Southern Cross University
ePublications@SCU

Theses

2009

Gene expression analysis of malting barley

Jessica Frances White
Southern Cross University

Publication details
White, JF 2007, 'Gene expression analysis of malting barley', PhD thesis, Southern Cross University, Lismore, NSW.
Copyright JF White 2009

ePublications@SCU is an electronic repository administered by Southern Cross University Library. Its goal is to capture and preserve the intellectual output of Southern Cross University authors and researchers, and to increase visibility and impact through open access to researchers around the world. For further information please contact epubs@scu.edu.au.

Gene expression analysis of malting barley

Jessica Frances White

B.App.Sc (Biotechnology)

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

Centre for Plant Conservation Genetics

Southern Cross University

September 2008

Thesis declaration

I certify that the work presented in this thesis is, to the best of my knowledge and belief, original, except as acknowledged in the text, and that the material has not been submitted, either in whole or in part, for a degree at this or any other university. I acknowledge that I have read and understood the University's rules, requirements, procedures and policy relating to my higher degree research award and to my thesis. I certify that I have complied with the rules, requirements, procedures and policy of the University (as they may be from time to time).

Name _____

Signature of the candidate _____

Date _____

Acknowledgements

I thank my principal supervisor, Professor Robert Henry, for providing the opportunity to undertake, and the support to complete this project. I would also like to thank my co-supervisors, Dr Toni Pacey-Miller and Dr Peter Bundock, for their support and advice.

I would like to thank my family, my parents, Margaret and Terry, and my brother and sisters, Andrew, Emily and Gabrielle, for their support, encouragement and understanding during the course of my study.

I acknowledge the journal reviewers and editors from the Plant Biotechnology Journal and the journal Plant Science whose comments assisted in the writing of papers published from this thesis.

The study was financially supported by a Grain Foods CRC Ltd research scholarship.

Abstract

Malting barley (*Hordeum vulgare*) is of significant importance and economic value to the malting and brewing industries. The biochemical and physiological processes of barley seed germination (malting) have subsequently been the focus of substantial research; however the underlying molecular processes require further investigation. The introduction of large-scale gene expression technologies has provided an opportunity to contribute significant genetic data to the understanding of the malting process. Long Serial Analysis of Gene Expression (LongSAGE) was applied to malting barley to characterise the transcriptional profile exhibited during grain germination. Eight LongSAGE tag libraries were constructed, one from whole dry seed and one from each of the time points 0, 12, 24, 48, 72, 96 and 120 hours post-steeping. Approximately 20 000 tags were sequenced from each LongSAGE library, generating 41 909 unique tag sequences. Annotations of the 100 most abundant LongSAGE tags from each library were analysed to reveal the major functional groups expressed in the malting transcriptome. Statistical analysis confidently identified 57 LongSAGE tags with significant changes in tag abundance across time. The genes that correspond to eight of these tags were targeted for validation by real-time RT-PCR. Real-time RT-PCR experiments were conducted using RNA extracted from the same grain sample from four of the time points examined by LongSAGE; whole dry seed, 12 hours post-steeping, 48 hours post-steeping and 96 hours post-steeping. Three of the transcripts show co-ordinated expression demonstrating significant up-regulation at 48 hours post-steeping and remain significantly elevated during malting. These transcripts α-amylase type B, (1-3,1-4)-β-D-glucanase and cysteine proteinase EP-B represent key starch, cell wall and protein modifying

enzymes known to play significant roles in the malting process. mRNA abundance levels observed by real-time RT-PCR correlate well with the data obtained from LongSAGE.

Microarray is another large scale gene expression technology with the potential to generate large amounts of genetic data quickly and efficiently. A CombiMatrix 12K ElectraSense™ Custom Array chip was designed from the unique LongSAGE tags from 24 hours post-steeping. RNA from the malting barley 12 hours post-steeping and 48 hours post-steeping samples was hybridised to the array chips. The CombiMatrix ElectraSense™ system incorporates electrochemical detection of a redox reaction proximal to probe specific electrodes. 113 probes show a greater than 10 fold difference between 12 and 48 hours post-steeping with 64 probes preferentially expressed at 12 hours post-steeping and 49 probes preferentially expressed at 48 hours post-steeping. The eight genes of interest analysed by real-time RT-PCR were represented by one or more probes on the malting barley CombiMatrix ElectraSense™ array. Fold changes observed between 12 and 48 hours post-steeping by the CombiMatrix ElectraSense™ microarray support the fold changes observed by LongSAGE and real-time RT-PCR.

List of publications included as part of the thesis

I warrant that I have obtained permission from the copyright owners to include my own published journal articles, in which the copyright is held by the publisher, in the thesis.

White J., Pacey-Miller T., Crawford A., Cordeiro G., Barbary D., Bundock P. and Henry R. J. (2006) Abundant transcripts of malting barley identified by serial analysis of gene expression (SAGE). Plant Biotechnology Journal 4:289-301.

White J., Pacey-Miller T., Bundock P. and Henry R. J. (2008) Differential LongSAGE tag abundance analysis in a barley seed germination time course and validation with relative real-time RT-PCR. Plant Science 175:858-867.

Statement of contribution of others

The input of co-authors to the published chapters of this thesis is acknowledged by the author. The author would like to recognize the contribution of co-authors to Chapter Two "Abundant transcripts of malting barley identified by serial analysis of gene expression (SAGE)", published in Plant Biotechnology Journal. T Pacey-Miller assisted with extraction of mRNA, A Crawford prepared the SAGE libraries, excluding HvTaM4 which was prepared by J White, G Cordeiro and D Barbary assisted with databases and annotations, and P Bundock prepared Table 2.5 and comments about antisense transcripts. Comments and suggestions for improving the manuscript were provided by T Pacey-Miller, P Bundock and R Henry, all other work was completed by J White.

The author would like to acknowledge the contribution of co-authors to Chapter Three "Differential LongSAGE tag abundance analysis in a barley seed germination time course and validation with relative real-time RT-PCR" accepted for publication in Plant Science. All laboratory work and analysis was carried out by J White. Comments and suggestions for improving the manuscript were provided by T Pacey-Miller, P Bundock and R Henry.

The author would like to recognize the contribution of co-authors to Chapter Four "Transcriptome analysis of malting barley; comparison of Microarray, LongSAGE and real-time RT-PCR" submitted for publication in Functional and Integrative Genomics. Slide preparation and chip hybridisation was carried out by G Good and P Bundock. All other work was carried out by J White and comments and suggestions

for improving the manuscript were provided by T Pacey-Miller, P Bundock and R Henry.

Signature of the candidate _____

Signature of the supervisor _____

List of additional publications by the candidate

Journal Publications

McIntosh S., Watson L., Bundock P., Crawford A., **White J.**, Cordeiro G., Barbary D., Rooke L. and Henry R. (2007) SAGE of the developing wheat Caryopsis. Plant Biotechnology Journal 5:69-83

Crawford A.C., **White J.**, Bundock P., Cordeiro G., McIntosh S., Pacey-Miller T., Rooke L., Henry R.J. (2005) Consistent production of cost-effective LongSAGE libraries. Plant Molecular Biology Reporter 23:139-143

Conference Papers

Pacey-Miller T., **White J.**, Crawford A.C., Cordeiro G.M., Barbary D., Bundock P.C. and Henry R.J. (2007) Unlocking the mysteries of malting using SAGE. International Plant and Animal Genomes Conference XV. San Diego CA, USA. 13-17 January 2007.

White J., Pacey-Miller T., Crawford A.C., Bundock P.C., Cordeiro G.M. and Henry R.J. (2007) Characterisation of the Malting Barley Transcriptome. 13th Australian Barley Technical Symposium. Fremantle WA, Australia. 26-30 August 2007.

Bundock P.C., Ablett G., Cordeiro G., Crawford A.C., Henry R.J., McIntosh S., Nguyen L., Pacey-Miller T., Watson L. and **White, J.** (2006) Using LongSAGE to

study cereal gene expression during grain development and germination. Plant and Animal Genomes Conference XIV. San Diego CA, USA. 14 - 18 January 2006.

Pacey-Miller T., Watson L., **White J.**, Crawford A.C., Bundock P.C., Cordeiro G.M., Barbary D., McIntosh S. & Henry R.J. (2006) Malting quality improvement using SAGE. 56th Australian Cereal Chemistry Conference. Fremantle WA, Australia. 10-14 September 2006.

White J., Pacey-Miller T., Crawford A.C., Bundock P.C., Cordeiro G.M. and Henry R.J. (2006) Characterisation of the Malting Barley Transcriptome. PRAXIS ’06, CRUX’S 5th Annual Interdisciplinary Postgraduate Conference. Lismore NSW, Australia. 16-17 November 2006.

White J., Pacey-Miller T., Crawford A.C., Cordeiro G.M., Barbary D., Bundock P.C. and Henry R.J. (2006) Serial analysis of gene expression of malting barley. ISPMB 8th International Congress of Plant Molecular Biology. Adelaide SA, Australia. 20-25 August 2006.

White J., Pacey-Miller T., Crawford A.C., Cordeiro G., Barbary D., Bundock P.C. and Henry R.J. (2006) Characterisation of the Malting Barley Transcriptome using SAGE. Plant and Animal Genomes Conference XIV. San Diego, USA. 14-18 January 2006.

Pacey-Miller T., **White J.**, Crawford A., Bundock P., Cordeiro G.M., Barbary D. and Henry R.J. (2005) Unravelling the Mysteries of Germination using SAGE (Serial

Analysis of Gene Expression). 18th Northern American Barley Researchers Workshop and 4th Canadian Barley Symposium. Red Deer, Alberta, Canada. 17 -20 July 2005.

Pacey-Miller T., **White J.**, Crawford A.C., Cordeiro G.M., Barbary D., Bundock P.C. and Henry R.J. (2005) SAGE and the Mysteries of Malting. 12th Australian Barley Technical Symposium. Hobart, Australia. 10-13 September 2005.

White J., Pacey-Miller T., Crawford A.C., Bundock P.C., Cordeiro G.M. and Henry R.J. (2005) Serial Analysis of Cereals. 12th Australian Barley Technical Symposium. Hobart, Australia. 10-13 September 2005.

Table of Contents

Thesis declaration i

Acknowledgements ii

Abstract iii

List of publications included as part of the thesis v

Statement of contribution of others vi

List of additional publications by the candidate viii

List of Tables xv

List of Figures xvii

Chapter One

Literature Review and General Introduction – Analysis of Gene Expression in Germinating Barley

1.1 Barley 1

1.1.1 The Importance of Barley 1

1.1.2 Barley Taxonomy and Origin 3

1.1.3 Morphology of barley inflorescence 4

1.2 Seed Germination 5

1.3 Barley Malting 6

1.3.1 Malting 6

1.3.2 Steeping 7

1.3.3 Germination 8

1.3.4 Kilning 11

1.4 Gene Expression 11

1.5 Molecular Biology Techniques 12

1.5.1 Genomics 12

1.5.2 Microarray 13

1.6 Serial Analysis of Gene Expression (SAGE) 16

1.6.1 Serial Analysis of Gene Expression 16

1.6.2 SAGE in plants 20

1.7 Real-time reverse transcriptase PCR 25

1.8 Gene Expression in Barley 27

1.9 Aims and Objectives 31

1.10 References 33

Chapter Two

Abundant transcripts of malting barley identified by serial analysis of gene expression (SAGE)

2.1 Summary 46

2.2 Introduction 47

2.3 Results and Discussion 49

2.3.1 Annotation of SAGE tags 52

2.3.2 Alternative mRNA transcripts 64

2.4 Conclusion 67

2.5 Experimental procedures 68

2.5.1 Tissue Collection and RNA Extraction 68

2.5.2 SAGE Library Construction and Sequencing 69

2.5.3 SAGE Data Analysis 70

2.6 References 70

Chapter Three

Differential gene expression analysis in a barley seed germination time course and confirmation with relative real-time RT-PCR

3.1 Abstract 77

3.2 Introduction 78

3.3 Materials and Methods 80

3.3.1 Tissue collection and LongSAGE library construction 80

3.3.2 LongSAGE data analysis 81

3.3.3 RNA extraction, cDNA synthesis and real-time RT-PCR analysis 81

3.4 Results and Discussion 84

3.4.1 SAGE data analysis 84

3.4.2 Real-time RT-PCR analysis 90

3.4.3 Comparison of SAGE and real-time RT-PCR data 97

3.5 References 103

Chapter Four

Transcriptome analysis of malting barley; comparison of Microarray, LongSAGE and real-time RT-PCR

4.1 Abstract 108

4.2 Introduction 109

4.3 Materials and Methods 112

4.3.1 Array design 112

4.3.2 Target preparation, array hybridisation, scanning and stripping 113

4.3.3 Quality control, background correction and data normalisation 114

4.4 Results and Discussion 116

4.4.1 Microarray analysis of malting barley 116

4.4.2 Probe identification 118

4.4.3 Microarray results compared to LongSAGE and real-time RT-PCR results 126

4.5 Acknowledgements 133

4.6 References 133

Chapter Five

General Conclusions

5.1 Conclusions and summary 137

5.2 References 142

Appendices

Appendix One 144

Acceptance letter of publication of "Differential gene expression analysis in a barley seed germination time course and confirmation with relative real-time RT-PCR" in the journal Plant Science

