the ancestor of the *Brachypodium*/Triticeae lineage. This duplication event would have occurred after the evolutionary divergence of the *Brachypodium*/Triticeae lineage from rice approximately 33.5 million years ago (MYA) but prior to the evolutionary divergence of *Brachypodium* and the Triticeae tribe approximately 24.5 MYA [57].

Our microsynteny analysis of the *TaGGP* genes revealed that the *TaGGP1* and *TaGGP2* SRs have no shared synteny except for the *TaGGP* genes themselves (Fig. 5). We also did not find any shared synteny between the *BdGGP1* and *BdGGP2* SRs in *Brachypodium* (Additional file 2: Table S2) or between the *HvGGP1* and *HvGGP2* SRs in barley (Additional file 2: Table S3). This lack of synteny between the *GGP1* and *GGP2* SRs in wheat, *Brachypodium*, and barley suggests that *GGP2* may have been duplicated via transposition in the ancestor of the *Brachypodium*/Triticeae lineage since other methods of gene duplication preserve gene synteny [58]. Given that the introns and gene structure of the wheat, *Brachypodium*, and barley *GGP2* genes have been preserved, a Class II transposable element (DNA transposon) that transposes via a DNA intermediate may have been responsible for such a transposition. We propose that a *Mutator*-like transposable element (MULE)—a family of DNA transposons that are prevalent in plants and capable of capturing, mobilizing, and amplifying gene fragments—may have been responsible for the duplication of *GGP2* [59]. However, given the proposed evolutionary time since *GGP2* was duplicated (24.5 to 33.5 MYA) any features of the transposable element would have degenerated by now and no longer be detectable [58].

We identified distinct expression patterns for the *TaGGP1* and *TaGGP2* genes with the *TaGGP1* homoeologs broadly expressed across different tissues and developmental stages, whereas the *TaGGP2* homoeologs were highly specialized with approximately 50- to 580-fold higher expression in anthers relative to other tissues (Fig. 6). Our *TaGGP* gene expression results are supported by gene expression data extracted from <http://bar.utoronto.ca/> which also demonstrates that the *TaGGP1* homoeologs are broadly expressed across different tissues and developmental stages, and that the *TaGGP2* homoeologs are highly expressed in anthers and leaves relative to other tissues (Additional file 1: Figure S5). Similar *GGP* expression profiles were also observed for other graminaceous species in terms of higher expression in anthers and/or leaves relative to other tissues (Additional file 1: Figures S5 and S6). To

our knowledge this is the first time that specialized expression of a *GGP* gene in anthers has been documented, however, a previous study found high ascorbate concentrations correlated with high RNA concentrations in the developing anthers of jimsonweed (*Datura stramonium* L.) [60]. Together this expression data suggests that ascorbate may play an important role in anther development in wheat and other graminaceous species, for example protecting pollen from ROS induced by environmental stresses [61]. Determining whether the high expression of *GGP* in anthers is correlated with higher ascorbate concentrations relative to other tissues warrants further study.

In regards to the origin of the *cis*-regulatory elements responsible for the high expression of *GGP2* genes in the anthers relative to other tissues in wheat, *Brachypodium*, and barley, we propose two hypotheses: (i) following duplication, the regulatory region of the *GGP2* gene diverged due to a reduction in selection pressure, which subsequently generated de novo *cis*-regulatory elements, or (ii) *cis*-regulatory elements in the regulatory region were inherited from a transposable element that may have been responsible for the duplication of *GGP2*. In support of the first hypothesis, sequence alignments of the 1-kb promoter region of the wheat, *Brachypodium*, and barley *GGP* genes demonstrate that the promoter region of the *GGP2* genes have commonalities with the promoter region of the *GGP1* genes but have otherwise diverged significantly (Additional file 1: Figure S7). It is possible that sequence changes to the promoter region of *GGP2* genes have generated de novo *cis*-regulatory elements responsible for the high expression of *GGP2* genes in anthers. On the other hand, transposable elements are known to influence the expression of transposed genes [58]. For example, MULEs—including any sequences that have been captured, mobilized, and amplified by a MULE—have been shown to be preferentially expressed in reproductive tissues [58, 62]. The terminal inverted repeat (TIR) of MULEs are proposed to serve as the regulatory region for the expression of MULEs, including any transposed gene fragments [58, 62]. If *GGP2* was transposed by a MULE in the ancestor of the *Brachypodium*/Triticeae lineage, it is possible that *GGP2* inherited *cis*-regulatory elements responsible for the high expression of *GGP2* in anthers directly from the TIR of the MULE.

Applications towards plant breeding

Abiotic stresses, such as high temperature, are responsible for major losses in global wheat productivity and are predicted to intensify due to climate change [63]. Identifying genetic strategies to mitigate or offset these losses will be essential for maintaining global wheat production levels. Exogenous application of ascorbate to wheat plants has been shown to alleviate salt- and drought-induced oxidative stress in many studies [64–69]. The development of

high-ascorbate bread wheat cultivars through overexpression of the *TaGGP* genes or genome editing of one or more *TaGGP* uORFs to disrupt their function could help to mitigate wheat production losses caused by abiotic stress without the need for exogenous applications.

