



THE UNIVERSITY OF
NEWCASTLE
AUSTRALIA

DOCTORAL THESIS

**The Molecular Characterisation of the
Vernalisation Response in Safflower via the
Development of Genomic and
Transcriptomic Resources**

Author:

Darren CULLERNE

BAppSci/BIT

MBiotech(Research)

Supervisors:

Dr. Andy EAMENS

Dr. Craig WOOD

Dr. Ben TREVASKIS

Prof. Christopher GROF

*A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy (Biological Sciences)*

in the

Faculty of Science and Information Technology
School of Environmental and Life Sciences

May 3, 2017

Statements of Collaboration and Originality

I, Darren CULLERNE, declare that:

- (a) This thesis, titled 'The Molecular Characterisation of the Vernalisation Response in Safflower via the Development of Genomic and Transcriptomic Resources' contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to the final version of my thesis being made available worldwide when deposited in the University's Digital Repository**, subject to the provisions of the Copyright Act 1968.

**Unless an Embargo has been approved for a determined period.

- (b) The work embodied in this thesis has been conducted in collaboration with and carried out at CSIRO Black Mountain in Canberra, in association with the University of Newcastle. As part of this collaboration, I have undertaken the research contained within this Thesis using CSIRO facilities under Dr Craig Wood in the Safflower Engineering laboratories.

Signed:

Date:

Acknowledgements

Firstly, I would like to thank my supervisors, Dr Craig Wood, Dr Ben Trevaskis, Dr Andy Eamens and Professor Chris Grof. Each of these people provided a unique depth of knowledge and expertise in their guiding contributions to this project.

I would like to acknowledge: The University of Newcastle, the Grains Research and Development Corporation and the CSIRO's Office of the Chief Executive for providing both operating funds and scholarship stipends. My involvement in this project would not have been possible without their financial support.

The bioinformatics community at CSIRO. In particular, my conversations with Dr Stuart Stephen, Mr Andrew Spriggs and Dr Jen Taylor, many of which were facilitated by the consumption of coffee. They proved to be invaluable for the correct planning and analysis of the countless samples and data files processed as part of this project.

The many postdoctoral scientists and technicians in the CSIRO Plant and Oil Engineering and the CSIRO Crop Genomics groups. Their suggestions of protocols, troubleshooting help, molecular biology techniques and assistance in getting my hands dirty on the bench and in the glasshouses is much appreciated.

The other PhD students, both within and external to CSIRO, some who have already graduated and others who yet to complete their projects. Their camaraderie and support helped me through some rough times during this project. I hope I have helped them in the same way. I would like to particularly thank Kyle Reynolds for his mental support, baked goods and coffee breaks throughout this project.

My family and friends for their ongoing support of my partner and I. Whether the support was financial, psychological, fermented or nutritional, I could not have accomplished this project without them.

Lastly, but most importantly, I'd like to thank my partner Michelle for support and encouragement throughout my studies and during this project in particular. Her endless patience and unwavering belief in me were crucial to getting me over the finishing line.

*To my grandparents, Eileen Dotridge Comans, Thomas Joseph
Bailey and Monica Ann Bailey and my great uncle Edmund
Theopolis Cullerne. You have uniquely influenced, supported, and
inspired me to excel.*

Contents

Statements of Collaboration and Originality	i
Acknowledgements	ii
List of Figures	x
List of Tables	xii
Abbreviations	xiv
Gene and Protein Abbreviations	xvi
Abstract	xvii
1 Introduction	1
1.1 Safflower: An Ancient Crop with a Future Based on Biotechnology	1
1.2 The Vernalisation Response	3
1.3 Mechanisms of the Vernalisation Response	6
1.3.1 <i>Arabidopsis thaliana</i> use FLC (and MAFs)	6
1.3.2 Legumes use a Family of FTs	9
1.3.3 Sugar Beet uses BTC1	10
1.3.4 Cereals use VRN1, VRN2 and FT	10
1.3.5 What Regulates the <i>Asteraceae</i> ?	12
1.4 Phylogenetic Analysis of Vernalisation Responsive Species	14
1.5 Further Questions on the Vernalisation Response in Safflower	15
1.6 Next Generation Sequencing in the Context of the Vernalisation Response in Safflower	16
1.7 Summary	17
2 Physiology of the Vernalisation Response in Safflower Varieties	19
2.1 Outline	19
2.2 Materials and Methods	19
2.2.1 Cultivars	19
2.2.2 Growth Conditions	20
2.2.2.1 Breaking Seed Dormancy	20
2.2.2.2 Petri Dish Germination and Vernalisation	20
2.2.2.3 Measuring Cylinder Germination and Vernalisation	20
2.2.2.4 Glasshouse Growth Conditions	20
2.2.2.5 Growth Cabinet Conditions	21
2.2.3 Generation of Crossing Population	21