Appendix Two 145

Copyright permission letters

List of Tables

Table		Page
Table 2.1	Summary of SAGE tag abundance	**51**
Table 2.2	Differential expression of SAGE tags between adjacent time points in a malting time course	**52**
Table 2.3	Summary of BLAST results for SAGE tag annotation	**54**
Table 2.4	The abundance, sequence and putative function of the 20 most abundant SAGE tags in each library	**54**
Table 2.5	Putative anti-sense LongSAGE tags from the 100 most abundant tags in each of the eight malting barley libraries	**67**
Table 3.1	PCR primers and product sizes for genes analysed by real-time RT-PCR	**83**
Table 3.2	LongSAGE tags found to be significantly differentially expressed between adjacent malting time-points using both Fishers exact test and a t-test at an overall 0.05 level	**89-90**
Table 3.3	Correlation of relative expression level obtained from LongSAGE and real-time RT-PCR for eight barley transcripts at four time-points during germination	**98**
Table 4.1	Signal range observed in normalised array data at the probe level and LongSAGE tag count data	**118**
Table 4.2	Ten probes with the highest signal intensity in malting barley at 12 hours post-steeping	**119**

Table 4.3	Ten probes with the highest signal intensity in malting barley at 48 hours post-steeping	**119**
Table 4.4	Probes with signal down-regulated more than 10 fold from 12 to 48 hours post-steeping	**123**
Table 4.5	Probes with signal up-regulated more than 10 fold from 12 to 48 hours post-steeping	**124**
Table 4.6	Malting barley microarray signal intensities of genes of interest and comparison to gene expression values of the same genes of interest using LongSAGE and Real-time RT-PCR	**127**

List of Figures

Figure		Page
Figure 1.1	Barley production in Australia	**2**
Figure 1.2	Morphological seed differences of wild and cultivated barley	**4**
Figure 1.3	Schematic of inflorescence of *Hordeum vulgare*	**5**
Figure 1.4	Electrochemical detection scheme of CombiMatrix ElectraSense™ microarray	**15**
Figure 1.5	Serial Analysis of Gene Expression	**18**
Figure 2.1	Whole grain images of germinating barley grains	**50**
Figure 2.2	The combined putative functional roles represented by the 100 most abundant tags from eight barley LongSAGE libraries	**58**
Figure 2.3	Profile of expression level of specific transcripts present in the top 100 tags	**59-60**
Figure 2.4	Alternative polyadenylation sites create multiple SAGE tags	**65**
Figure 3.1	A heat map showing experimental treatments along the horizontal axis, and hierarchical clustering of LongSAGE tags along the vertical axis	**85**
Figure 3.2	K-mean clusters of temporal expression profiles generated using GeneSpring	**86-87**
Figure 3.3	Relative level of transcript abundance of genes of interest compared to internal control genes obtained from real-time RT-PCR	**92-93**

Figure 3.4	Relative expression level of LongSAGE and Real-time RT-PCR analysis	**99-101**
Figure 4.1	Representative ElectraSense™ image of slide one hybridised with RNA from 12 hours post-steeping	**115**
Figure 4.2	Reproducibility of CombiMatrix ElectraSense™ Custom microarrays	**116-117**
Figure 4.3	Comparison of the normalized signal intensities obtained from barley CombiMatrix ElectraSense™ 12K Custom microarrays	**121**

Chapter One

Literature Review and General Introduction – Analysis of Gene Expression in Germinating Barley

1.1 Barley

1.1.1 The Importance of Barley

Cultivated barley (*Hordeum vulgare* L.) is an important grain crop, the world's fourth major cereal crop after maize, wheat and rice (Langridge and Barr, 2003). Barley provides a major contribution to the world food supply in the areas of human food, stock feed and malt production, while also playing an important role as an experimental and model plant species for study (Nilan and Ullrich, 1993). Barley malt is used to add flavour to a variety of different foods; however the most extensive use of barley malt is for brewing (Bamforth and Barclay, 1993).

Breeding strategies and farming practices in the last few decades have resulted in profitable yield increases in barley production (Langridge and Barr, 2003). In Australia, barley production and area have both increased over time. Between the 1996/97 and 2003/04 winter crop seasons, Australian barley production increased 1.3 fold although the sown area was only 1.1 times greater (Figure 1.1) (ABARE, 2003). For Australia the growing seasons for 2002/03, 2004/05 and 2006/07 resulted in poor production due to extremely dry conditions in the main growing states. The forecast for 2007/08, while higher than 2006/07, is still relatively low as most cropping regions of Australia have recorded below average winter rainfall. These extremely dry

conditions have been confounded by above average daytime temperatures which have placed additional pressure on the moisture stressed crops (Brown *et al.*, 2007). It is becoming increasingly obvious that research in breeding and farming practices needs to focus on barley production in dryer, hotter conditions if improvements in yield over area are to continue.

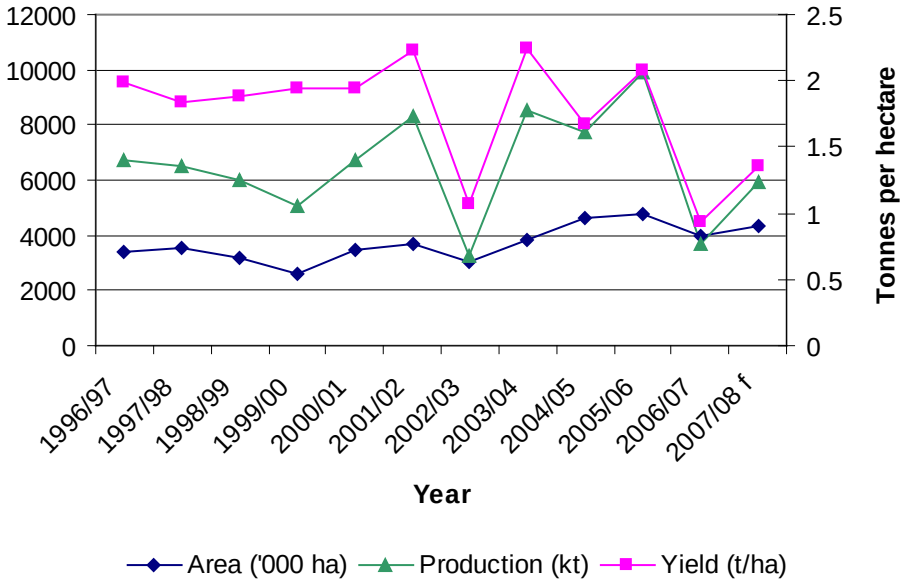


Figure 1.1 Barley production in Australia. Based on data from ABARE Australian Crop Reports (Brown *et al.*, 2007; Hogan *et al.*, 2004) and Australian Commodity Statistics (ABARE, 2003; ABARE, 2007). f ABARE forecast.

The amount of barley used for domestic malting and malting barley exports have also increased over time. Australia is a large exporter of barley malt, according to the United Nations Statistics Division, Australia was the 4th largest exporter of barley in 2008 (UNdata, 2009). The five year average to 2006/07 for total Australian barley production used for export was approximately 66% (Brown *et al*., 2007). The average price per tonne of Australian barley is consistently increasing for each of domestic

barley, export feed barley and export malt barley. The increases in the value of Australian barley and barley malt are likely to be a result of increased crop quality, increased likelihood of the harvested grain achieving malt quality, and improved barley varieties and farming practices in the last few decades (Langridge and Barr, 2003).

1.1.2 Barley Taxonomy and Origin

Barley is a member of the grass family, the Poaceae, the barley genus Hordeum belongs to the Triticeae tribe (Nilan and Ullrich, 1993). Archaeological remains of barley grains found in the Israel-Jordan region of the Fertile Crescent indicate that barley was domesticated there about 10 000 years ago when the cultivated form, *Hordeum vulgare*, is believed to have split from its wild relative *Hordeum spontaneum* (Badr *et al*., 2000). Modern nomenclature recognises the two forms as belonging to different subspecies of *Hordeum vulgare* (H.v.) (Nilan and Ullrich, 1993). They are closely related and are inter-fertile, though *H.v.vulgare* can be both six- and two-rowed with non-brittle rachis, whereas the wild *H.v.spontaneum* is two-rowed with a brittle rachis (Nilan and Ullrich, 1993). Morphological seed differences such as groove depth and endosperm width are some of the characters used to discriminate wild from cultivated barley lines (Figure 1.2) (Badr *et al.*, 2000). One of 31 Hordeum species, cultivated barley is a diploid with 2n=14 chromosomes. Diploid, tetraploid and hexaploid varieties exist amongst the wild Hordeum species (Nilan and Ullrich, 1993).

Figure removed due to copyright restrictions

Figure 1.2 Morphological seed differences of wild and cultivated barley

Reproduced from (Badr *et al.,* 2000)

1.1.3 Morphology of barley inflorescence

The inflorescence of *H.v.vulgare* (Figure 1.3), which as the focus of this review will herein be referred to as barley, is a spike with an axis known as a rachis (Duffus and Cochrane, 1993). The spike has three spikelets alternating on opposite sides of each node of the flat rachis of the head. There is one central and two lateral spikelets (Nilan and Ullrich, 1993) each with a ventral glume or palea, and a dorsal glume, or lemma, which make up the husk, or hull (Duffus and Cochrane, 1993). When all three spikelets are fertile, it is referred to as a six-rowed spike. If only the central spikelet is fertile, the spike is known as two-rowed (Nilan and Ullrich, 1993). Each spikelet is a single floret composed of a lemma, palea, a single ovary and style, and three stamens and anthers. Cultivated barley is effectively a self-pollinator with an extremely low occurrence of cross pollination (Nilan and Ullrich, 1993).

Figure removed due to copyright restrictions

Figure 1.3 Schematic of inflorescence of *Hordeum vulgare*

USDA-NRCS PLANTS Database / Hitchcock, A.S. (rev. A. Chase). 1950. Manual of the grasses of the United States. USDA Misc. Publ. No. 200. Washington, DC. (USDA, 2004)

1.2 Seed Germination

Germination begins with uptake of water by the dry seed, and ends with the elongation of the radicle, or embryonic axis (Bewley, 1997). Most of the water

penetrates the seed near the embryo and spreads at different rates through various regions of the seed (Fincher and Stone, 1993). Initially imbibition is the dominant process of water uptake, as it is not dependent on metabolic activity, however once the protoplasm is hydrated and metabolic activity resumes, osmotic forces become more important in controlling the water content (Street and Opik, 1970). Substrates and enzymes for some metabolic processes appear to be present within the dry seed (Bewley, 1997; Street and Opik, 1970), and respiratory activity can be detected within minutes of imbibition (Bewley, 1997). The synthesis of ribosomal proteins under conditions of transcription inhibition suggests that pre-formed transcripts are stored in the embryonic axis of the dry seed (Beltran-Pena *et al.*, 1995). These ready to use mRNA enable the synthesis of proteins and assist in the production of new mRNAs during germination (Bewley, 1997).

1.3 Barley Malting

1.3.1 Malting

Malting is the industrial process applied to cereal grains to generate malt. Malt is used to add flavour, colour, sweetness and even nutrition to many foods. The most common use for malt is to provide a source of fermentable sugars to produce beer and whiskey through alcoholic fermentation (Bamforth and Barclay, 1993). Barley varieties differ widely in their potential as malting barleys. The production of malt for whichever purpose is dependent on the viability of the barley grain. Suitable malt barley will have low dormancy, low water sensitivity and high vigour (Bamforth and Barclay, 1993). Husked barleys only are used for malting (Briggs, 1998) with the

husks providing an efficient filter bed in the brewing process, through which wort is separated from residual solid particles after mashing (Bamforth and Barclay, 1993; Kasha *et al.*, 1993).

Irrespective of the intended use of malt, malting processes are similar. The malting process begins with steeping the barley in water to achieve a sufficient moisture level to encourage germination (Bamforth and Barclay, 1993). Steeping, or imbibition, activates metabolism in the aleurone and embryonic tissues which synthesise hydrolytic enzymes that catalyse the degradation of storage proteins in the starchy endosperm (Fincher and Stone, 1993). Once the starch granules are exposed and the grain is evenly modified or germinated, the grain is ready for kilning. Kilning arrests germination and stabilizes the malt by lowering the moisture levels ready for milling (Bamforth and Barclay, 1993).

1.3.2 Steeping

Steeping, the soaking of barley grains in water, is performed to increase the moisture content of the grain, to ensure rapid and uniform germination while minimising embryo growth and respiration (Bamforth and Barclay, 1993). The moisture content must be taken from 11-12% up to 42-46% in 48-52 hours (Bamforth and Barclay, 1993). To induce germination requires the thorough wetting of the grain, interrupted by air rests that provide the grain access to oxygen for respiration and helps to overcome the water sensitivity of some grain (Bamforth and Barclay, 1993). Grain temperatures are controlled during air rests, as respiration produces a considerable output of heat energy (Briggs, 1998).

During imbibition, the grain swells, increasing its volume by approximately one quarter, the width and depth of the grain increase but the length does not (Briggs, 1998). The acrospire, or sprout, and rootlet begin to grow, and these are traditionally used as a guide to the malting progress. The growing acrospire presses a visible groove into the outer tissues and its length is noted as a fraction of the grain's (Briggs, 1998). When the average length of the acrospire is approximately ¾ the length of the grain, and the endosperm is easily rubbed out between the finger and thumb, the malt is ready to be kilned (Bamforth and Barclay, 1993).

When imbibition begins, the embryo and husk absorb water more rapidly than the endosperm, and it has been thought that the embryo regulates water uptake (Bamforth and Barclay, 1993). More recently it has been suggested that structural differences in the starchy endosperm also affect water uptake during germination (Chandra *et al*., 1999). Structural differences in barley grains have been described as mealy or steely. Mealy endosperms contain air spaces between their loosely packed cells and starch granules, and appear to display faster distribution of water during steeping, whereas steely endosperms have a more densely packed protein matrix and smaller starch granules (Chandra *et al*., 1999).