With respect to human nutrition, the development of bread wheat products with elevated levels of ascorbate may be limited due to the decline of ascorbate to negligible levels during grain maturation that may be difficult to overcome [70, 71]. However, wheat products made from germinated grain may provide a vehicle for ascorbate biofortification since germinated grain contains considerably higher levels of ascorbate relative to non-germinated grain [72, 73]. Germinated grain from bread wheat cultivars with an enhanced capacity to biosynthesize ascorbate could represent a novel way for improving human dietary intake of ascorbate.

Conclusions

We identified six *TaGGP* genes in the bread wheat genome, each with a highly conserved uORF. Phylogenetic and microsynteny analyses revealed that *GGP2* may have been duplicated via a transposable element in the ancestor of the *Brachypodium*/Triticeae lineage. The *TaGGP2* homoeologs had specialized expression in anthers which may have been inherited from the TIR of a MULE. Transient transformation of the *TaGGP* coding sequences in *N. benthamiana* leaf increased ascorbate concentrations up 5.3-fold, confirming their activity *in planta*. We propose that the six *TaGGP* genes and uORFs identified in this study will provide a valuable genetic resource for increasing ascorbate biosynthesis in bread wheat to improve wheat abiotic stress tolerance and human nutrition.

Methods

Identification and naming of the *GGP* genes and uORFs

The *OsGGP* coding sequence (LOC4351698) was used as a query in BLASTN searches against the International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v1.0 assembly of the wheat genome cv. Chinese Spring [48] to identify *TaGGP* coding sequences. Each *TaGGP* gene was assigned a unique name based on homoeologous grouping and their respective subgenomes and followed the recommended rules for gene symbolization in wheat (<http://wheat.pw.usda.gov/ggpages/wgc/98/Intro.htm>). Gene prediction models including predicted coding and proteins sequences were obtained from the IWGSC database. The *OsGGP* coding sequence was also used as a query in BLASTN searches against the genome sequence databases for barley (IBSC_v2), *Brachypodium* (*Brachypodium distachyon*_v3.0), *A. tauschii* (Aet_v4.0) from <https://plants.ensembl.org> and for sorghum (*Sorghum bicolor*_NCBIv3) from <https://www.ncbi.nlm.nih.gov>. The structure of the *GGP* genes was determined with Splign (<https://www.ncbi>.

nml.nih.gov/sutils/splign/splign.cgi) [74]. The *GGP* uORF coding and peptide sequences were manually identified in the 5'UTR of the *GGP* genes starting from the highly conserved AUC codon encoding isoleucine as previously described [33].

Cloning and gene sequencing

Full-length gene sequences of *TaGGP1-A* and *TaGGP2-A* could not be retrieved from the IWGSC RefSeq v1.0 assembly of the wheat genome cv. Chinese Spring [48] due to each gene having a gap in intron 2 where sequence was missing. PCR primers were designed to amplify gene fragments spanning the gap of the *TaGGP1-A* and *TaGGP2-A* genes from cv. Chinese Spring genomic DNA using the respective IWGSC gene sequences and Primer3 software [75, 76]. The *TaGGP1* gene fragment was PCR amplified with forward primer 5' - GGCAATCTGTTAGGCAAGCA - 3' and reverse primer 5' - TGTCAAAAACAGGTATCA GCAATTT - 3'. The *TaGGP2* gene fragment was PCR amplified with forward primer 5' - AGTTCTTCCTAAAT TCTCTCCTCCTT - 3' and reverse primer 5' - TTGAGG AATCACCTCACCTA - 3'. PCR amplification cycles consisted of 1 cycle = 1 m 95 °C; 35 cycles = 20 s 95 °C, 30 s 61 °C, 30 s 72 °C; 1 cycle = 5 m 72 °C. PCRs were performed in a final volume of 20 µL for MyTaq™ HS Red DNA Polymerase (Bioline, MA, USA) according to manufacturer's instructions. The PCR products were purified with DNA Clean & Concentrator-5 (Zymo Research, CA, USA) and cloned into the pGEM®-T Vector System (Promega, WI, USA) according to manufacturer's instructions. Restriction enzyme digestion of the cloned PCR products in pGEM®-T with PvuII (New England Biolabs, MA, USA) according to the manufacturer instructions was used to discriminate the *TaGGP1-A* and *TaGGP2-A* gene fragments from their respective B and D homoeologous gene fragments prior to Sanger sequencing at the Australian Genome Research Facility Ltd. (<http://www.agrf.org.au/>).

Sequence alignment and phylogenetic analysis of the *GGP* and uORF sequences

Sequence alignment and phylogenetic analysis of the *GGP* coding, protein, and uORF peptide sequences were performed in Geneious 10.2.2 using ClustalW alignment with default parameters and the PhyML Geneious plugin [77] with a K80 substitution model for nucleotide sequences or a LG substitution model for amino acid sequences and a bootstrap value of 1000. The *GGP* genes used in the sequence alignments and phylogenetic analyses are provided in Additional file 2: Table S4.