2.2.4	Characterisation of Vernalisation and Day Length Effect	23
2.2.5	Time to Vernalisation Saturation	23
2.2.6	Optimum Vernalisation Temperature	23
2.3	Results	24
2.3.1	Initial Characterisation of the Vernalisation Response in Winter Hardy Safflower	24
2.3.2	Resetting of the Vernalisation Response in Safflower	29
2.3.3	Vernalisation Exposure Timecourse to Determine the Saturation Point of the Vernalisation Response	29
2.3.4	Temperature Timecourse to Determine the Vernalisation Temperature for Safflower	31
2.3.5	Inheritance of the Vernalisation Phenotype in Safflower	33
2.4	Discussion	36
2.4.1	The Vernalisation Response is Observed in Winter Safflower Only	36
2.4.2	The Vernalisation Response in Safflower is Recessive	37
2.4.3	The Vernalisation Response in Safflower is Epigenetic and Resets in the Next Generation	38
2.5	Conclusion	39
3	Transcriptomic Analysis of the Vernalisation Response in Safflower	40
3.1	Outline	40
3.2	Materials and Methods	40
3.2.1	Selection of RNA Extraction Protocol	40
3.2.1.1	PureLink Based Method	41
3.2.1.2	Qiagen RNeasy Kit	41
3.2.1.3	TRIzol Based Method (Manufacturers Protocol)	41
3.2.1.4	Cetyl Trimethyl Ammonium Bromide (CTAB) Based Method	41
3.2.1.5	Hot Phenol Based Method	41
3.2.1.6	TRIzol Based Method (Modified from Manufacturers Method)	42
3.2.2	Primer Design and RT-qPCR Protocol	42
3.2.3	Assembly of <i>De Novo</i> Transcriptomic References for Safflower . . .	43
3.2.3.1	RNA Extraction and Sequencing of Spring Safflower . . .	43
3.2.3.2	Pre-processing of Spring Safflower Reads	44
3.2.3.3	<i>De Novo</i> Assembly of Spring Safflower Reads	44
3.2.3.4	Quality Assessment of the <i>De Novo</i> Spring Safflower Assembly	44
3.2.3.5	Back Alignment of Spring Safflower Reads to the <i>De Novo</i> Reference	45
3.2.4	Assembly of a <i>De Novo</i> Winter Safflower Assembly	45
3.2.5	Aligning the Spring and Winter <i>De Novo</i> Assemblies	45