1.3.3 Germination

During the malting process germination is targeted to maximise the production of extractable material for malt, through the process of endosperm modification. Modification describes the physical and chemical changes that occur during malting,

including the degradation of the starchy endosperm and the associated biochemical changes (Briggs, 1998). The rate of modification depends on several factors; the rate of moisture distribution through the endosperm, the rate of enzyme synthesis, the extent of enzyme release and the structure of the endosperm which will determine its digestibility (Bamforth and Barclay, 1993). Hydrolytic enzymes, initially released from the scutellum and later the aleurone layer, degrade cell walls, allowing the matrix proteins to be broken down, freeing the starch granules and exposing them to degradation (Briggs, 1998).

The hormone activity of gibberellic acid (GA) and abscisic acid (ABA) plays a large part in endogenous regulation of germination by inducing and repressing the synthesis and secretion of hydrolytic enzymes respectively (Fincher and Stone, 1993). GA is released from the scutellum and embryo after imbibition; it diffuses through the endosperm to the aleurone layer where it stimulates the production of hydrolytic enzymes, including α-amylase and β-glucanase (Briggs, 1998; Fincher and Stone, 1993; Russell, 1998). α-Amylase acts to break down starch to generate soluble sugars and is the most extensively studied aleurone enzyme (Gibson, 2004; Jones and Jacobsen, 1991). Several α-amylase isoenzymes have been identified and divided into two groups based on their isoelectric points (Fincher, 1989; MacGregor *et al.*, 1988). The group of enzymes with low-pI values degrade starch more rapidly than the high-pI group (Fincher, 1989).

The cell walls within the barley endosperm contain a great quantity of the polysaccharide β-glucan (Nuutila *et al.*, 2002). Cell wall (1-3,1-4)-β-glucan is broken down by (1-3,1-4)-β-glucanases during germination (Fincher, 1989). There are two

closely related isoenzymes of (1-3,1-4)-β-glucanase in barley, EI and EII (Slakeski and Fincher, 1992; Woodward and Fincher, 1982). Beta-glucan concentration is relatively low compared to starch, however it can cause difficulties with malting and brewing, and affect malt quality (Nuutila *et al*., 2002; Wang *et al*., 2004). The rate of modification is controlled by water uptake, the production of hydrolytic enzymes and by the content of cell wall β-glucan and more significantly, the ability to synthesise β-glucanase (Bamforth and Barclay, 1993). The native barley β-glucanase is susceptible to inactivation at temperatures above 55°C (Nuutila *et al*., 2002). The development of a genetically modified, thermotolerant β-glucanase expressed during germination, has increased degradation activity during germination and mashing, which in turn decreases β-glucan concentration in wort and improves the filtration of beer (Nuutila *et al*., 2002; Wang *et al*., 1998).

During malting germination is controlled by maintaining moisture, oxygen supply, carbon dioxide removal and elimination of excess heat. Germination normally occurs between 16 and 20ºC (Bamforth and Barclay, 1993). Higher temperatures can increase growth rate, however this cannot be sustained over time as the rate of enzyme formation decreases. Malting time is limited; therefore the grain must germinate quickly and evenly, in the dark, and in modern malteries is expected to reach its maximum germination within three days (Briggs, 1998).

In higher plants, germination is usually preceded by a period of dormancy (Kuhlemeier *et al*., 1987). Seed dormancy in grain crops is not usually considered an advantageous characteristic. Bewley (1997) and Bove *et al*. (2001) review several studies that identify proteins, mRNAs and hormones involved with dormancy and

seed germination. The mechanisms by which the embryo emerges from the seed after germination and how this process is blocked in a dormant seed remain to be fully understood (Bewley, 1997; Bove *et al*., 2001).

1.3.4 Kilning

When the grain is completely and evenly modified it should be friable and thus ready for kilning (Bamforth and Barclay, 1993). Kilning gradually dries the grain, reducing the high moisture levels required for germination down to less than 5%. This must be regulated carefully to ensure the survival of enzymes that are crucial to brewing and distilling processes (Bamforth and Barclay, 1993). Some enzymes are more heat tolerant at lower moisture contents so drying begins at a relatively low temperature and increases progressively affecting flavour and colour changes over an approximate 24 hour kilning period (Bamforth and Barclay, 1993). On completion of kilning the malt grain is milled, ready to begin the brewing process.

1.4 Gene Expression

The germination rate of barley is a major contributor to malt quality. The mechanisms that drive biological events and developmental processes, such as germination, can be revealed by monitoring the pattern of gene expression under various physiological and pathological conditions (Schnable *et al*., 2004; Snustad and Simmons, 2000). Gene expression can be controlled at several levels in eukaryotes including transcription, RNA processing, RNA transport, translation, RNA degradation and protein degradation (Russell, 1998). The ability of a cell to control these processes allows the

production of a variety of diverse proteins in different cells at separate times (Kuhlemeier *et al*., 1987; Russell, 1998). Molecular biology techniques including RNA extraction, reverse transcription and northern blot analysis, have allowed us to examine the expression of single genes in different tissues or at different developmental time points. These types of analyses are beneficial; however they can be difficult and allow characterisation of only one gene at a time. Modern molecular biology techniques allow researchers to analyse the entire transcriptome of a cell or tissue by characterising differential expression of mRNAs (Donson *et al*., 2002; Dunwell *et al*., 2001).

1.5 Molecular Biology Techniques

1.5.1 Genomics

The genomes of eukaryotes are generally much larger than prokaryotes, however in most eukaryote genomes only a small proportion of the DNA encodes amino acid sequences. To focus on the protein-coding content of a genome, it is common to only analyse expressed genes or DNA sequences that are transcribed into messenger RNA (mRNA) copies (Snustad and Simmons, 2000). A specialized enzyme, reverse transcriptase, is used to catalyse the synthesis of DNA molecules complementary to mRNA templates called cDNAs (complementary DNAs) (Snustad and Simmons, 2000). Molecular biologists create cDNA libraries which contain the coding regions of the expressed genes. To construct a library, cDNA sequences are inserted into cloning vectors to form recombinant DNA molecules (Snustad and Simmons, 2000). These are introduced into host bacterial cells which are grown into colonies that will

contain just one cDNA insert that can be isolated and sequenced. Expressed genes can be identified from their DNA sequence. Single pass sequencing of cDNA clones produces Expressed Sequence Tags (ESTs) that are used to characterise the genes expressed in a tissue (Snustad and Simmons, 2000).

The EST approach to functional genomics can be expensive and time consuming and may miss genes expressed at low levels or in only a few specialized cells unless a normalization step is carried out and the library is extensively sequenced. A method to identify patterns of gene expression across different tissues, throughout development, or in response to external stimuli, is northern blot hybridization (Snustad and Simmons, 2000). In this method, RNAs from specific starting material are separated by gel electrophoresis and transferred to membranes for hybridization with a labelled probe. The specific sequence probe will bind with complementary transcripts on the membrane and display a level of expression for that transcript. The presence of certain gene transcripts, in different cells at various or specific times, can indicate functions and interactions of genes during development (Schnable *et al*., 2004).

1.5.2 Microarray

Northern blot hybridisation is designed to analyse the expression of one transcript at a time. To analyse the expression of large sets of genes, dot blot hybridisation or microarray technology was developed (Schena *et al*., 1995; Snustad and Simmons, 2000). A microarray is a reverse northern, in which gene specific nucleotide sequences or probes are bound in a specific array pattern to a solid surface, and these are hybridised with fluorescent cDNA or cRNA preparations (Snustad and Simmons,

2000). The amount of cDNA that hybridises to each probe/dot can be measured by scanning the array with an imaging system and analysing the results with computer programs (Snustad and Simmons, 2000). The microarray can monitor many thousands of genes by using high speed robotic printing onto microscope slides (Schena *et al*., 1995). Microarrays contain replicated features to test the reproducibility of the process (Schena *et al.*, 1995) and allow differential gene expression to be measured by simultaneous two-colour hybridisation (Aharoni and Vorst, 2002; Schena *et al*., 1995). Two mRNA sources are prepared by reverse transcription to create two cDNA probes with different fluorophores (Schena *et al*., 1995). Samples of known concentration are used to calibrate the fluorescence of the probes so that expression can be quantified and compared (Schena *et al*., 1995). Different types of microarray platforms are available for different applications. The major systems include two-colour or one-colour platforms designed to detect the hybridisation of probes to oligonucleotide or cDNA features.

A recent improvement to the development of microarrays is the ElectraSense™ electrochemical detection system developed by CombiMatrix (Ghindilis *et al*., 2007; Roth *et al*., 2006). Microarrays have previously been manufactured by spotting pre-synthesized molecules, using inkjet or light-directed synthesis processes and the signal intensities of subsequently hybridised molecules measured with fluorescent detection. CombiMatrix uses computer software to guide the simultaneous synthesis of different oligonucleotides at thousands of platinum microelectrodes using a modified semiconductor (Ghindilis *et al*., 2007). This electronic system can subsequently be used for electrochemical detection (ElectraSense™) of target molecules. ElectraSense™ technology is based on detection of a redox reaction

proximal to the specific electrode. Target molecules are labelled with biotin, hybridised to the array, the biotin is bound to a horseradish peroxidase (HRP) streptavidin marker and electrodes are exposed to tetramethylbenzidine (TMB) and hydrogen peroxide (Ghindilis *et al*., 2007; Roth *et al*., 2006). The oxidized TMB is reduced at the electrode surface which generates an electrochemical signal (Figure 1.4) (Ghindilis *et al*., 2007; Roth *et al*., 2006).

Figure removed due to copyright restrictions

Figure 1.4 Electrochemical detection scheme of CombiMatrix ElectraSense™ microarray. Figure taken from (Ghindilis *et al*., 2007).

Microarray analysis is a valuable tool for plant functional genomics and has been thoroughly reviewed as an effective technique for identification and characterisation of differentially expressed genes (Aharoni and Vorst, 2002; Desprez *et al*., 1998; Donson *et al*., 2002; Dunwell *et al*., 2001; Schnable *et al*., 2004; Wullschleger, 2003; Zhu, 2003). Differentially expressed genes are identified by comparing two or more biological samples. A typical experiment will compare samples of different varieties, different times in development or treated versus untreated samples. It is common to keep almost all attributes of the experiment the same except for one feature of interest. Two samples can then be compared to determine the level of gene expression and

how it is affected in each sample. A differentially expressed gene would present high levels of expression in one sample and low levels of expression in the other. Microarray technology is an important tool for differential expression analysis however it is limited to the analysis of genes with known sequence. Serial Analysis of Gene Expression (SAGE) allows all transcripts in a sample to be profiled without prior knowledge of what the transcripts encode (Dunwell *et al*., 2001; Velculescu *et al.*, 1995).

1.6 SAGE

1.6.1 SAGE

The SAGE method developed by Velculescu *et al.* (1995) enables analysis of gene expression of a large number of transcripts in a quantitative and unbiased manner. SAGE is based on the notion that a short nucleotide sequence tag, isolated from a defined position within a transcript, contains enough information to uniquely identify it (Velculescu *et al*., 1995). Efficient analysis of transcripts is achieved by concatenation of the short sequence tags into a single clone to be sequenced (Velculescu *et al*., 1995). SAGE has an advantage over EST based methods, in being able to detect low-abundance transcripts (Sun *et al*., 2004) as it overcomes the obstacle of redundancy present in EST based methods, making it more likely to uncover novel transcripts (Chen *et al*., 2002b).

In SAGE, as in cDNA library construction, double-stranded cDNA is synthesised from mRNA, using OligodT primers, from the tissue/material of interest. In SAGE the

cDNA is cut with a 4bp anchoring restriction enzyme NlaIII, which would be expected to cleave most transcripts at least once (Figure 1.5) (Velculescu *et al*., 1995). The cDNA is divided into two pools and each ligated to a different adapter containing a type IIS restriction enzyme priming site for MmeI. Type IIS restriction enzymes (tagging enzymes) cleave at a defined distance downstream of the recognition site (Tucholski *et al*., 1995). This digestion leaves the adapter attached to a small sequence tag of the original transcript. The two pools are combined at this stage and ligated to form ditags (Velculescu *et al*., 1995). Production of uniform ditags allows amplification of the sample without introducing PCR bias. Cleavage of the PCR product with the original anchoring enzyme allows the ditags to be isolated and adapter molecules removed (Velculescu *et al.*, 1995). The resulting ditags are ligated to form concatemers to be cloned into the vector pZErO-1 and sequenced. During sequence analysis the use of the anchoring enzyme site for ligation acts as a punctuation boundary for each tag (Velculescu *et al*., 1995).

Since it was first described, several modifications to the SAGE method have been suggested to improve yield and transcript identification in SAGE libraries (Powell, 1998; Velculescu *et al*., 2000; Zhang and Gilles, 2003). A major theme among these modifications has been to improve the accuracy of tag identification. A different restriction endonuclease has been used to produce longer tags of 21bp in LongSAGE (Saha *et al*., 2002; Zhang and Gilles, 2003) and 26bp in the SuperSAGE methods (Matsumura *et al*., 2003). The most recent and comprehensive review of the SAGE method describes Robust-LongSAGE (RL-SAGE), in which major improvements to the protocol overcome significant technical difficulties including low cloning efficiency and small insert sizes (Gowda *et al*., 2004).