Microsynteny analysis of the *TaGGP*, *BdGGP*, and *HvGGP* genes

Ten high confidence genes upstream and downstream of the six *TaGGP*, two *BdGGP*, and two *HvGGP* genes

were retrieved from the IWGSC RefSeq v1.0 assembly of the wheat genome cv. Chinese Spring [48], the *Brachypodium distachyon* v3.0 assembly of the *Brachypodium* genome (<https://plants.ensembl.org>), and the IBSC_v2 assembly of the barley genome (<https://plants.ensembl.org>), respectively, excluding genes associated with retro-transposons and retrovirus related polypeptides, and used as query sequences in BLASTN searches against the respective databases with an E-value threshold of $1e^{-10}$. The genes were considered syntenic based on their genomic location and sequence identity ($\geq 90\%$). The list of the genes used in the microsynteny analysis of the six *TaGGP*, two *BdGGP*, and two *HvGGP* genes, and their respective position and gene function, are provided in Additional file 2: Tables S1, S2, and S3, respectively.

Quantitative reverse transcription-PCR analysis of the *TaGGP* genes

The cDNA library of 10 different tissues and development stages of bread wheat cv. Chinese Spring was previously generated as described [78]. The quantitative reverse transcription-PCR analysis of the six *TaGGP* genes was carried out using subgenome specific PCR primer pairs (Additional file 2: Table S5) for each *TaGGP* gene as previously described [79]. Subgenome specificity of primer pairs was verified using cv. Chinese Spring nulli-tetrasomic DNA [80]. Primer efficiency was $\geq 97\%$ for all primer pairs. A three gene normalisation factor of the most stable housekeeping genes *TaActin*, *TaCyclophilin*, and *TaELF* was used to normalise qRT-PCR expression data as previously described [78, 81]. Supplementary *GGP* gene expression data was extracted using the corresponding gene identifiers provided in Additional file 2: Table S4 from <http://bar.utoronto.ca/> for wheat [82], barley [83], and maize [84], and from <https://www.ebi.ac.uk> for *Brachypodium*, rice, and sorghum [85].

Analysis of *TaGGP* promoter regions

For each *TaGGP* gene, the first exon as well as the 1-kb upstream region, which was considered the promoter region, was extracted from the genome sequence databases for *T. aestivum* (IWGSC) from <https://plants.ensembl.org>. These sequences were used as a query in BLASTN searches against the NCBI expressed sequence tag (EST) database for *T. aestivum* (<https://blast.ncbi.nlm.nih.gov>) to identify *TaGGP* cap-trapped full-length cDNA sequences and the TTS. The identification of anther/pollen *cis*-acting elements was performed in Geneious 10.2.2 as previously described [79]. Sequence alignment of the 1-kb promoter region of the *GGP* genes was performed in Geneious 10.2.2 using the MAFFT Geneious plugin [86] with a L-INS-i algorithm.

Transient transformation of *N. benthamiana* leaves

Bread wheat cv. Gladius was sourced from the Australian Grains Genebank (<https://www.seedpartnership.org.au/associates/agg>). PCR primers were designed to amplify the coding sequence of the *TaGGP1* homoeologs and *TaGGP2* homoeologs from bread wheat cv. Gladius cDNA using Primer3 software [75, 76]. The *TaGGP1* homoeologs were PCR amplified with forward primer 5′ – ATGAAGCTGACGATTAAGAGGGTA – 3′ and reverse primer 5′ – TCACGGAATTACGAGGCAG – 3′. The *TaGGP2* homoeologs were PCR amplified with forward primer 5′ – ATGAGATGAAGCTGACGAT – 3′ and reverse primer 5′ – TCACTGAAGGACAAGGCAAC – 3′. PCR amplification cycles consisted of 1 cycle = 30 s 98 °C; 35 cycles = 10 s 98 °C, 15 s 61 °C, 45 s 72 °C; 1 cycle = 5 m 72 °C. PCRs were performed in a final volume of 50 µL for Phusion High-Fidelity DNA Polymerase (New England Biolabs) according to the manufacturer's instructions. The PCR products were purified with DNA Clean & Concentrator-5 (Zymo Research), A overhangs added with MyTaq™ HS DNA Polymerase (Bioline), ligated into the pCR8/GW/TOPO vector system using the pCR™8/GW/TOPO® TA Cloning Kit (Invitrogen, CA, USA), and recombined into the pMDC32 vector system [87] using Gateway™ LR Clonase™ II Enzyme mix (Invitrogen) according to the manufacturer's instructions. Cloning of the correct gene family member and gene orientation was confirmed by Sanger sequencing at the Australian Genome Research Facility Ltd. *A. tumefaciens* (GV3101 pMP90) containing the constructs of interest (except *TaGGP2-D*) were co-infiltrated with *A. tumefaciens* containing a P19 construct [88] to prevent post-transcriptional gene silencing into the younger leaves of 6-week-old *N. benthamiana* plants (LAB strain) grown in a plant growth room (16 h daylength, 24 °C) [89]. Infiltration conditions were conducted as previously described [90]. Infiltration with *A. tumefaciens* containing the P19 construct alone was used as a control. Infiltration with *A. tumefaciens* containing the kiwifruit *AcGGP* gene in the pGreen vector system [24, 91] was used as a positive control. Leaves were harvested for ascorbate measurements 7 d after infiltration and frozen on dry ice.