3.2.6	Differential Expression (DE) Analysis	46
3.2.6.1	Growth Conditions for DE	46
3.2.6.2	Experiment 1: Winter Safflower Before and After Vernalisation	46
3.2.6.3	Experiment 2: Vernalisation of Safflower at Five Time Points	46
3.2.6.4	Analysis of Back Alignment Data	47
3.2.6.5	Annotating Differentially Expressed Transcripts	48
3.3	Results	49
3.3.1	Assessment of RNA Extraction Methods for Safflower Leaf Tissue	49
3.3.2	Assembly of the <i>De Novo</i> Spring Safflower Transcriptomic Reference	50
3.3.2.1	Pre-processing of Reads from Safflower Tissues	50
3.3.2.2	Assembly of the <i>De Novo</i> Transcriptomic Reference	50
3.3.2.3	Quality Assessment with CEGMA and BUSCO	51
3.3.2.4	Quality Assessment using the <i>CtFAD2</i> Gene Family	51
3.3.2.5	Assessment of the Winter Safflower <i>De Novo</i> Assembly	54
3.3.2.6	Aligning the Spring and Winter Safflower Transcriptomic Assemblies	55
3.3.3	Experiment 1: Differentially Expressed Transcripts Before and After Vernalisation	55
3.3.4	Experiment 2: Vernalisation at Five Time Points	63
3.3.5	Alignment of Annotated Sequences from Spring and Winter Safflower	68
3.3.5.1	<i>APETALA 1-LIKE (CtAP1-LIKE)</i>	68
3.3.5.2	<i>FLOWERING LOCUS T-LIKE (CtFT-LIKE)</i>	68
3.3.5.3	<i>MADS-BOX DOMAIN CONTAINING 1 (CtMADS1)</i>	69
3.3.5.4	<i>VERNALISATION 1-LIKE (CtVRN1-LIKE)</i>	69
3.3.6	RT-qPCR Validation of RNA-seq Data	69
3.4	Discussion	74
3.4.1	The Creation of a Reference Transcriptomic Assembly for Safflower	74
3.4.2	Differentially Expressed Transcripts During the Vernalisation Response	75
3.4.3	Characterisation of Genes in the Vernalisation Response	76
3.5	Further Investigations	77
3.6	Conclusion	78
4	Genomic Basis of the Vernalisation Response in Safflower	80
4.1	Outline	80
4.2	Materials and Methods	81
4.2.1	Cultivars and Growth Conditions	81
4.2.2	Extraction of Nuclear Genomic DNA	81
4.2.2.1	Preparation of the Nuclear Extraction Buffer (NEB)	81

4.2.2.2	Isolation of the Nuclei	81
4.2.2.3	Extraction of Nuclear Genomic DNA	82
4.2.3	<i>De Novo</i> Assembly using Illumina Reads	83
4.2.3.1	Illumina Sequencing	83
4.2.3.2	Pre-Processing of Illumina Reads	83
4.2.3.3	Assembly of Illumina Reads	83
4.2.3.4	Scaffolding Using Library Information	83
4.2.3.5	Scaffolding Using the Spring Safflower <i>De Novo</i> Transcriptome	84
4.2.3.6	Back Alignment of Illumina Reads	84
4.2.4	<i>De Novo</i> Assembly using Pacific Biosciences (PacBio) Reads	84
4.2.4.1	PacBio Sequencing	84
4.2.4.2	Error Correction of PacBio Reads	86
4.2.4.3	Assembly of Error Corrected PacBio Reads	86
4.2.4.4	Analysis of the Assembly of Library 1	86
4.2.5	Generation of SNP-based Markers for the Vernalisation Response .	87
4.2.5.1	Scoring of F ₃ Phenotypes	87
4.2.5.2	Generation of Markers by DArT	89
4.2.5.3	Comparison of Markers Across Families	89
4.2.5.4	Mapping of Markers	89
4.3	Results	91
4.3.1	A High Quality Draft Assembly of the Safflower Genome	91
4.3.1.1	Pre-processing of Illumina Reads	91
4.3.1.2	Assembly and Scaffolding	91
4.3.1.3	Quality Assessment of the Assembly	92
4.3.1.4	Back Alignment as a Method of Quality Assessment . . .	92
4.3.2	Determining Intron/Exon Boundaries for Identified Vernalisation Genes	93
4.3.2.1	<i>APETALA 1-LIKE (CtAP1-LIKE)</i>	93
4.3.2.2	<i>FLOWERING LOCUS T-LIKE (CtFT-LIKE)</i>	94
4.3.2.3	<i>MADS BOX DOMAIN CONTAINING 1 (CtMADS1)</i> . . .	95
4.3.2.4	<i>VERNALISATION 1-LIKE (CtVRN1-LIKE)</i>	96
4.3.3	The PacBio <i>De Novo</i> Assemblies	96
4.3.4	Library 1: A Draft Safflower Chloroplast	99
4.3.5	Library 2: A Work in Progress	104
4.3.6	DNA Based Markers of Vernalisation in Safflower	106
4.3.6.1	Aligning Differentially Expressed Transcripts onto the Genetic Map	107
4.4	Discussion	109
4.4.1	The <i>De Novo</i> Assemblies	109
4.4.2	Intron/Exon Boundaries for Genes Annotated in the Vernalisation Response	111