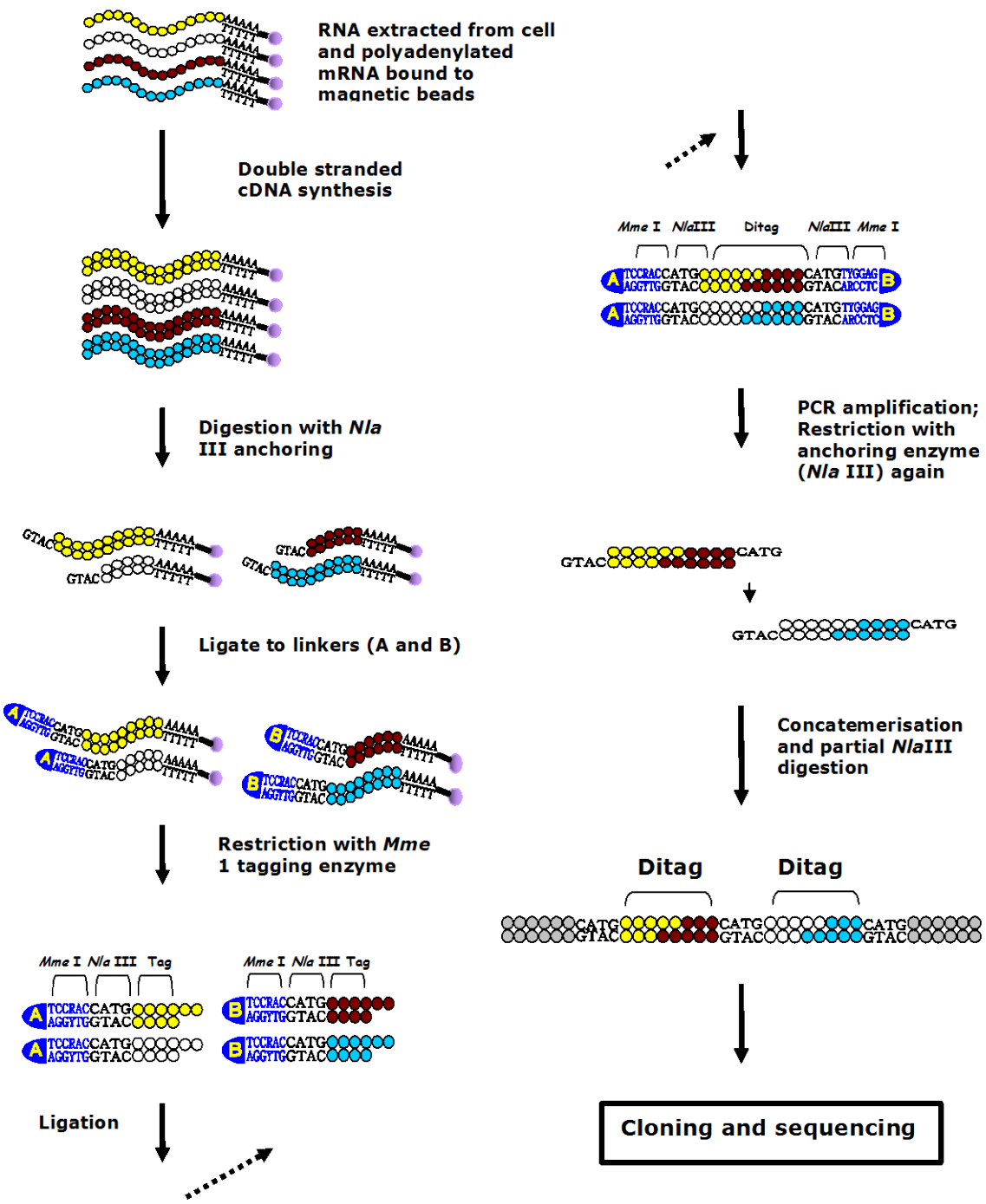


Figure 1.5 Serial Analysis of Gene Expression

Modified protocols have been described to reduce the amount of mRNA required for cDNA synthesis (Velculescu *et al*., 2000). To produce sufficient ditags for library construction, these modifications have relied on additional PCR amplifications, which

may introduce bias in gene expression patterns (Gowda *et al*., 2004). In addition, a large quantity of mRNA can lead to incomplete digestion of cDNA which may generate multiple tags from the same transcript making it difficult to calculate gene expression from tag abundance (Gowda *et al*., 2004). Modifications present in the RL-SAGE protocol allow efficient libraries to be constructed with limited starting material, as little as 50ng compared with 2-5µg used in conventional and LongSAGE protocols, by extending the incubation times for cDNA digestion, cDNA adapter ligation and ditag ligation (Gowda *et al*., 2004). Improvements in ditag amplification and gel purification, including steps to decrease linker contamination as recommended by Powell (1998), have increased the cloning efficiency of subsequent steps. Partial digestion of concatemers, suspected to circularise during ligation, significantly improves the efficiency of cloning and increases insert sizes of clones (Gowda *et al*., 2004). Critical improvement of SAGE has reduced the time required for library construction by avoiding colony PCR screening to remove empty clones (Gowda *et al*., 2004). Crawford *et al*. (2005) provide a method for ensuring consistent production of high-molecular weight SAGE concatemers. They demonstrated that the concentration of PCR amplified ditag molecules present in concatemerisation and partial digestion reactions affects the proportion of low-molecular weight contaminants present in the resulting library (Crawford *et al*., 2005).

SuperSAGE, which uses a type III restriction endonuclease EcoP15l to isolate 26bp tags, has been used to compare two interacting organisms in a host-pathogen relationship by unambiguous gene-to-tag assignment (Matsumura *et al*., 2003). Unlike previous pathogen transcript profiles that rely on the use of artificial membranes to prepare suitably uncontaminated experimental material (Thomas *et al*., 2002),

SuperSAGE will be useful for understanding the real molecular mechanisms of both interacting species especially with the advantage of completely sequenced host and pathogen genomes (Matsumura *et al*., 2003).

SAGE studies have produced vast numbers of tags that cannot be matched to known genes, especially when there is a lack of genomic sequence data for the organism of interest (Dean and Lorenz, 2004; Saha *et al*., 2002). RT-PCR techniques have been developed to further analyse unknown tags (Chen *et al*., 2002a; Chen *et al*., 2000; Lee *et al*., 2002; Matsumura *et al*., 2003; van den Berg *et al*., 1999). The GLGI technique (Generation of Longer complementary DNA [cDNA] fragments from SAGE tags for Gene Identification), uses a SAGE tag sequence as the sense primer, and an anchored oligo(dT) primer as the antisense primer to amplify the original cDNA template from which the SAGE tag was derived (Chen *et al*., 2002a; Chen *et al*., 2000; Lee *et al*., 2002). Unlike the short tags produced by conventional SAGE which prove inadequate in efforts to clone unknown genes (Lorenz and Dean, 2002), the longer tags from SuperSAGE have been established as useful primers for the 3'RACE technique (Rapid Amplification of cDNA Ends) to produce longer cDNA fragments from unknown tags (Matsumura *et al.*, 2003).

1.6.2 SAGE in plants

SAGE was first described in plants by Matsumura *et al*. (1999). In this study of rice seedlings the 10 122 tags analysed included 5 921 unique tag sequences expressed in rice seedlings (*Oryza sativa* L.). While only 23.1% of these tags matched to rice cDNA or EST sequence databases, most of the highly expressed tags could be

identified, and matched primarily to housekeeping genes. This quantitative study of gene expression in rice, further applied the SAGE method to identify genes that were differentially expressed between anaerobically treated and untreated seedlings. Twenty-four genes, including eighteen anaerobically induced genes and six that were repressed, showed expression at significantly different levels in each tag pool (Matsumura *et al*., 1999).

Lorenz and Dean (2002) produced the first large scale SAGE libraries in a higher plant. They produced a combined data set of more than 150 000 tags with approximately 42 000 different tag sequences using libraries constructed from lignifying xylem in the crown and base portions of loblolly pine (*Pinus taeda*). In this study, it was decided that the omission of singleton tags would not affect the analysis of differentially expressed tags between the SAGE data sets. The possibility that a tag seen only once could represent a PCR artefact or sequencing error makes it difficult to determine true unique genes (Lorenz and Dean, 2002). Zhang *et al.* (1997) and Velculescu *et al.* (1997) however, estimate the sequence error rate for SAGE tag concatemers as approximately 0.7% or a tag error of 6.8% which would equate to at least 93% of unique tags representing true transcripts.

The SAGE method was used to analyse gene expression patterns in rice leaf and seed, and a significant number of tags coding for antisense transcripts were found (Gibbings *et al*., 2003). SAGE technology allows the position and orientation of a tag to be determined with respect to an mRNA molecule, thus allowing the detection of antisense transcription (Gibbings *et al*., 2003; Patankar *et al*., 2001). In Gibbings *et al*. (2003), two of the 100 most abundant tags were derived from the non-coding strand in

the 5'direction, and antisense tags were even more frequent at lower abundances. Many of the tags of low abundance were found to be derived from antisense transcripts of more abundant genes, indicating that the gene number generated by this SAGE analysis is an overestimation. However, there was no association between the frequencies of sense and corresponding antisense tags, indicating that the presence of antisense tags is not an artifact of the SAGE method (Gibbings *et al*., 2003). Antisense transcription in plants is more widespread than previously thought and may represent an important mechanism for gene regulation (Gibbings *et al*., 2003; Terryn and Rouze, 2000). Antisense transcripts have a role in RNA silencing. Small antisense RNAs were first discovered in plants as a mechanism of defence against invading nucleic acids such as transgenes and viruses; however endogenous silencing pathways and microRNA molecules have been found to have widespread roles including gene regulation at transcriptional, RNA stability and translational levels (Brodersen *et al.*, 2006; Bartel, 2004).

SAGE has been used to analyse the biological response of the transcriptome to abiotic stimuli (Jung *et al*., 2003). Comparison of the expression level of tags identified from cold-treated Arabidopsis leaves to normal leaves identified 272 differentially expressed genes (Jung *et al*., 2003). The total number of expressed genes in this study was notably reduced after cold stress, but the average copy number per gene was increased (Jung *et al*., 2003). Of the genes showing more than a ten fold difference in expression, 87 were newly identified as cold-stress related genes (Jung *et al*., 2003). To confirm the validity of the SAGE data, five putative cold-responsive genes were further analysed by northern blots. The expression data obtained by northern blot analysis were consistent with, and thus confirmed, the differential expression patterns

derived from SAGE (Jung *et al*., 2003). Several other SAGE studies have used northern blot analyses to validate the expression data produced by SAGE (Patankar *et al*., 2001; Trendelenburg *et al*., 2002).

The model plant Arabidopsis has been analysed in many SAGE studies (Fizames *et al.*, 2004; Jung *et al*., 2003; Lee and Lee, 2003). Lee and Lee (2003) were able to reveal the groups of proteins and metabolic pathways likely to be active during Arabidopsis pollen development. A SAGE library was also constructed from cold stressed pollen of Arabidopsis (Lee and Lee, 2003). The expression of known cold-inducible genes, all of which are highly induced in the Arabidopsis leaf under cold stress (Jung *et al*., 2003), showed no significant increases in the pollen library (Lee and Lee, 2003). Microarray based analysis was used to support this observation, and the study suggests that the inability to induce expression of genes important for cold acclimation is linked to the cold sensitivity of pollen (Lee and Lee, 2003).

Fizames *et al.* (2004) conducted a large investigation of the Arabidopsis root transcriptome by SAGE. Taking advantage of the complete sequence of the genome, the study used a computer-based approach to characterise expression of more than 3 000 genes that did not exist previously in EST or cDNA databases (Fizames *et al*., 2004). EST clusters or cDNA sequences had previously been used to identify the gene corresponding to the SAGE tag transcript (Gibbings *et al*., 2003; Matsumura *et al*., 1999). This has been a major limitation of SAGE tag to gene assignment as not all transcripts exist in cDNA and EST databases and many tags are left unmatched and unidentified (Fizames *et al*., 2004; Lash *et al*., 2000; Pleasance *et al*., 2003). Despite having the full genome sequence a large proportion of tags in the Arabidopsis root

transcriptome study could not be assigned to any gene (Fizames *et al*., 2004). The unassigned tags could be artifactual, however erroneous tags are discarded by the exclusion of single tags and duplicate ditags, which suggests that unassigned tags originate from transcripts with incorrect or absent virtual tags in the computer generated tag reference library (Fizames *et al*., 2004).

SAGE has been employed to investigate barley in whole seed and through a 120 hour malting time course (White *et al*., 2006) and in a direct comparison to the barley Affymetrix GeneChip (Ibrahim *et al*., 2005). The most abundant and most significantly up-regulated LongSAGE tag in the study by White *et al.* (2006), matched to two previously reported genes, B22EL8 and pZE40, which have been described as metallothionein-like proteins with proposed roles in synthesis and transport of metabolites (Klemsdal *et al*., 1991; Smith *et a*l., 1992). Ibrahim *et al.* (2005) also reported pZE40 as the most abundant SAGE tag identified in their study. They generated 82,122 tags from two libraries representing whole developing barley caryopsis and dissected embryos, at 12 days post-anthesis (Ibrahim *et al*., 2005). Difficulties associated with multiple tags matching to one transcript were reported. Sources of multiple tags per transcript evolve from alternative splicing, incomplete digestion with the anchoring enzyme during library construction, alternative polyadenylation sites and possibly non-specific poly dT anchoring of reverse transcription during cDNA synthesis (Ibrahim *et al*., 2005; Pauws *et al*., 2001; Welle *et al*., 1999).