Ascorbate measurement

Total ascorbate was extracted and measured as previously described with modifications [92]. Briefly, total ascorbate was extracted from ground, lyophilized tissue homogenised in extraction fluid containing 8% metaphosphoric acid, 2 mM EDTA, and 2 mM TCEP (Sigma-Aldrich, MO, USA) at 40 °C for 2 h. The extract was centrifuged and 5 µl of supernatant was injected onto a C18 3 µm 33 × 7 mm Alltima Rocket column (Hichrom Limited, UK) with a flow rate of 1 mL/m. The concentration was determined by reverse phase HPLC with UV detection at 245 nm on a Shimadzu (Japan) 8050 instrument

(RT 1.43 m) using ascorbic acid as a standard (Sigma-Aldrich). The elution procedure was gradient with mobile phases A: MS grade water with 1% formic acid (Sigma-Aldrich) and B: MS grade acetonitrile with 1% formic acid (Sigma-Aldrich). The peak was validated as ascorbic acid by using L-ascorbic acid analytical standard (Sigma-Aldrich) with HPLC-MS Multi Reaction Monitoring (MRM) scan, qualifier transition m/z (+) = 177.05 > 95.00.

Statistical analysis

Statistically significant differences between means were detected using a one-way ANOVA followed by a Tukey post-hoc test with a 95% confidence level using Minitab® 17.1.0 (<http://www.minitab.com/en-us/>).

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12870-019-2123-1>.

Additional file 1: Figure S1. Gene structure of the *AetGGP*, *BdGGP*, *HvGGP*, *OsGGP*, *SbGGP*, and *ZmGGP* genes. **Figure S2.** An unrooted phylogenetic tree of *GGP* coding sequences from a range of graminaceous species. **Figure S3.** An unrooted phylogenetic tree of *GGP* uORF peptides from a range of graminaceous and non-graminaceous species. **Figure S4.** Annotation of anther/pollen *cis*-acting elements within the 1-kb promoter region of the *TaGGP* genes. **Figure S5.** *GGP* gene expression data extracted from <http://bar.utoronto.ca/> for wheat and barley and <https://www.ebi.ac.uk> for *Brachypodium*. **Figure S6.** *GGP* gene expression data extracted from <https://www.ebi.ac.uk> for rice and sorghum and <http://bar.utoronto.ca/> for maize. **Figure S7.** Nucleotide sequence alignment of the 1-kb promoter regions of wheat, *Brachypodium*, and barley *GGP* genes.

Additional file 2: Table S1. Ten high confidence genes upstream and downstream of the six *TaGGP* genes. **Table S2.** Ten high confidence genes upstream and downstream of the two *BdGGP* genes. **Table S3.** Ten high confidence genes upstream and downstream of the two *HvGGP* genes. **Table S4.** The *GGP* genes utilized in the manuscript for sequence alignments and phylogenetic analyses. **Table S5.** Primers used for the quantitative reverse transcription-PCR analysis of the six *TaGGP* and housekeeping genes.

Abbreviations

5'UTR: 5′ untranslated region; EST: Expressed sequence tag; GGP: GDP-L-galactose phosphorylase; HIT: Histidine triad; mORF: Major open reading frame; MULE: *Mutator*-like transposable element; NLS: Nuclear localization signal; ROS: Reactive oxygen species; SR: Subchromosomal regions; TIR: Terminal inverted repeat; TSS: Transcription start site; uORF: Upstream open reading frame

Acknowledgements

The authors would like to thank Margaret Pallotta for providing wheat cv. Chinese Spring nulli-tetrasomic DNA used in this study.

Author's contributions

JPB, RPH, AATJ conceived and designed the research. JPB identified the *GGP* genes and uORFs and RCB visualized their structures. RCB and JTB conducted the phylogenetic analyses. JPB and UB conducted the microsynteny analysis. JTB confirmed the specificity of the subgenome primers for the quantitative reverse transcription-PCR analysis and analysed the data. JPB analysed the supplementary gene expression data. JPB, RCB, and UB analysed the promoter regions. RCB, SR, and JGP conducted the transient transformation of *N. benthamiana* and quantified the ascorbate concentrations. JPB, RPH, and AATJ supervised the research. RCB drafted the manuscript. All authors read and approved the manuscript.

Funding

No funding was obtained for this study.