4.4.3	Genetic Markers of the Vernalisation Response in Safflower	112
4.4.4	The Curious Case of the Safflower Chloroplast	113
4.4.5	Future Directions	114
4.5	Conclusion	116
5	Overall Discussion	118
	References	125
	Appendix A Significantly Differentially Expressed Spring Safflower Transcripts from Experiment 2	136
	Appendix B Differential Expression Plots	141
	Appendix C Characterised Differentially Expressed Vernalisation Transcripts	145
C.1	Spring Safflower Vernalisation Transcripts	145
C.2	Winter Safflower Vernalisation Transcripts	147
	Appendix D Winter safflower transcripts	150
	Appendix E Multiple Sequence Alignments of Annotated Safflower Transcripts	151
	Appendix F PCR Primers	155
	Appendix G Read Alignments	156
	Appendix H Spring:Winter Segregation Ratios for Crossing Population	160
	Appendix I Molecular Markers of the Vernalisation Response in Safflower	165
	Appendix J Software Parameters (Assembly)	171
J.1	Safflower Transcriptome (Spring Reference)	171
J.1.1	Trinity (Inchworm, Chrysalis, Butterfly)	171
J.1.2	Biokanga 'Align'	171
J.2	Safflower Transcriptome (Winter cultivar)	172
J.2.1	Biokanga 'Assemb'	172
J.2.2	Biokanga 'Scaffold'	172
J.2.3	Biokanga 'Scaffold'	173
J.3	Safflower Genome (Illumina)	173
J.3.1	Biokanga - 'Assemb' (PE)	173
J.3.2	Biokanga - 'Assemb' (MP)	174
J.3.3	Biokanga - 'Scaffold' (PE)	174
J.3.4	Biokanga - 'Scaffold' (MP)	174
J.3.5	Biokanga - 'Blitz'	175
J.3.6	Biokanga - 'Align' (fixed insert length)	175
J.3.7	Biokanga - 'Align' (varied insert length)	176

J.4	Safflower Chloroplast (PacBio)	176
J.4.1	PacBiokanga - 'Ecreads' (Pass 1)	176
J.4.2	PacBiokanga - 'Ecreads' (Pass 2)	177
J.4.3	PacBiokanga - 'Ecreads' (Build Overlap File)	177
J.4.4	PacBiokanga - 'Contigs'	178
J.4.5	PacBiokanga - 'Econtigs'	178
J.5	Safflower Genome (PacBio)	178
J.5.1	PacBiokanga - 'Ecreads' (Pass 1)	178
J.5.2	PacBiokanga - 'Ecreads' (Build Overlaps - EC read samples)	179

List of Figures

1.1	Global safflower production in 2014.	2
1.2	The effect of vernalisation on <i>Arabidopsis thaliana</i>	4
1.3	Gene interaction models from different angiosperms	5
1.4	The effect of vernalisation on days to bolting in lettuce.	13
1.5	Phylogenetic tree of six families of angiosperms.	14
2.1	Procedure for crossing safflower.	22
2.2	Seeds of both vernalised and unvernalsed spring and winter safflower.	24
2.3	Expressed vernalisation phenotypes.	25
2.4	Boxplot of the Q ₀ generation of spring and winter safflower.	26
2.5	Characterisation of aspects of the vernalisation response in the Q ₁ generation of winter safflower.	28
2.6	Vernalisation timecourse: Fixed 4°C temperature, lengthening exposure time.	30
2.7	Vernalisation timecourse: Fixed 28 day exposure time, varying temperatures.	32
2.8	Effects of field conditions on phenotyping the vernalisation effect on safflower F ₂ crosses.	34
3.1	Results from testing six different RNA extraction methods.	49
3.2	Phylogenetic tree of the previously characterised <i>CtFAD2</i> family of genes (Cao et al. 2013) and transcripts from the spring safflower <i>de novo</i> transcriptome.	53
3.3	All significantly differentially expressed transcripts from Experiment 2.	64
3.4	Differential expression of transcripts in spring and winter safflower from Experiment 2.	67
3.5	Expression of the four transcripts from Experiment 1 using normalised RNA-Seq data.	71
3.6	Expression of four transcripts from Experiment 1 using RT-qPCR, normalised with <i>CtACTIN1-LIKE</i>	71
3.7	Expression of <i>CtMADS1</i> and <i>CtFT-LIKE</i> transcripts from Experiment 2 using normalised RNA-Seq data.	72
3.8	Expression of <i>CtMADS1</i> and <i>CtFT-LIKE</i> transcripts from Experiment 2 using RT-qPCR, normalised with <i>CtACTIN1-LIKE</i>	73
4.1	Method for assembly of Illumina and PacBio data.	85
4.2	The preparation of the F ₃ crossing population.	88
4.3	The process for creating digest and SNP markers.	90
4.4	Gene model for <i>CtAP1-LIKE</i>	94