1.7 Real-time reverse transcription PCR

Real-time PCR is an efficient method for determining DNA quantity as it combines DNA amplification and detections steps in one assay. This eliminates the need for post PCR processing steps such as gel electrophoresis, Southern blotting and DNA sequencing. Reverse transcription PCR (RT-PCR) involves the use of reverse transcriptase to synthesise complementary DNA (cDNA) to an mRNA template (Snustad and Simmons, 2000). Monitoring the amplification reaction in real time allows the accurate determination of template quantity in the starting material. Amplification is monitored using fluorescent molecules. There are effectively two types of fluorescent chemistry, template-specific fluorescent probes and non-specific DNA binding fluorophores (eg SYBR Green) (Bustin, 2005b). DNA binding fluorophores fluoresce when incorporated into double stranded DNA. The fluorescent signal produced by either chemistry is directly proportional to the amount of amplified product present. The linear relationship between the fluorescence output and sample concentration allows the amount of template present at the start of the reaction to be calculated (Bustin, 2005a). Real-time RT-PCR results are often referred to in terms of the cycle threshold (Ct). This refers to the cycle number at which a PCR product is first detected. Ct is determined in the exponential phase of the reaction and is inversely proportional to target copy number. Therefore the higher the starting copy number, the sooner the PCR product is detected and the lower the Ct value (Bustin, 2005a).

Real-time RT-PCR reactions are a simple, specific, sensitive and highly reproducible method for the detection of RNA. The most important determining factor in the reproducibility of RT-PCR results is the RNA template quality (Bustin, 2005b; Fleige

and Pfaffl, 2006). The introduction of lab-on-chip technologies such as the Bioanalyser allows the accurate assessment of RNA quality and quantity and provides a statistic for this assessment, the RNA Integrity Number (RIN) (Fleige and Pfaffl, 2006).

Real-time RT-PCR results can be obtained by relative or absolute quantification. Relative quantification represents the change in mRNA levels of a gene of interest relative to the levels of a control mRNA, for example, a housekeeping gene such as GAPDH (Bustin, 2005b). Target Ct values are compared to the internal control and expressed as a ratio (Bustin, 2005b). Absolute quantification is an attempt to determine absolute copy number in terms of amount of total RNA or weight of tissue in extraction based on a standard curve. A standard dilution series of known concentrations is used to generate a standard curve by plotting Ct versus known copy number (Bustin, 2005b). The copy numbers of unknown samples can be determined by the position of corresponding Ct values on the curve.

Real-time RT-PCR has been used extensively in the biomedical field and the utility of the technology in the development of diagnostic assays is rapidly increasing (Mackay *et al*., 2002; Valasek and Repa, 2005). Real-time RT-PCR has also proven to be a valuable research tool to measure gene expression levels for a wide variety of genes in various plant species; for example, Arabidopsis (Klok *et al*., 2002), canola (Yang *et al*., 2007), pea (Ayele *et al*., 2006), peanut (Luo *et al*., 2005) and barley (Burton *et al*., 2004; Dean *et al*., 2002; Liu *et al*., 2005; Potokina *et al*., 2006; Wolbang *et al*., 2004).

1.8 Gene Expression in Barley

In recent times, microarray technology has been exploited to examine gene expression in barley (Close *et al*., 2004; Potokina *et al*., 2002; Sreenivasulu *et al*., 2002; Sreenivasulu *et al*., 2004; Watson *et al*., 2001). Most notably a 22K Barley1 GeneChip probe array has become publicly available (Close *et al*., 2004). This cereal genomics resource, a model for plants without a fully sequenced genome, allows parallel expression analysis and gene discovery in disease defence, abiotic stress response, evolutionary diversity and development (Close *et al*., 2004). Microarray content was derived from a worldwide contribution of 350 000 ESTs from 84 cDNA libraries, together with 1 145 non-redundant gene sequences from NCBI (Close *et al*., 2004). The parameters used in designing the array allow a high degree of specificity to enable very similar genes, such as gene family members, to be differentiated (Close *et al*., 2004). Other features include the use of multiple probe pairs per sequence to reduce false positives and enable expression quantification, and the use of probes representing housekeeping or conserved genes to facilitate comparative studies among grasses (Close *et al*., 2004). A public database allowing access to the data makes possible the integration of expression profiling experiments from around the world, links annotations and alignments with data from other plant databases, and the option to query a gene of interest to discover the conditions under which it showed significant change (Close *et al*., 2004). The 22K Barley1 GeneChip is a useful tool for the analysis of barley and other cereal genomes. The GeneChip has been used to profile barley-stem rust interactions (Zhang *et al.*, 2008), maturation, desiccation and germination (Sreenivasilu *et al.*, 2008), genotyping (Walia *et al.*, 2007), Giberellin

and abscisic acid responses (Chen *et al.*, 2006), and most recently to analyse malting barley phenotypes (Lapitan *et al.*, 2009).

To investigate the regulation of gene expression during germination of barley grains, Potokina *et al*. (2002) developed a cDNA macroarray with 1 440 spotted clones derived from cDNA libraries of developing caryopsis, roots and etiolated leaves. It was their goal to analyse the spatial and temporal patterns of gene expression in embryo, scutellum and endosperm tissue, including the aleurone, during different stages of germination (4, 12, 36, 52 hours after the start of imbibition) (Potokina *et al*., 2002). The scutellum was considered to be a transitional tissue between embryo and endosperm from a spatial and functional point of view and was used as a comparator for each (Potokina *et al*., 2002). Genes expressed in the endosperm were clustered according to three distinct behaviours, up-regulated, down-regulated and constant expression. Significant down-regulation of genes appears to be specific to the endosperm, as it was not observed in the embryo or scutellum (Potokina *et al*., 2002). A considerable portion of genes are up-regulated in each of the three tissues during the late stages of germination (36 and 52 hours after imbibition), while a few genes identified only in the scutellum were up-regulated only in the final stage 52 hours after imbibition (Potokina *et al*., 2002). Most of the genes preferentially expressed in the scutellum correspond with expected functions of the tissue, including production of hydrolytic enzymes for storage protein and starch catabolism (Potokina *et al*., 2002). This study provides some information about the metabolic processes of germination. The array was derived mainly from a cDNA library of developing barley seeds however, and to perform a more detailed analysis of germination would require the incorporation of transcripts from germinating barley grains in the array design.

Sreenivasulu *et al*. (2002; 2004) has produced cDNA macroarrays to infer molecular physiology from maternal pericarp and filial embryo/endosperm tissues of developing barley grains. The gene expression profiles in developing barley caryopses were studied from fertilization to storage phase and it was discovered that expression in the pericarp mainly encodes enzymes involved in carbohydrate and lipid metabolism while expression in the embryo is related to degradation and processing of proteins or the accumulation of starch (Sreenivasulu *et al*., 2002). An intermediate phase separating the cell division and accumulation processes associated with grain development was identified, during which a considerable rise in gene expression related to photosynthesis and energy production was observed (Sreenivasulu *et al*., 2004). Like the array by Potokina *et al*. (2002), this study was very useful; however it provides no expression data specific to the germination of barley grains.

In a comparative study of SAGE and Affymetrix array technologies, Ibrahim *et al*. (2005) investigated gene expression in barley caryopsis and dissected embryos at 12 days post-anthesis. The Barley 1 Affymetrix GeneChip contains 22 791 probe sets, approximately half of which were given a present call in both tissues ($p \leq 0.05$) (Ibrahim *et al*., 2005). Similar trends were observed between the SAGE and GeneChip scores with only a small proportion of transcripts accounting for the majority of signal. The 100 most abundant transcripts determined by each method were identified and compared, and strong agreement was found between the SAGE and GeneChip technologies (Ibrahim *et al*., 2005).

Germination rate plays an important role in determining malt quality. Watson *et al*. (2004; 2001) used cDNA microarrays to examine gene expression in the embryos of germinating barley grain. The expression profiles of barley embryos (including scutellum) from 4, 8, 12, 24, 48, 72 and 96 hours post-imbibition were compared to a reference expression profile of the 24 hour post-imbibition time-point in a series of microarray experiments (Watson and Henry, 2004). Genes with similar gene expression patterns were separated into groups of potentially similar function. Only 514 of the total data set (4 945 clones) could be classified by distinctive patterns of expression (Watson and Henry, 2004). Of this subset, 16.9% of clones showed ten fold or greater differences in expression prior to 24 hours. More than 50% of the subset however, showed a ten fold or greater difference in expression in the period 24-96 hours post-imbibition and a high proportion of these were down-regulated (Watson and Henry, 2004). One of the few clones showing sustained up-regulation in the period after 24 hours post-imbibition was the 23kDa stress-response Jasmonate-Induced Protein (JIP-23). JIP-23 most likely plays a role in mediating osmotic stress in germinating barley embryos (Watson and Henry, 2004). This work provides a useful beginning for the analysis of gene expression in germinating barley embryos by determining temporal differences in gene expression patterns over the first four days post-imbibition. However, a comprehensive analysis of genetic components and mechanisms for developmental, regulatory and response processes in whole germinating barley has yet to be investigated in a full genome examination.

Malting is a highly controlled process that ensures the efficient release of fermentable sugars from the barley grain. The biochemical basis of the malting process has been studied in considerable depth. This knowledge has been used to optimise the quality

and quantity of output from malt barley. The genetic basis of the malting process however has not yet been examined in much detail. Modern molecular biology techniques enable the analysis of thousands of transcripts simultaneously. The high-throughput techniques of SAGE and Microarray will be applied to a time course of malting barley samples. This study aims to provide large data sets for the basis of accurate gene expression analysis with widespread coverage of the malting barley transcriptome.

1.9 Aims and Objectives

Malting is a highly controlled process that ensures the efficient release of fermentable sugars from the barley grain. The biochemical basis of the malting process has been studied in considerable depth. This knowledge has been used to optimise the quality and quantity of output from malt barley. However the genetic basis of the malting process has not been studied in detail. Modern molecular biology techniques enable the analysis of thousands of gene transcripts simultaneously, providing an opportunity to improve knowledge of the malting process at the gene level. This study aims to determine what genes are expressed during the malting of barley and the temporal patterns in which they are expressed and to explore different gene expression analysis technologies.

The high-throughput techniques of SAGE and Microarray will be applied to a time course of malting barley samples to investigate the genes that control and regulate this process. This time course includes samples of dry dormant seeds, post-steeped grain, and grain germinating at time points up to 120 hours post-steeping when the shoots

and roots are emerging from the grain and the endosperm is completely modified. This study aims to provide primary resources of transcription data in the form of a large malting barley LongSAGE tag database and a CombiMatrix 12K Custom ElectraSense™ barley microarray chip.

Chapter one of this thesis is a review of the relevant literature on malting and gene expression analysis. Chapter two describes the generation of LongSAGE libraries from eight time points along a time-course of malting barley. SAGE is an open system capable of capturing global transcript expression data that does not rely on known sequence information, a disadvantage of most other gene expression technologies. SAGE tags are isolated from a defined position within a transcript which enables tags to be annotated by similarity matching to databases of known sequences. Analysis of the most abundant LongSAGE tags highlights the functional groups of dominant mRNA transcripts. This is the largest study to date of global gene expression during grain germination.

Chapter three examines the identification and real-time RT-PCR validation of LongSAGE tags showing differential expression across the malting time course. Transcripts identified with roles in metabolism, particularly in carbohydrate and protein modification, were selected as having a direct relationship to barley seed germination and malting. Real-time RT-PCR was conducted using RNA extracted from the same malting grain samples used in the LongSAGE study. Real-time RT-PCR experiments were performed at four time points corresponding to the LongSAGE study and the mRNA abundance levels observed using real-time RT-PCR compared with the LongSAGE data.

Chapter four describes a microarray study in which the unique tags of the 24 hours post-steeping LongSAGE barley library were used to design a CombiMatrix 12K Custom ElectraSense™ array. Chips were hybridised with cRNA from two important malting time points. The biological sample used to conduct this experiment was the same as that used for the LongSAGE and Real-time RT-PCR experiments. The probes identified as differentially expressed between time points in the malting barley time course are compared between the three gene expression technologies; ElectraSense™ microarray, LongSAGE and real-time RT-PCR.

Chapter five is a general discussion that summarises the findings of the thesis. It also indicates further directions of study and the potential applications of the resources developed from this work.

1.10 References

ABARE. (2003) Australian Commodity Statistics. Canberra: Australian Bureau of Agricultural and Resource Economics.

ABARE. (2007) Australian Commodity Statistics: September Quarter (Wright, A., ed). Canberra: Australian Bureau of Agricultural and Resource Economics.

Aharoni, A. and Vorst, O. (2002) DNA microarrays for functional plant genomics. *Plant Mol. Biol.*, **48**, 99 - 118.

Ayele, B.T., Ozga, J.A. and Reinecke, D.M. (2006) Regulation of GA Biosynthesis Genes during Germination and Young Seedling Growth of Pea (*Pisum sativum* L.). *J. Plant Growth Regu.*, **25**, 219 - 232.

Badr, A., Muller, K., Schafer-Pregl, R., Rabey, H.E., Effgen, S., Ibrahim, H.H., Pozzi, C., Rohde, W. and Salamini, F. (2000) On the Origin and Domestication History of Barley (*Hordeum vulgare*). *Mol. Biol. and Evol.*, **17**, 499 - 510.

Bamforth, C.W. and Barclay, A.H.P. (1993) Malting technology and the uses of malt. In *Barley: Chemistry and Technology*. (MacGregor, A. W. and Bhatty, R. S., eds). Minnesota: American Association of Cereal Chemists, Inc., pp. 297 - 354.

Bartel, D. (2004) MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*, **116**, 281 – 297.

Beltrán-Peña, E., Ortíz-López, A. and Sánchez de Jiménez, E. (1995) Synthesis of Ribosomal-Proteins from Stored Messenger-RNAs Early in Seed-Germination. *Plant Mol. Biol.*, **28**, 327 - 336.