Availability of data and materials

The GGP and uORF sequences, as well as the supplementary gene expression data, analysed during this study are available from public databases as outlined in the material and methods and supplementary material. The GenBank accession numbers for the bread wheat cv. Gladius *TaGGP* coding sequences generated during the current study are: *TaGGP1-A*, MK514258; *TaGGP1-B*, MK514259; *TaGGP1-D*, MK514260; *TaGGP2-A*, MK514261; *TaGGP2-B*, MK514262; and *TaGGP2-D*, MK514263. The GenBank accession numbers will be searchable after Feb 13, 2020 or once the accession numbers appear in print. The remaining datasets generated during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of BioSciences, The University of Melbourne, Melbourne, Victoria 3010, Australia. ²Centre for Tropical Crops and Biocommodities, Institute for Future Environments, Queensland University of Technology, Brisbane, Queensland 4001, Australia. ³School of Agriculture, The University of Adelaide, Adelaide, South Australia 5064, Australia.

Received: 15 August 2019 Accepted: 7 November 2019

Published online: 26 November 2019

References

- Foyer CH, Noctor G. Managing the cellular redox hub in photosynthetic organisms. *Plant Cell Environ.* 2012;35(2):199–201.
- Foyer CH, Shigeoka S. Understanding oxidative stress and antioxidant functions to enhance photosynthesis. *Plant Physiol.* 2011;155(1):93–100.
- Foyer CH, Noctor G. Ascorbate and glutathione: the heart of the redox hub. *Plant Physiol.* 2011;155(1):2–18.
- Lisko KA, Aboobucker SI, Torres R, Lorence A. Engineering elevated vitamin C in plants to improve their nutritional content, growth, and tolerance to abiotic stress. In: Jetter R, editor. *Phytochemicals—Biosynthesis, Function and Application*. Switzerland: Springer; 2014. p. 109–28.
- Macknight RC, Laing WA, Bulley SM, Broad RC, Johnson AA, Hellens RP. Increasing ascorbate levels in crops to enhance human nutrition and plant abiotic stress tolerance. *Curr Opin Biotech.* 2017;44:153–60.
- De Tullio M, Arrigoni O. Hopes, disillusion and more hopes from vitamin C. *Cell Mol Life Sci.* 2004;61(2):209–19.
- Debolt S, Melino V, Ford CM. Ascorbate as a biosynthetic precursor in plants. *Ann Bot.* 2006;99(1):3–8.
- Potters G, De Gara L, Asard H, Horemans N. Ascorbate and glutathione: guardians of the cell cycle, partners in crime? *Plant Physiol Biochem.* 2002;40(6–8):537–48.
- Horemans N, Foyer CH, Potters G, Asard H. Ascorbate function and associated transport systems in plants. *Plant Physiol Biochem.* 2000;38(7–8):531–40.
- Barth C, De Tullio M, Conklin PL. The role of ascorbic acid in the control of flowering time and the onset of senescence. *J Exp Bot.* 2006;57(8):1657–65.
- Drouin G, Godin J-R, Pagé B. The genetics of vitamin C loss in vertebrates. *Curr Genomics.* 2011;12(5):371–8.
- Lane DJ, Richardson DR. The active role of vitamin C in mammalian iron metabolism: much more than just enhanced iron absorption! *Free Radic Biol Med.* 2014;75:69–83.
- Camarena V, Wang G. The epigenetic role of vitamin C in health and disease. *Cell Mol Life Sci.* 2016;73(8):1645–58.
- Troesch B, Hoefl B, McBurney M, Eggersdorfer M, Weber P. Dietary surveys indicate vitamin intakes below recommendations are common in representative Western countries. *Br J Nutr.* 2012;108(4):692–8.
- Christie-David D, Gunton J. Vitamin C deficiency and diabetes mellitus—easily missed? *Diabet Med.* 2017;34(2):294–6.
- Dowdle J, Ishikawa T, Gatzek S, Rolinski S, Smirnov N. Two genes in *Arabidopsis thaliana* encoding GDP-L-galactose phosphorylase are required for ascorbate biosynthesis and seedling viability. *Plant J.* 2007;52(4):673–89.
- Lim B, Smirnov N, Cobbett CS, Golz JF. Ascorbate-deficient *vrc2* mutants in *Arabidopsis* do not exhibit decreased growth. *Front Plant Sci.* 2016;7:1025.
- Conklin PL, Saracco SA, Norris SR, Last RL. Identification of ascorbic acid-deficient *Arabidopsis thaliana* mutants. *Genetics.* 2000;154(2):847–56.