4.5	The gene model for <i>CtFT-LIKE</i>	94
4.6	The gene model for <i>CtMADS1</i>	95
4.7	The gene model for <i>CtVRN1</i>	96
4.8	The distribution of the read length in PacBio Libraries 1 and 2.	97
4.9	Alignment of error corrected (pass 1) PacBio read against the draft safflower genome.	98
4.10	Alignment of the safflower <i>de novo</i> chloroplast and the chloroplast from <i>Arabidopsis</i> and sunflower.	100
4.11	A visualisation of the assembled genome of the draft safflower chloroplast.	102
4.12	A high resolution image of the draft safflower chloroplast.	103
4.13	An error corrected (pass 1) PacBio read aligned against the draft safflower genome.	105
4.14	The locations of DArT markers on the genetic map of Chromosome 8 of the safflower genome.	107
4.15	A hypothetical alignment of different assemblies.	115
B.1	Mean expression counts of every contig in the spring safflower <i>de novo</i> transcriptome.	141
B.2	Mean expression counts of every contig in the winter safflower <i>de novo</i> transcriptome.	142
B.3	Volcano plot of expression of spring safflower transcripts.	143
B.4	Volcano plot of expression of winter safflower transcripts.	144
E.1	Multiple Sequence Alignment of amino acid sequences with homology to <i>CtFT-LIKE</i>	151
E.2	Multiple Sequence Alignment of amino acid sequences with homology to <i>CtAP1-LIKE</i>	152
E.3	Multiple Sequence Alignment of amino acid sequences with homology to <i>CtMADS1</i> and other MADS-box containing sequences.	153
E.4	Multiple Sequence Alignment of amino acid sequences with homology to <i>CtVRN1-LIKE</i>	154
G.1	Back alignment of short reads generated from spring safflower tissues.	156
G.2	Reads generated for each replicate for vernalised and unvernalsed winter safflower aligned against the <i>de novo</i> transcriptome.	157
G.3	Reads generated for each replicate for winter and spring safflower in vernalisation timecourse.	158
G.4	Back alignments of each unfiltered Illumina genomic library against the <i>de novo</i> spring safflower genome.	159

List of Tables

2.1	Expressed phenotypes and growth attributes for vernalised and unvernalsed Q ₁ safflower.	27
2.2	Different gene models explaining the number of loci responsible for the vernalisation response	35
3.1	Coverage of each pair of read libraries from different safflower tissues . .	50
3.2	Attributes of the <i>de novo</i> spring safflower transcriptome.	51
3.3	CEGMA analysis on the <i>de novo</i> spring safflower transcriptome.	51
3.4	BUSCO analysis on the <i>de novo</i> spring safflower transcriptome.	51
3.5	Attributes of the winter safflower <i>de novo</i> transcriptome.	54
3.6	CEGMA analysis on the winter safflower <i>de novo</i> transcriptome.	54
3.7	BUSCO analysis on the winter safflower <i>de novo</i> transcriptome.	54
3.8	Differentially Expressed Spring Safflower Transcriptomic Contigs.	56
3.9	Key differentially expressed transcripts in winter safflower aligned to the spring safflower transcriptome.	58
3.10	Differentially expressed winter safflower transcriptomic contigs.	60
3.11	Results of the three very significantly differentially expressed winter safflower transcripts.	62
3.12	Comparison of previously published safflower trasncriptomes	74
4.1	Attributes of the CSIRO draft safflower genome.	91
4.2	CEGMA analysis on the draft safflower genome.	92
4.3	BUSCO analysis on the draft safflower genome.	92
4.4	Dimensions of the PacBio Genomic Libraries.	97
4.5	Attributes of the PacBio Library 1 assembly using PacBiokanga.	99
4.6	Attributes of the partially error corrected PacBio Library 2.	104
4.7	Digest and SNP markers reported by DArT.	106
4.8	The CSIRO draft safflower genome constructed using Biokanga compared against the draft safflower genome presented in Bowers et al. (2016).	110
A.1	Significantly differentially expressed spring safflower transcriptomic contigs from Experiment 2.	137
F.1	PCR Primers used in RT-qPCR experiments.	155
H.1	The segregation ratios of the F ₃ population of crossed safflower plants. . .	160
I.1	Illumina genomic contigs, containing digest markers, that align to SNP-containing Bowers contigs.	165