Bewley, J.D. (1997) Seed Germination and Dormancy. *Plant Cell*, **9**, 1055 - 1066.

Bove, J., Jullien, M. and Grappin, P. (2001) Functional genomics in the study of seed germination. *Genome Biol.*, **3**, reviews1002.1001 - reviews1002.1005.

Briggs, D.E. (1998) Malts and Malting. London: Blackie Academic and Professional, Chapman and Hall.

Brodersen, P. and Voinnet, O. (2006) The diversity of RNA silencing pathways in plants. *Trends Genet.*, **22**, 268 – 277.

Brown, A., Drum, F. and Lawrance, L. (2007) Australian Crop Report. Canberra: Australian Bureau of Agricultural and Resource Economics.

Burton, R.A., Shirley, N.J., King, B.J., Harvey, A.J. and Fincher, G.B. (2004) The *CesA* Gene Family of Barley. Quantitative Analysis of Transcripts Reveals Two Groups of Co-Expressed Genes. *Plant Physiol.*, **134**, 224 - 236.

Bustin, S. (2005a) Real-Time PCR. *Encyclopedia of Diagnostic Genomics and Proteomics*, **1**, 1117 - 1125.

Bustin, S. (2005b) Real-Time Reverse Transcription PCR. *Encyclopedia of Diagnostic Genomics and Proteomics*, **1**, 1131 - 1135.

Chandra, G.S., Proudlove, M.O. and Baxter, E.D. (1999) The structure of barley endosperm - An important determinant of malt modification. *J. Sci. Food Agr.*, **79**, 37 - 46.

Chen, J.J., Lee, S.G., Zhou, G.L. and Wang, S.M. (2002a) High-throughput GLGI procedure for converting a large number of serial analysis of gene expression tag sequences into 3 ' complementary DNAs. *Gene Chromosome Canc.*, **33**, 252 - 261.

Chen, J.J., Rowley, J.D. and Wang, S.M. (2000) Generation of longer cDNA fragments from serial analysis of gene expression tags for gene identification. *Proc. Natl. Acad. Sci. U. S. A*, **97**, 349 - 353.

Chen, J.J., Sun, M., Lee, S.G., Zhou, G.L., Rowley, J.D. and Wang, S. (2002b) Identifying novel transcripts and novel genes in the human genome by using novel SAGE tags. *Proc. Natl. Acad. Sci. U. S. A*, **99**, 12257 - 12262.

Chen, K.G. and An, Y.Q.C. (2006) Transcriptional responses to gibberellin and abscisic acid in barley aleurone. *J. Integr. Plant Biol.*, **48**, 591-612.

Close, T.J., Wanamaker, S.I., Caldo, R.A., Turner, S.M., Ashlock, D.A., Dickerson, J.A., Wing, R.A., Muehlbauer, G.J., Kleinhofs, A. and Wise, R.P. (2004) A New Resource for Cereal Genomics: 22K Barley GeneChip Comes of Age. *Plant Physiol.*, **134**, 960 - 968.

Crawford, A.C., White, J.F., Bundock, P., Cordeiro, G., McIntosh, S., Pacey-Miller, T., Rooke, L. and Henry, R.J. (2005) A modified protocol for consistent production of cost-effective LongSAGE libraries. *Plant Mol. Biol. Rep.*, **23**, 139 - 143.

Dean, J.D., Goodwin, P.H. and Hsiang, T. (2002) Comparison of relative RT-PCR and northern blot analyses to measure expression of β-1,3-glucanase in *Nicotiana benthamiana* infected with *Colltotrichum destructivum*. *Plant Mol. Biol. Rep.*, **20**, 347 - 356.

Dean, J.F.D. and Lorenz, W.W. (2004) Transcriptional profiling in plants using Serial Analysis of Gene Expression (SAGE). *AgBiotechNet*, **6**, 1 - 7.

Desprez, T., Amselem, J., Caboche, M. and Hofte, H. (1998) Differential gene expression in Arabidopsis monitored using cDNA arrays. *Plant J.*, **14**, 643 - 652.

Donson, J., Fang, Y., Espiritu-Santo, G., Xing, W., Salazar, A., Miyamoto, S., Armendarez, V. and Volkmuth, W. (2002) Comprehensive gene expression analysis by transcript profiling. *Plant Mol. Biol.*, **48**, 75 - 97.

Duffus, C.M. and Cochrane, M.P. (1993) Formation of the Barley Grain - Morphology, Physiology, and Biochemistry. In *Barley: Chemistry and Technology* (MacGregor, A. W. and Bhatty, R. S., eds). Minnesota: American Association of Cereal Chemists, Inc, pp. 31 - 72.

Dunwell, J.M., Moya-Leon, M.A. and Herrera, R. (2001) Transcriptome analysis and crop improvement (A review). *Biol. Res.*, **34**, 153 - 164.

Fincher, G.B. (1989) Molecular and cellular biology associated with endosperm mobilization in germinating cereal grains. *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, **40**, 305 - 346.

Fincher, G.B. and Stone, B.A. (1993) Physiology and Biochemistry of Germination in Barley. In *Barley: Chemistry and Technology* (MacGregor, A. W. and Bhatty, R. S., eds). Minnesota: American Association of Cereal Chemists, Inc., pp. 247 - 295.

Fizames, C., Munos, S., Cazettes, C., Nacry, P., Boucherez, J., Gaymard, F., Piquemal, D., Delorme, V., Commes, T., Doumas, P. et al. (2004) The Arabidopsis Root Transcriptome by Serial Analysis of Gene Expression. Gene Identification Using the Genome Sequence. *Plant Physiol.*, **134**, 67 - 80.

Fleige, S. and Pfaffl, M.W. (2006) RNA integrity and the effect on the real-time qRT-PCR performance. *Mol. Aspects Med.*, **27**, 126 - 139.

Ghindilis, A.L., Smith, M.W., Schwarzkopf, K.R., Roth, K.M., Peyvan, K., Munro, S.B., Lodes, M.J., Stover, A.G., Bernards, K., Dill, K. et al. (2007) CombiMatrix oligonucleotide arrays: Genotyping and gene expression assays employing electrochemical detection. *Biosens. Bioelectron.*, **22**, 1853-1860.

Gibbings, J.G., Cook, B.P., Dufault, M.R., Madden, S.L., Khuri, S., Turnbull, C.J. and Dunwell, J.M. (2003) Global transcript analysis of rice leaf and seed using SAGE technology. *Plant Biotech. J.*, **1**, 27- 285.

Gibson, S.I. (2004) Sugar and phytohormone response pathways: navigating a signalling network. *J. Exp. Bot.*, **55**, 253 - 264.

Gowda, M., Jantasuriyarat, C., Dean, R.A. and Wang, G.L. (2004) Robust-LongSAGE (RL-SAGE): a substantially improved LongSAGE method for gene discovery and transcriptome analysis. *Plant Physiol*, **134**, 890 - 897.

Hogan, J., Barrett, D., Berry, P., Lawrance, L., Connell, P., Nelson, R. and Want, S. (2004) Australian Crop Report. Canberra: Australian Bureau of Agricultural and Resource Economics.

Ibrahim, A.F.M., Hedley, P.E., Cardle, L., Kruger, W., Marshall, D.F., Muehlbauer, G.J. and Waugh, R. (2005) A comparative analysis of transcript abundance using SAGE and Affymetrix arrays. *Funct. Integrative Genomics*, **5**, 163 - 174.

Jones, R.L. and Jacobsen, J.V. (1991) Regulation of Synthesis and Transport of Secreted Proteins in Cereal Aleurone. *International Review of Cytology*, 126, 49 - 88.

Jung, S.H., Lee, J.Y. and Lee, D.H. (2003) Use of SAGE technology to reveal changes in gene expression in Arabidopsis leaves undergoing cold stress. *Plant Mol. Biol.*, **52**, 553-567.

Kasha, K.J., Falk, D.E. and Ziauddin, A. (1993) Potential Improvement of Barley Quality Through Genetic Engineering. In *Barley: Chemistry and Technology* (MacGregor, A. W. and Bhatty, R. S., eds). Minnesota: American Association of Cereal Chemists, Inc., pp. 419 - 436.

Klemsdal, S.S., Hughes, W., Lonneborg, A., Aalen, R.B. and Olsen, O.A. (1991) Primary structure of a novel barley gene differentially expressed in immature aleurone layers. *Mol. Gen. Genet.*, **228**, 9 - 16.

Klok, E.J., Wilson, I.W., Wilson, D., Chapman, S.C., Ewing, R.M., Somerville, S.C., Peacock, W.J., Dolferus, R. and Dennis, E.S. (2002) Expression Profile Analysis of the Low-Oxygen Response in Arabidopsis Root Cultures. *Plant Cell*, **14**, 2481 - 2494.

Kuhlemeier, C., Green, P.J. and Chua, N.H. (1987) Regulation of Gene Expression in Higher Plants. *Ann. Rev. Plant Physiol*., **38**, 221 - 257.

Langridge, P. and Barr, A.R. (2003) Preface. *Aust. J. Agr. Res.*, **54**, I - IV.

Lapitan, N.L.V., Hess, A., Cooper, B., Botha, A.M., Badillo, D., Iyer, H., Menert, J., Close, T., Wright, L., Hanning, G., Tahir, M. and Lawrence, C. (2009) Differentially expressed genes during malting and correlation with malting quality phenotypes in barley (Hordeum vulgare L.). *Theor. Appl. Genet.*, **118**, 937-952.

Lash, A.E., Tolstoshev, C.M., Wagner, L., Schuler, G.D., Strausberg, R.L., Riggins, G.J. and Altschul, S.F. (2000) SAGEmap: A public gene expression resource. *Genome Res.*, **10**, 1051 - 1060.

Lee, J.Y. and Lee, D.H. (2003) Use of serial analysis of gene expression technology to reveal changes in gene expression in Arabidopsis pollen undergoing cold stress. *Plant Physiol.*, **132**, 517 - 529.

Lee, S., Clark, T., Chen, J.J., Zhou, G.L., Scott, L.R., Rowley, J.D. and Wang, S.M. (2002) Correct identification of genes from serial analysis of gene expression tag sequences. *Genomics*, **79**, 598 - 602.

Liu, H., Hedley, P., Cardle, L., Wright, K.M., Hein, I., Marshall, D. and Waugh, R. (2005) Characterisation and functional analysis of two barley caleosins expressed during barley caryopsis development. *Planta*, **221**, 513 - 522

Lorenz, W.W. and Dean, J.F.D. (2002) SAGE Profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (*Pinus taeda*). *Tree Physiol.*, **22**, 301 - 310.

Luo, M., Liang, X.Q., Dang, P., Holbrook, C.C., Bausher, M.G., Lee, R.D. and Guo, B.Z. (2005) Microarray-based screening of differentially expressed genes in peanut in response to *Aspergillus parasiticus* infection and drought stress. *Plant Sci.*, **169**, 695-703.

MacGregor, A.W., Marchylo, B.A. and Kruger, J.E. (1988) Multiple alpha-amylase components in germinated cereal grains determined by isoelectric focusing and chromatofocusing. *Cereal Chem.*, **65**, 326 - 333.

Mackay, I.M., Arden, K.E. and Nitsche, A. (2002) Real-time PCR in virology. *Nucleic Acids Res.*, **30**, 1292 - 1305.

Matsumura, H., Nirasawa, S. and Terauchi, R. (1999) Transcript profiling in rice (*Oryza sativa* L.) seedlings using serial analysis of gene expression (SAGE). *Plant J.*, **20**, 719 - 726.

Matsumura, H., Reich, S., Ito, A., Saitoh, H., Kamoun, S., Winter, P., Kahl, G., Reuter, M., Kruger, D.H. and Terauchi, R. (2003) Gene expression analysis of plant host-pathogen interactions by SuperSAGE. *Proc. Natl. Acad. Sci. U. S. A*, **100**, 15718 - 15723.

Nilan, R.A. and Ullrich, S.E. (1993) Barley: Taxonomy, Origin, Distribution, Production, Genetics, and Breeding. In *Barley: Chemistry and Technology*. (MacGregor, A. W. and Bhatty, R. S., eds). Minnesota: American Association of Cereal Chemists, Inc., pp. 1 - 29.

Nuutila, A.M., Ritala, A., Salmenkallio-Marttila, M., Aspegren, K., Aikasalo, R., Kurtèn, U., Tammisola, J., Teeri, T.H., Mannonen, L. and Kauppinen, V. (2002) Improvement of malting quality of barley by complementing the malt enzyme spectrum. *Phytochemistry Reviews*, **1**, 135 - 140.

Patankar, S., Munasinghe, A., Shoaibi, A., Cummings, L.M. and Wirth, D.F. (2001) Serial Analysis of Gene Expression in *Plasmodium falciparum* reveals the global expression profile of erythrocytic stages and the presence of anti-sense transcripts in the malarial parasite. *Mol. Biol. Cell*, **12**, 3114 - 3125.

Pauws, E., van Kampen, A.H.C., van de Graaf, S.A.R., de Vijlder, J.J.M. and Ris-Stalpers, C. (2001) Heterogeneity in polyadenylation cleavage sites in mammalian mRNA sequences: implications for SAGE analysis. *Nucleic Acids Res.*, **29**, 1690 - 1694.

Pleasance, E.D., Marra, M.A. and Jones, S.J.M. (2003) Assessment of SAGE in transcript identification. *Genome Res.*, **13**, 1203 - 1215.

Potokina, E., Prasad, M., Malysheva, L., Roder, M. and Graner, A. (2006) Expression genetics and haplotype analysis reveal cis regulation of serine carboxypeptidase I (Cxp1), a candidate gene for malting quality in barley (*Hordeum vulgare* L.). *Funct. Integrative Genomics*, **6**, 25 - 35.