- Baldet P, Bres C, Okabe Y, Mauxion J-P, Just D, Bournonville C, et al. Investigating the role of vitamin C in tomato through TILLING identification of ascorbate-deficient tomato mutants. *Plant Biotechnol.* 2013;30(3):309–14.
- Höller S, Ueda Y, Wu L, Wang Y, Hajirezaei M-R, Ghaffari M-R, et al. Ascorbate biosynthesis and its involvement in stress tolerance and plant development in rice (*Oryza sativa* L.). *Plant Mol Biol.* 2015;88(6):545–60.
- Vidal-Meireles A, Neupert J, Zsigmond L, Rosado-Souza L, Kovács L, Nagy V, et al. Regulation of ascorbate biosynthesis in green algae has evolved to enable rapid stress-induced response via the *VTC2* gene encoding GDP-L-galactose phosphorylase. *New Phytol.* 2017;214(2):668–81.
- Laing WA, Wright MA, Cooney J, Bulley SM. The missing step of the L-galactose pathway of ascorbate biosynthesis in plants, an L-galactose guanylttransferase, increases leaf ascorbate content. *Proc Natl Acad Sci U S A.* 2007;104(22):9534–9.
- Linster CL, Gomez TA, Christensen KC, Adler LN, Young BD, Brenner C, et al. *Arabidopsis VTC2* encodes a GDP-L-galactose phosphorylase, the last unknown enzyme in the Smirnoff-wheeler pathway to ascorbic acid in plants. *J Biol Chem.* 2007;282(26):18879–85.
- Bulley SM, Rassam M, Hoser D, Otto W, Schünemann N, Wright M, et al. Gene expression studies in kiwifruit and gene over-expression in *Arabidopsis* indicates that GDP-L-galactose guanylttransferase is a major control point of vitamin C biosynthesis. *J Exp Bot.* 2009;60(3):765–78.
- Zhang G-Y, Liu R-R, Zhang C-Q, Tang K-X, Sun M-F, Yan G-H, et al. Manipulation of the rice L-galactose pathway: evaluation of the effects of transgene overexpression on ascorbate accumulation and abiotic stress tolerance. *PLoS One.* 2015;10(5):e0125870.
- Zhou Y, Tao Q, Wang Z, Fan R, Li Y, Sun X, et al. Engineering ascorbic acid biosynthetic pathway in *Arabidopsis* leaves by single and double gene transformation. *Biol Plantarum.* 2012;56(3):451–7.
- Bulley S, Laing W. The regulation of ascorbate biosynthesis. *Curr Opin Plant Biol.* 2016;33:15–22.
- Bulley S, Wright M, Rommens C, Yan H, Rassam M, Lin-Wang K, et al. Enhancing ascorbate in fruits and tubers through over-expression of the L-galactose pathway gene GDP-L-galactose phosphorylase. *Plant Biotechnol J.* 2012;10(4):390–7.
- Li X, Ye J, Munir S, Yang T, Chen W, Liu G, et al. Biosynthetic gene pyramiding leads to ascorbate accumulation with enhanced oxidative stress tolerance in tomato. *Int J Mol Sci.* 2019;20(7):1558.
- Wang L, Meng X, Yang D, Ma N, Wang G, Meng Q. Overexpression of tomato *GDP-L-galactose phosphorylase* gene in tobacco improves tolerance to chilling stress. *Plant Cell Rep.* 2014;33(9):1441–51.
- Ali B, Pantha S, Acharya R, Ueda Y, Wu L-B, Ashrafuzzaman M, et al. Enhanced ascorbate level improves multi-stress tolerance in a widely grown *indica* rice variety without compromising its agronomic characteristics. *J Plant Physiol.* 2019;40:152998.
- Hellens RP, Brown CM, Chisnall MA, Waterhouse PM, Macknight RC. The emerging world of small ORFs. *Trends Plant Sci.* 2016;21(4):317–28.
- Laing WA, Martínez-Sánchez M, Wright MA, Bulley SM, Brewster D, Dare AP, et al. An upstream open reading frame is essential for feedback regulation of ascorbate biosynthesis in *Arabidopsis*. *Plant Cell.* 2015;27(3):772–86.
- Li T, Yang X, Yu Y, Si X, Zhai X, Zhang H, et al. Domestication of wild tomato is accelerated by genome editing. *Nat Biotechnol.* 2018;36:1160–3.
- Zhang H, Si X, Ji X, Fan R, Liu J, Chen K, et al. Genome editing of upstream open reading frames enables translational control in plants. *Nat Biotechnol.* 2018;36:894–8.
- Hayden CA, Jorgensen RA. Identification of novel conserved peptide uORF homology groups in *Arabidopsis* and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. *BMC Biol.* 2007;5(1):32.
- Tran MK, Schultz CJ, Baumann U. Conserved upstream open reading frames in higher plants. *BMC Genomics.* 2008;9(1):361.