I.2	Differentially expressed transcripts from Experiment 1 that map to SNP-containing Bowers contigs.	167
I.3	Differentially expressed transcripts from Experiment 2 that map to SNP-containing Bowers contigs.	168

Abbreviations

AGRF	Australian Genome Research Facility
<i>Arabidopsis</i>	<i>Arabidopsis thaliana</i>
AraTha	<i>Arabidopsis thaliana</i>
<i>At</i>	<i>Arabidopsis thaliana</i>
BLAST	Basic Local Algorithm Search Tool (Software)
BLASTN	BLAST Nucleotide (Software)
BLASTP	BLAST Protein (Software)
BUSCO	Benchmarking Universal Single-Copy Orthologs (Software)
bp	base pair
<i>Bv</i>	<i>Beta vulgaris</i>
CarTin	<i>Carthamus tinctorius</i>
CEGMA	Core Eukaryotic Genes Mapping Approach (Software)
ChrLav	<i>Chrysanthemum lavandulifolium</i>
ChrMor	<i>Chrysanthemum morifolium</i>
<i>Ci</i>	<i>Chicory intybus</i>
cM	Centimorgans
CSIRO	Commonwealth Scientific and Industrial Research Organisation
<i>Ct</i>	<i>Carthamus tinctorius</i>
CTAB	Cetyl trimethyl ammonium bromide
DArT	Diversity Arrays Technology
DEPC	Diethylpyrocarbonate
EDTA	Ethylenediaminetetraacetic acid
Eg	<i>Eustoma grandiflorum</i>
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
E/O	Eocene / Oligocene boundary
EST	Expressed Sequence Tag
<i>g</i>	gravitational force
Gbp	Gigabase pair (1 Gbp = 1,000,000,000 bp)
GLA	gamma linolenic acid
GM	Genetic Modification
GRDC	Grains Research Development Corporation
ha	hectare
HorVar	<i>Hordeum vulgare</i>
<i>Hv</i>	<i>Hordeum vulgare</i>
MP	Mate pair (reads)
MSA	Multiple sequence alignment
<i>Mt</i>	<i>Medicago truncatula</i>
<i>M. truncatula</i>	<i>Medicago truncatula</i>

mya	Millions of years (Annus)
NCBI	National Center for B iotechnology I nformation
NEB	Nuclear E xtraction B uffer
NGS	Next g eneration s equencing
OD_{x/y}	Optical d ensity (absorbance) at wavelength <i>x</i> and <i>y</i>
PacBio	P acific B iosciences (sequencing technology)
PCR	polymerase chain r eaction
PE	Paired e nd (reads)
PHD-PRC2	Plant H omeo-Domain P olycomb R epression C omplex 2
PIPES	piperazine-N,N'-bis(2-ethanesulfonic acid)
PVP	polyvinylpyrrolidinone
RIN	RNA i ntegrity score
RPKM	R eads p er k ilobase of transcript per m illion mapped reads
RT-qPCR	R everse t ranscriptase q uantitative p olymerase chain r eaction
SAM	Shoot a pical m eristem
SCUBAT	Scaffolding C ontigs U sing B LAST-like A lignment T ool
SDS	Sodium d odecyl sulfate
SHO	Super h igh o leic
SNP	single n ucleotide p olymorphism
TAIR	The <i>A</i> rabidopsis I nformation R esource
Tris-HCl	T ris(hydroxymethyl)aminomethane hydrochloride
UTR	Untranslated r egion