Potokina, E., Sreenivasulu, N., Altschmied, L., Michalek, W. and Graner, A. (2002) Differential gene expression during seed germination in barley (*Hordeum vulgare* L.). *Funct. Integrative Genomics*, **2**, 28 - 39.

Powell, J. (1998) Enhanced concatemer cloning - a modification to the SAGE (Serial Analysis of Gene Expression) technique. *Nucleic Acids Res.*, **26**, 3445 - 3446.

Roth, K.M., Peyvan, K., Schwarzkopf, K.R. and Ghindilis, A. (2006) Electrochemical detection of short DNA oligomer hybridization using the CombiMatrix ElectraSense Microarray reader. *Electroanalysis*, **18**, 1982-1988.

Russell, P.J. (1998) Genetics. California: The Benjamin/Cummings Publishing Company, Inc.

Saha, S., Sparks, A.B., Rago, C., Akmaev, V., Wang, C.J., Vogelstein, B., Kinzler, K.W. and Velculescu, V.E. (2002) Using the transcriptome to annotate the genome. *Nature Biotech.*, **20**, 508 - 512.

Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative Monitoring of Gene-Expression Patterns with a Complementary-DNA Microarray. *Science*, **270**, 467 - 470.

Schnable, P.S., Hochholdinger, F. and Nakazono, M. (2004) Global expression profiling applied to plant development. *Curr. Opin. Plant Biol.*, **7**, 50 - 56.

Slakeski N. and Fincher G.B. (1992) Barley (1->3,1->4)-β-glucanase isoenzyme EI gene expression is mediated by auxin and gibberellic acid, *Febs Letters*, **306**, 98-102.

Smith, L.M., Handley, J., Li, Y., Martin, H., Donovan, L. and Bowles, D.J. (1992) Temporal and spatial regulation of a novel gene in barley embryos. *Plant Mol. Biol.*, **20**, 255 - 266.

Snustad, D.P. and Simmons, M.J. (2000) Principles of Genetics. New York: John Wiley & Sons, Inc.

Sreenivasulu, N., Altschmied, L., Panitz, R., Hähnel, U., Michalek, W., Weschke, W. and Wobus, U. (2002) Identification of genes specifically expressed in maternal and filial tissues of barley caryopses: a cDNA array analysis. *Mol. Genet. Genomics*, **266**, 758 - 767.

Sreenivasulu, N., Altschmied, L., Radchuk, V., Gubatz, S., Wobus, U. and Weschke, W. (2004) Transcript profiles and deduced changes of metabolic pathways in maternal and filial tissues of developing barley grains. *Plant J.*, **37**, 539 - 553.

Sreenivasulu, N., Usadel, B., Winter, A., Radchuk, V., Scholz, U., Stein, N., Weschke, W., Strickert, M., Close, T.J., Stitt, M., Graner, A. and Wobus, U. (2008) Barley grain maturation and germination: Metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. *Plant Physiol.*, **146**, 1738 - 1758.

Street, H.E. and Opik, H. (1970) The Physiology of Flowering Plants: Their Growth and Development. London: Edward Arnold (Publishers) Ltd.

Sun, M., Zhou, G., Lee, S., Chen, J., Shi, R.Z. and Wang, S.M. (2004) SAGE is far more sensitive than EST for detecting low-abundance transcripts. *BMC Genomics*, **5:1**.

Terryn, N. and Rouze, P. (2000) The sense of naturally transcribed antisense RNAs in plants. *Trends Plant Sci.*, **5**, 394 - 396.

Thomas, S.W., Glaring, M.A., Rasmussen, S.W., Kinane, J.T. and Oliver, R.P. (2002) Transcript profiling in the barley mildew pathogen *Blumeria graminis* by serial analysis of gene expression (SAGE). *Mol. Plant Microbe Interact.*, **15**, 847 - 856.

Trendelenburg, G., Prass, K., Priller, J., Kapinya, K., Polley, A., Muselmann, C., Ruscher, K., Kannbley, U., Schmitt, A.O., Castell, S. et al. (2002) Serial Analysis of Gene Expression identifies Metallothionein-II as major neuroprotective gene in mouse focal cerebral ischemia. *J. Neurosci.*, **22**, 5879 - 5888.

Tucholski, J., Skowron, P.M. and Podhajska, A.J. (1995) *Mme*I, a class-IIS restriction endonuclease: Purification and characterization. *Gene*, **157**, 87 - 92.

UNdata (2009) http://data.un.org/Default.aspx

USDA, N. (2004) The PLANTS Database, Version3.5 (http://plants.usda.gov). In *National Plant Data Center*. Baton Rouge, LA 70874-4490 USA.

Valasek, M.A. and Repa, J.J. (2005) The power of real-time PCR. *Adv. Physiol Educ.*, **29**, 151 - 159.

van den Berg, A., van der Leij, J. and Poppema, S. (1999) Serial analysis of gene expression: rapid RT-PCR analysis of unknown SAGE tags. *Nucleic Acids Res.*, **27**, e17 i - iii.

Velculescu, V.E., Vogelstein, B. and Kinzler, K.W. (2000) Analysing uncharted transcriptomes with SAGE. *Trends Genet.*, **16**, 423 - 425.

Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K.W. (1995) Serial analysis of gene expression. *Science*, **270**, 484 - 487.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Hieter, P., Vogelstein, B. and Kinzler, K.W. (1997) Characterization of the yeast transcriptome. *Cell*, **88**, 243 - 251.

Walia, H., Wilson, C., Condamine, P., Ismail, A.M., Xu, J., Cui, X.P. and Close T.J, (2007) Array-based genotyping and expression analysis of barley cv. Maythorpe and Golden Promise. *BMC Genom.*, **8**:87.

Wang, J.M., Zhang, G.P., Chen, J.X. and Wu, F.B. (2004) The changes of beta-glucan content and beta-glucanase activity in barley before and after malting and their relationships to malt qualities. *Food Chem.*, **86**, 223 - 228.

Wang, M., Oppedijk, B., Caspers, M., Lamers, G., Boot, M., Geerlings, D., Bakhuizen, B., Meijer, A. and Van Duijn, B. (1998) Spatial and temporal regulation of DNA fragmentation in the aleurone of germinating barley. *J. Exp. Bot.*, **49**, 1293 - 1301.

Watson L., Henry R.J. (2005) Microarray analysis of gene expression in germinating barley embryos (*Hordeum vulgare* L.). *Funct. Integrative Genomics*, **5**, 155 - 162 .

Watson, L., Holton, T.A. and Henry, R.J. (2001) Exploring Gene Expression Profiles in Germinating Barley. *Proceedings of the 10th Australian Barley Technical Symposium*. Canberra.

Welle, S., Bhatt, K. and Thornton, C.A. (1999) Inventory of High-Abundance mRNAs in Skeletal Muscle of Normal Men. *Genome Res.*, **9**, 506 - 513.

White, J., Pacey-Miller, T., Crawford, A., Cordeiro, G., Barbary, D., Bundock, P. and Henry, R. (2006) Abundant transcripts of malting barley identified by serial analysis of gene expression (SAGE). *Plant Biotech. J.*, **4**, 289 - 301.

Wolbang, C.M., Chandler, P.M., Smith, J.J. and Ross, J.J. (2004) Auxin from the Developing Inflorescence Is Required for the Biosynthesis of Active Gibberellins in Barley Stems. *Plant Physiol.*, **134**, 769 - 776.

Woodward, J. R. and Fincher, G. B. (1982) Purification and Chemical Properties of Two 1,3;1,4-βGlucan Endohydrolases from Germinating Barley. *Eur. J. Biochem.*, **121**, 663 - 669.

Wullschleger, S.D. (2003) Emerging use of gene expression microarrays in plant physiology. *Comp. Funct. Genom.*, **4**, 216 - 224.

Yang, B., Srivastava, S., Deyholos, M.K. and Kav, N.N.V. (2007) Transcriptional profiling of canola (*Brassica napus* L.) responses to the fungal pathogen *Sclerotinia sclerotiorum*. *Plant Science*, **173**, 156 -171.

Zhang, L., Zhou, W., Velculescu, V.E., Kern, S.E., Hruban, R.H., Hamilton, S.R., Vogelstein, B. and Kinzler, K.W. (1997) Gene expression profiles in normal and cancer cells. *Science*, **276**, 1268 - 1272.

Zhang, Y. and Gilles, P. (2003) I-SAGETM Long Kit for digital genome-wide expression profiles from smaller sample sizes. *Focus*, **25**, 11 - 15.

Zhang, L., Castell-Miller, C., Dahl, S., Steffenson, B. and Kleinhofs, A. (2008) Parallel expression profiling of barley-stem rust interactions. *Funct. Integr. Genom.*, **8**, 187-198.

Zhu, T. (2003) Global analysis of gene expression using GeneChip microarrays. *Curr. Opin. Plant Biol.*, **6**, 418 - 425.

Chapter Two

Abundant transcripts of malting barley identified by serial analysis of gene expression (SAGE)

This chapter has been published and the publishers hold Copyright.

White J., Pacey-Miller T., Crawford A., Cordeiro G., Barbary D., Bundock P. and Henry R. J. (2006) Abundant transcripts of malting barley identified by serial analysis of gene expression (SAGE). Plant Biotechnology Journal 4:289-301.

http://onlinelibrary.wiley.com/doi/10.1111/j.1467-7652.2006.00181.x/abstract

Chapter 3

Differential LongSAGE tag abundance analysis in a barley seed germination time course and validation with relative real-time RT-PCR

This chapter has been published and the publishers hold Copyright.

White J., Pacey-Miller T., Bundock P. and Henry R. J. (2008) Differential LongSAGE tag abundance analysis in a barley seed germination time course and validation with relative real-time RT-PCR. Plant Science 175:858-867.

doi:10.1016/j.plantsci.2008.08.008

Chapter 4

Transcriptome analysis of malting barley; comparison of Microarray, LongSAGE and real-time RT-PCR

4.1 Abstract

A 12K CombiMatrix Electrochemical CustomArray was employed to compare gene expression of the Australian malting barley variety Tallon at 12 and 48 hours post-steeping. One hundred and thirteen probes show a greater than ten fold difference in signal intensity, with 64 preferentially expressed at 12 hours post-steeping and 49 at 48 hours post-steeping. The probes with up-regulated expression at 48 hours post-steeping include transcripts associated with stress response, pathological defence and metabolism. The 23kDa Jasmonate-induced protein, JIP-23, is noted as the probe with the highest and fifth highest fold up-regulation from 12 to 48 hours post-steeping. The CombiMatrix ElectraSense™ microarray signal intensity values were compared to expression data from LongSAGE and real-time RT-PCR experiments conducted on the same malting barley samples. The fold change expression data correlated well between the three gene expression technologies. The CombiMatrix ElectraSense™ microarray was the most cost effective and reliable methods for measuring thousands of transcripts simultaneously with the ability to compare many different samples with multiple replicates and relatively lower cost.

4.2 Introduction

Malt barley is the major raw material apart from water used by the brewing industry. The malting of barley involves the controlled germination of the barley grain to promote expression of genes altering the grain composition and generating enzymes for the brewing process. Here we compare the use of different tools to establish the underlying patterns of gene expression during the malting process. High throughput gene expression technologies such as SAGE (Velculescu *et al.*, 1995) and microarrays (Schena *et al.*, 1995) allow us to examine large numbers of gene transcripts in specific samples of malt barley. This provides the opportunity to study changes in gene expression in particular varieties, under various conditions and at specific time points. A time course of malting barley variety Tallon was examined by LongSAGE (White *et al.*, 2006). Eight LongSAGE libraries were constructed from seed sampled at 0, 12, 24, 48, 72, 96 and 120 hours post-steeping and also from dry seed (White *et al.*, 2006). Statistical analysis revealed a set of 57 tags that show significant differential expression across the time course (White *et al.*, 2006). Eight of these were identified as candidates for real-time RT-PCR validation. Real-time RT-PCR was carried out for each of the genes of interest on four of the seed samples analysed in the LongSAGE study. Real-time RT-PCR analysis successfully confirmed the patterns of expression seen in the LongSAGE study (White *et al*., 2006).

Microarray technology allows the simultaneous expression analysis of large sets of genes of known sequence (Schena *et al*., 1995). Differential gene expression has most commonly been measured using two-colour hybridisation. Two mRNA sources are prepared to create two cDNA probes with different fluorophores. Samples of known

concentration are used to calibrate the fluorescence, which is captured by scanning the array with an imaging system (Schena *et al*., 1995). The amount of fluorescence represents the amount of transcript expressed in each sample.

CombiMatrix have developed an electrochemical detection system for oligonucleotide arrays (Ghindilis *et al*., 2007; Roth *et al*., 2006). The CombiMatrix system involves a semiconductor matrix of 12 544 individually addressable platinum microelectrodes on which different oligonucleotides can be simultaneously synthesised via digital control (Ghindilis *et al*., 2007; Roth *et al*., 2006). The electronics used for the oligonucleotide synthesis are subsequently utilized for the detection of redox active chemistries associated with hybridised target molecules (Ghindilis *et al*., 2007; Roth *et al*., 2006). Biotin bound target molecules are labelled with a streptavidin horse radish peroxidase (HRP), the array is exposed to the substrate tetramethylbenzidine (TMB) and hydrogen peroxide, oxidised TMB is reduced at the electrode surface which generates an electrochemical signal that is read with the ElectraSense™ microarray reader (Ghindilis *et al*., 2007; Roth *et al*., 2006).