38. Guerrero-González ML, Rodríguez-Kessler M, Jiménez-Bremont JF. uORF, a regulatory mechanism of the Arabidopsis polyamine oxidase 2. *Mol Biol Rep.* 2014;41(4):2427–43.
39. Hanfrey C, Elliott KA, Franceschetti M, Mayer MJ, Illingworth C, Michael AJ. A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. *J Biol Chem.* 2005;280:39229–39237.
40. Hanfrey C, Franceschetti M, Mayer MJ, Illingworth C, Michael AJ. Abrogation of upstream open reading frame-mediated translational control of a plant S-adenosylmethionine decarboxylase results in polyamine disruption and growth perturbations. *J Biol Chem.* 2002;277:44131–44139.
41. Hu W-W, Gong H, Pua EC. The pivotal roles of the plant S-adenosylmethionine decarboxylase 5' untranslated leader sequence in regulation of gene expression at the transcriptional and posttranscriptional levels. *Plant Physiol.* 2005;138(1):276–86.
42. Kwak S-H, Lee SH. The regulation of ornithine decarboxylase gene expression by sucrose and small upstream open reading frame in tomato (*Lycopersicon esculentum* mill). *Plant Cell Physiol.* 2001;42(3):314–23.
43. Rahmani F, Hummel M, Schuurmans J, Wiese-Klinkenberg A, Smeekens S, Hanson J. Sucrose control of translation mediated by an upstream open reading frame-encoded peptide. *Plant Physiol.* 2009;150(3):1356–67.
44. Thaler SK, Berberich T, Lee SS, Yang SH, Zhu X, Imai R, et al. Deregulation of sucrose-controlled translation of a bZIP-type transcription factor results in sucrose accumulation in leaves. *PLoS One.* 2012;7(3):e33111.
45. Alatorre-Cobos F, Cruz-Ramírez A, Hayden CA, Pérez-Torres C-A, Chauvin A-L, Ibarra-Laclette E, et al. Translational regulation of *Arabidopsis XIPOT1* is modulated by phosphocholine levels via the phylogenetically conserved upstream open reading frame 30. *J Exp Bot.* 2012;63(14):5203–21.
46. Zhu X, Li Y, Fang W, Kusano T. Galactinol is involved in sequence-conserved upstream open reading frame-mediated repression of *Arabidopsis HsfB1* translation. *Environ Exp Bot.* 2018;156:120–9.
47. Sanahuja G, Farré G, Bassie L, Zhu C, Christou P, Capell T. Ascorbic acid synthesis and metabolism in maize are subject to complex and genotype-dependent feedback regulation during endosperm development. *Biotechnol J.* 2013;8(10):1221–30.
48. IWGSC. Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. *Science.* 2018;361:661.
49. Abhinandan K, Skori L, Stanic M, Hickerson NM, Jamshed M, Samuel MA. Abiotic stress signaling in wheat—an inclusive overview of hormonal interactions during abiotic stress responses in wheat. *Front Plant Sci.* 2018;9:734.
50. USDA Food Composition Databases. <https://ndb.nal.usda.gov/ndb/>.
51. Liu C, Atkinson M, Chinoy C, Devos K, Gale M. Nonhomoeologous translocations between group 4, 5 and 7 chromosomes within wheat and rye. *Theor Appl Genet.* 1992;83(3):305–12.
52. Akhunov ED, Sehgal S, Liang H, Wang S, Akhunova AR, Kaur G, et al. Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. *Plant Physiol.* 2013;161(1):252–65.
53. Brenner C, Hint, Fhit, and GalT: function, structure, evolution, and mechanism of three branches of the histidine triad superfamily of nucleotide hydrolases and transferases. *Biochemistry.* 2002;41(29):9003–14.
54. Müller-Moulé P. An expression analysis of the ascorbate biosynthesis enzyme VTC2. *Plant Mol Biol.* 2008;68:31–41.
55. Wang L-Y, Li D, Deng Y-S, Lv W, Meng Q-W. Antisense-mediated depletion of tomato GDP-L-galactose phosphorylase increases susceptibility to chilling stress. *J Plant Physiol.* 2013;170(3):303–14.
56. Yang D-Y, Ma N-N, Zhuang K-Y, Zhu S-B, Liu Z-M, Yang X-H. Overexpression of tomato *SIGGP-LIKE* gene improves tobacco tolerance to methyl viologen-mediated oxidative stress. *J Plant Physiol.* 2017;209:31–41.
57. Wicker T, Wing RA, Schubert I. Recurrent sequence exchange between homeologous grass chromosomes. *Plant J.* 2015;84(4):747–59.
58. Cerbin S, Jiang N. Duplication of host genes by transposable elements. *Curr Opin Genet Dev.* 2018;49:63–9.
59. Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR. Pack-MULE transposable elements mediate gene evolution in plants. *Nature.* 2004;431:569–79.
60. Hegde RR. Differential distribution of ascorbic acid and RNA in the developing anthers of *Datura stramonium* L. *Bot Mag Tokyo.* 1985; 98(3):219–23.
61. Kumar RR, Goswami S, Gadpayle KA, Singh K, Sharma SK, Singh G, et al. Ascorbic acid at pre-anthesis modulate the thermotolerance level of wheat (*Triticum aestivum*) pollen under heat stress. *J Plant Biochem Biot.* 2014; 23(3):293–306.
62. Zhao D, Hamilton JP, Vaillancourt B, Zhang W, Eizenga GC, Cui Y, et al. The unique epigenetic features of pack-MULEs and their impact on chromosomal base composition and expression spectrum. *Nucleic Acids Res.* 2018;46(5):2380–97.