Gene and Protein Abbreviations

AP1	APETALA 1
CAL	CAULIFLOWER-A
BTC1	BOLTING TIME CONTROL 1
FAD2	FATTY ACID DESATURASE 2
FD	BASIC-LEUCINE ZIPPER (bZIP) TRANSCRIPTION FACTOR
FL	FLC-LIKE (<i>Beta vulgaris</i>)
FLC	FLOWERING LOCUS C
FLCL	FLC-LIKE (<i>Eustoma</i> spp.)
FL1	FLC-LIKE 1 (<i>Chicory intybus</i>)
FRI	FRIGIDA
FT	FLOWERING LOCUS T
FTL	FT-LIKE
LFY	LEAFY
MAF1-5	MADS AFFECTING FLOWERING 1-5
OS2	ODDSOC2
PEP1	PERPETUAL FLOWERING 1
SOC1	SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1
SOC1L	SOC1-LIKE
VIN3	VERNALISATION INSENSITIVE 3
VRN1-2	VERNALISATION 1-2

Nomenclature 1 Italicised represents the DNA locus or RNA transcript of the gene.
Non-italicised represents the encoded protein.

Nomenclature 2 A lower case gene or encoded protein represents a mutant or recessive allele of that gene or protein.

Nomenclature 3 A two letter abbreviation before a gene represents the species where it is found, e.g. *At* = *Arabidopsis*.

Abstract

The Molecular Characterisation of the Vernalisation Response in Safflower via the Development of Genomic and Transcriptomic Resources

by Darren CULLERNE

Safflower (*Carthamus tinctorius*) is an oilseed grown globally. There is an interest in modifying safflower to cope with climate change and to adapt to new agronomic trends. While almost all cropped safflower are spring varieties, a number of wild safflower varieties are noted as 'winter hardy'. These display characteristics found in plants that respond to vernalisation (an extended period of non-freezing cold). Because flowering traits, including the vernalisation response, are linked to yield and adaptability to climate, this PhD project sought to understand the vernalisation response in safflower, as a component of flowering time.

The vernalisation response in 'winter hardy' and spring safflower was investigated. It was confirmed that winter safflower does respond to vernalisation conditions, similar to other plant species. The vernalisation response in winter safflower is saturated after approximately 2 weeks exposure to vernalisation conditions. It is inheritable, epigenetic and appears to be dependent on a single recessive allele, as shown by segregation ratios in a crossing population created from winter and spring parents.

Two approaches were developed to characterise this vernalisation response. Firstly, RNA sequencing (RNA-Seq) was performed on RNA extracted from winter and spring safflower before, during and after exposure to vernalisation conditions. Differential expression analyses on the resulting RNA-Seq datasets tentatively identified four genes as directing functional roles in the vernalisation response of safflower: *APETALA 1-LIKE* (*CtAP1-LIKE*), *MADS-BOX DOMAIN CONTAINING 1* (*CtMADS1*), *FLOWERING LOCUS T-LIKE* (*CtFT-LIKE*) and *VERNALISATION 1-LIKE* (*CtVRN1-LIKE*). This analysis also identified 33 additional gene products (annotated transcripts or transcripts of unknown function) as candidates for further experimental investigation.

In addition to the transcriptomic data, genomic resources were developed to further characterise the molecular basis of the vernalisation response. A high quality *de novo* assembly was constructed using Illumina reads from spring safflower, covering approximately 80% of the estimated 1,400,000,000 base pair (1.4 Gbp) spring safflower genome. Using this draft genome in combination with F₃ crossing families and a genetic marker approach, 27 genetic markers for vernalisation were identified. Furthermore, these markers were mapped to a recent genetic map of safflower (Bowers 2016), clustering in close proximity to one another on chromosome 8. A single differentially expressed transcript, identified in the transcriptomic analyses, was located on the same chromosome. However, the transcript of interest was mapped to a chromosome 8 position some distance away from the identified markers.

These high quality transcriptomic and genomic resources were used to identify the molecular basis for vernalisation in safflower. The investigative approaches developed in this project can also be utilised to characterise the molecular mechanisms of other traits in safflower.