63. Asseng S, Ewert F, Martre P, Rötter RP, Lobell D, Cammarano D, et al. Rising temperatures reduce global wheat production. *Nat Clim Chang.* 2015;5:143–7.
64. Afzal I, Basra SM, Farooq M, Nawaz A. Alleviation of salinity stress in spring wheat by hormonal priming with ABA, salicylic acid and ascorbic acid. *Int J Agric Biol.* 2006;8(1):23–8.
65. Al-Hakimi A, Hamada A. Counteraction of salinity stress on wheat plants by grain soaking in ascorbic acid, thiamin or sodium salicylate. *Biol Plantarum.* 2001;44(2):253–61.
66. Athar H-U-R, Khan A, Ashraf M. Inducing salt tolerance in wheat by exogenously applied ascorbic acid through different modes. *J Plant Nutr.* 2009;32(11):1799–817.
67. Hafez E, Gharib H. Effect of exogenous application of ascorbic acid on physiological and biochemical characteristics of wheat under water stress. *Int J Plant Prod.* 2016;10(4):579–96.
68. Khan A, Ashraf M. Exogenously applied ascorbic acid alleviates salt-induced oxidative stress in wheat. *Environ Exp Bot.* 2008;63(1–3):224–31.
69. Malik S, Ashraf M. Exogenous application of ascorbic acid stimulates growth and photosynthesis of wheat (*Triticum aestivum* L.) under drought. *Soil & Environ.* 2012;31(1).
70. Every D, Griffin W, Wilson P. Ascorbate oxidase, protein disulfide isomerase, ascorbic acid, dehydroascorbic acid and protein levels in developing wheat kernels and their relationship to protein disulfide bond formation. *Cereal Chem.* 2003;80(1):35–9.
71. Paradiso A, De Pinto M, Locato V, De Gara L. Galactone-γ-lactone-dependent ascorbate biosynthesis alters wheat kernel maturation. *Plant Biol.* 2012;14(4):652–8.
72. De Gara L, De Pinto M, Arrigoni O. Ascorbate synthesis and ascorbate peroxidase activity during the early stage of wheat germination. *Physiol Plant.* 1997;100(4):894–900.
73. Yang TK, Basu B, Ooraikul F. Studies on germination conditions and antioxidant contents of wheat grain. *Int J Food Sci Nutr.* 2001;52(4):319–30.
74. Kapustin Y, Souvorov A, Tatusova T, Lipman D. Splign: algorithms for computing spliced alignments with identification of paralogs. *Biol Direct.* 2008;3(1):20.
75. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics.* 2007;23(10):1289–91.
76. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 2012; 40(15):e115.
77. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 2010;59(3):307–21.
78. Schreiber AW, Sutton T, Caldo RA, Kalashyan E, Lovell B, Mayo G, et al. Comparative transcriptomics in the Triticeae. *BMC Genomics.* 2009;10(1):285.
79. Beasley JT, Bonneau JP, Johnson AA. Characterisation of the nicotianamine aminotransferase and deoxymugineic acid synthase genes essential to Strategy II iron uptake in bread wheat (*Triticum aestivum* L.). *PLoS One.* 2017;12(5):e0177061.
80. Sears ER. The aneuploids of common wheat. *Mo Agr Exp Sta Res Bull.* 1954; 572:1–58.
81. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 2002; 3(7):research0034.
82. Ramírez-González R, Borrill P, Lang D, Harrington S, Brinton J, Venturini L, et al. The transcriptional landscape of polyploid wheat. *Science.* 2018;361:662.
83. Druka A, Muehlbauer G, Druka I, Caldo R, Baumann U, Rostoks N, et al. An atlas of gene expression from seed to seed through barley development. *Funct Integr Genomics.* 2006;6(3):202–11.
84. Sekhon RS, Lin H, Childs KL, Hansey CN, Buell CR, de Leon N, et al. Genome-wide atlas of transcription during maize development. *Plant J.* 2011;66(4):553–63.
85. Davidson RM, Gowda M, Moghe G, Lin H, Vaillancourt B, Shiu SH, et al. Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. *Plant J.* 2012;71(3):492–502.

86. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013; 30(4):772–80.
87. Curtis MD, Grossniklaus U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* 2003; 133(2):462–9.
88. Lindbo JA. High-efficiency protein expression in plants from agroinfection-compatible tobacco mosaic virus expression vectors. *BMC Biotechnol.* 2007;7(1):52.
89. Bally J, Jung H, Mortimer C, Naim F, Philips JG, Hellens R, et al. The rise and rise of *Nicotiana benthamiana*: a plant for all reasons. *Annu Rev Phytopathol.* 2018;56:405–26.
90. Philips JG, Dudley KJ, Waterhouse PM, Hellens RP. The rapid methylation of T-DNAs upon *Agrobacterium* inoculation in plant leaves. *Front Plant Sci.* 2019;10:312.
91. Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol Biol.* 2000;42(6):819–32.
92. Rassam M, Laing W. Variation in ascorbic acid and oxalate levels in the fruit of *Actinidia chinensis* tissues and genotypes. *J Agric Food Chem.* 2005;53(6):2322–6.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