CombiMatrix have demonstrated that the ElectraSense™ system is a sensitive and reproducible detection platform with effective reusability and performance at a level that is at least competitive with the performance of fluorescence based techniques (Ghindilis *et al*., 2007; Roth *et al*., 2006). The CombiMatrix ElectraSense™ system also offers advantages over existing technologies including lower reagent and equipment costs, a digital data format that does not require image adjustments, the absence of negative effects related to fluorescent dye quenching, ozone-mediated

signal degradation and autofluorescence, and the potential for a high degree of automation (Ghindilis *et al*., 2007; Roth *et al*., 2006).

Microarray analysis is an effective tool for plant functional genomics and has already been successfully used to explore different aspects of the barley transcriptome (Close *et al*., 2004; Potokina *et al*., 2004; Potokina *et al*., 2002; Sreenivasulu *et al*., 2002; Sreenivasulu *et al*., 2004; Watson and Henry, 2005). The recent development and public availability of the Affymetrix 22K Barley1 GeneChip is a major contribution to cereal genomics resources (Close *et al*., 2004). With content derived from a worldwide contribution of 350 000 ESTs from 84 cDNA libraries the array design parameters included high specificity for very similar genes, such as gene family members, so that they might be differentiated (Close *et al*., 2004). In a comparative study of SAGE and Affymetrix array technologies, Ibrahim *et al*. (2005) investigated gene expression in barley caryopses and dissected embryos at 12 days post-anthesis using the 22K Barley1 GeneChip. The Affymetrix GeneChip contains 22 791 probe sets, approximately half of which were given a present call in both the caryopsis and dissected embryo tissues ($p \leq 0.05$) (Ibrahim *et al*., 2005). The 100 most abundant transcripts determined by each method were identified and compared, and strong agreement was found between the SAGE and GeneChip technologies (Ibrahim *et al*., 2005).

This study employs a 12K CombiMatrix Electrochemical CustomArray to compare gene expression in the Australian malting barley variety Tallon at 12 and 48 hours post-steeping. The array was designed using unigene and contig sequences that match to LongSAGE tags from Tallon at 24 hours post-steeping.

4.3 Materials and Methods

4.3.1 Array design

CombiMatrix 12K Custom ElectraSense™ arrays were designed based on data from a LongSAGE library constructed from malt barley at 24 hours post-steeping (White *et al*., 2006). LongSAGE tags were blasted to the *Hordeum vulgare* Gene Index database (Version 9.0, release September 15, 2004 The Gene Index Databases, Dana Farber Cancer Institute, Boston, MA USA http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=barley). Probes were designed using CombiMatrix CustomArray® Software based on the Tentative Consensus sequences (TC's) and singletons that were returned as perfect match hits to the LongSAGE tags. The optimal oligonucleotide length for the CombiMatrix 12k CustomArray is 35-40 base pairs in length. The array includes 6,390 probes in the sense orientation, 1 417 antisense probes and 3 030 features representing 1 515 probes in both orientations. This equates to 10 837 different probes on the array. There are 12 544 total features on a CombiMatrix 12k CustomArray® with 460 features on an ElectraSense™ array used for quality control, 12 000 user defined probes and the remainder being blank features. Eight bacterial QC probes were used for the hybridisation of complementary spike-in oligonucleotides labelled at the 5' end with biotin to enable determination of the linearity of concentration versus signal.

4.3.2 Target preparation, array hybridisation, scanning and stripping

Tissue was collected, RNA extracted and samples from 12 and 48 hours post-steeping used from studies previously described (White *et al*., 2006; White *et al*., 2008). Target RNA was amplified and labelled with a Kreatech RNA ampULSe: Amplification and Labelling Kit for CombiMatrix arrays with Biotin ULS (Cat. no. EA-026; Kreatech Biotechnology, Amsterdam, The Netherlands). All steps were carried out as per the protocol with aRNA fragmentation carried out using reagents from Ambion (Cat. no. AM8740; Ambion, Austin, TX, USA). Hybridization and electrochemical detection was carried out as per the CombiMatrix Protocol for ElectraSense™ 12K Microarray Hybridization and Electrochemical Detection PTL007 and the CombiMatrix ElectraSense™ Detection Kit (Cat. no. 610027; CombiMatrix, Mukilteo, WA, USA). Arrays were scanned using an ElectraSense™ reader from CombiMatrix. Arrays were stripped and rehybridised two to three times each using the CombiMatrix ElectraSense™ Stripping Kit for 12K (Cat. no. 610029; CombiMatrix) as per the CombiMatrix Stripping and Preparation of ElectraSense™ 12K Microarrays for Re-hybridisation protocol PTL003.

Two microarray slides were used to conduct all of the hybridisations for this series of experiments. The RNA samples were hybridised to the arrays in an alternating pattern. Slide one was hybridised with RNA from 12 hours post-steeping and slide two was hybridised with RNA from 48 hours post-steeping, the slides were stripped and prepared for rehybridisation. The slides were then swapped so that slide one was hybridised with RNA from 48 hours post-steeping and slide two was hybridised with RNA from 12 hours post-steeping. This pattern was repeated and in total, slide one

was hybridised twice with RNA from 12 hours post-steeping and twice with RNA from 48 hours post-steeping. Slide two was hybridised twice with RNA from 48 hours post-steeping but only once with RNA from 12 hours post-steeping. Each hybridisation was scanned so that at least two visually acceptable scans were produced.

4.3.3 Quality control, background correction and data normalisation

The image synthesised from signal data for each hybridisation by the ElectraSense™ software was analysed by visual inspection (Figure 4.1). As an initial step in quality control of hybridisations, the concentration of the spike in controls was plotted against the log of the signal of those controls to determine if an approximate linear relationship resulted (Data not shown). Data for the final analysis did not include the scans from the first hybridisation of slide one to RNA from 48 hours post-steeping or the spike-in controls from any of the hybridisations. Data files for replicate scans and hybridisations were merged and average probe signal intensities used to reduce experimental bias. The data were processed using an RMA (Robust Multi-Array Average) background correction, normalised using quantile normalisation and the data from features of replicate probes were averaged.

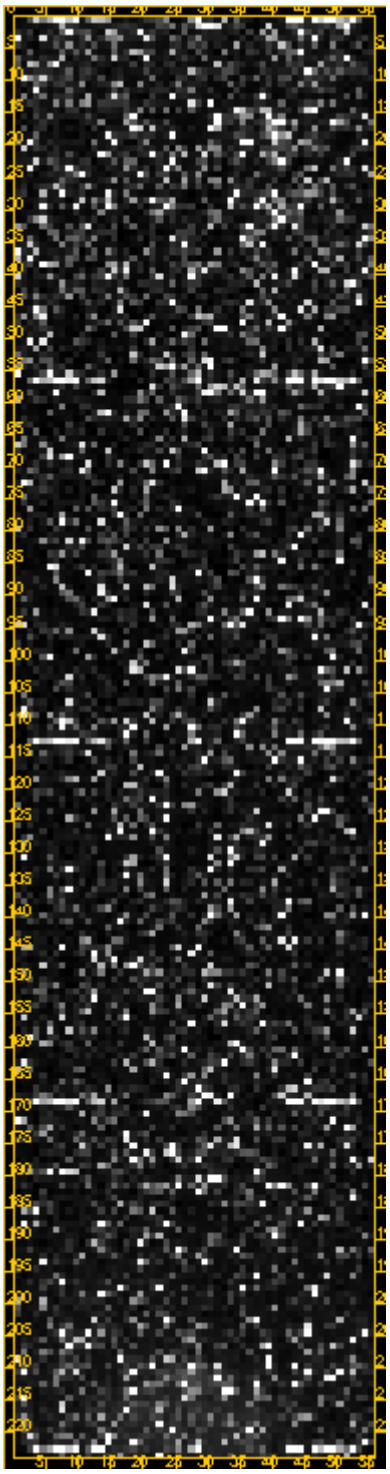


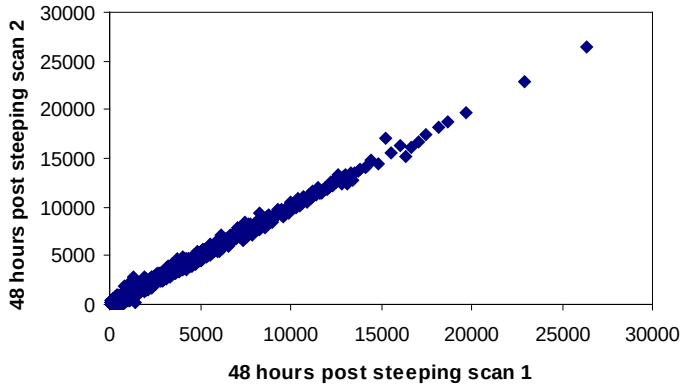
Figure 4.1 Representative ElectraSense™ image of slide one hybridised with RNA from 12 hours post-steeping. The spike-in controls are present in five rows spread evenly across the chip and visible due to their relatively higher signal intensity.

4.4 Results

4.4.1 Microarray analysis of malting barley

Two samples of malting barley from important and physiologically contrasting time points, 12 and 48 hours post-steeping, were chosen for a microarray gene expression analysis study. Microarrays were designed based on SAGE tag results from the 24 hours post-steeping LongSAGE library described by White *et al*. (2006) and malting barley RNA samples from related experiments were used to enable direct comparisons with the relevant results. This analysis assumes that the re-use of slides by stripping and re-hybridising, effects all probes evenly. Scatter plots comparing the probe intensities between replicate scans and between hybridisations on different chips produced very consistent results (Figure 4.2). The median correlation co-efficient for the probe intensities between hybridisations at the M2 time point was 0.931 and for the M4 time point was 0.968 with a very high level of reproducibility observed between replicate scans (median 0.989).

(a)



(b)

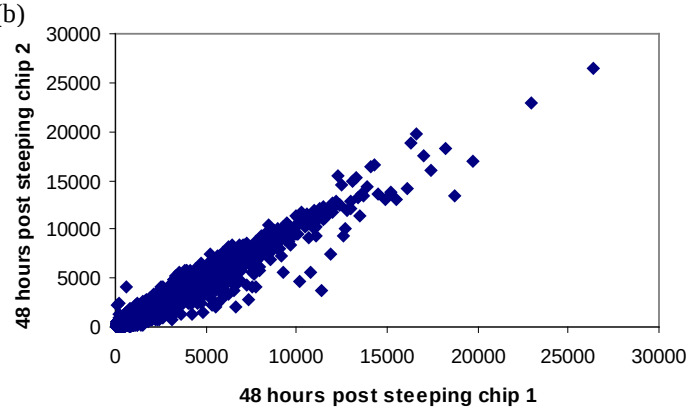


Figure 4.2 Reproducibility of CombiMatrix ElectraSense™ Custom microarrays. Intensity data for each probe is plotted from (a) replicate scans of the same hybridisation, and (b) arrays of the same sample hybridised to different chips. The correlation coefficient (r) for plot (a) is 0.998 and for plot (b) 0.978.

The data from both the LongSAGE and CombiMatrix ElectraSense™ microarray experiments produce a very similar signal range from highest to lowest signal intensity for the samples, 12 and 48 hours post-steeping (Table 4.1). The patterns between the technologies are also very similar. For example at each time point, analysed by each technology, the lowest signal intensity value represents approximately 0.002% of the highest intensity value. However the microarray results appear to provide a better range of signal intensity at low levels. The signal intensity of microarray probes observed at probe numbers 1000 and 2000, in a rank of intensity, represent approximately 0.025% and 0.01% of the total signal on the array respectively. The LongSAGE tag frequency of tag number 1000 is only 0.015% of total tags. The SAGE technology appears more limited in its ability to differentiate expression intensity at the lower end of the frequency range. This limitation may be somewhat overcome by more extensive sequencing of each SAGE library.

Table 4.1 Signal range observed in normalised array data at the probe level and LongSAGE tag count data

Microarray Signal range	12 hours post-steeping	48 hours post-steeping	LongSAGE Signal range	12 hours post-steeping	48 hours post-steeping
Total signal intensity of all probes	7,303,233	7,318,885	**Total tags analysed**	20,042	20,251
Highest intensity value	22,704 (0.31%)	24,641 (0.34%)	**Highest tag count**	499 (2.49%)	203 (1.00%)
Lowest intensity value	61.4 (0.0008%)	58 (0.0008%)	**Lowest tag count**	1 (0.005%)	1 (0.005%)
Signal intensity of probe #1000 in rank of intensity	1868.8 (0.0256%)	1854.3 (0.0253%)	**Tag count of tag #1000 in rank of frequency**	3 (0.015%)	3 (0.015%)
Signal intensity of probe #2000 in rank of intensity	797.6 (0.0109%)	758.6 (0.0104%)	**Tag count of tag #2000 in rank of frequency**	2 (0.01%)	2 (0.01%)

While looking at the LongSAGE results alone, some differences between the time points can be observed. Similar total numbers of LongSAGE tags were analysed in each library, however 12 hours post-steeping has a much higher top ranking frequency count than 48 hours post-steeping. The highest count tag in the 12 hours post-steeping library accounts for almost 2.5% of the total tags analysed in that library. While the highest tag count in the 48 hours post-steeping library accounts for only 1% of total tags analysed. This may indicate that 12 hours post-steeping has a greater diversity of tags at higher frequencies while M4 has a greater diversity of tags at lower frequencies.

4.4.2 Probe identification for Microarray

Probe sequences were compared with complete coding sequences at the NCBI database for *H. vulgare*, and the NCBI database for all barley sequences using BLASTN, to assist in assigning a putative function. Annotations with sequences with perfect matches to probe sequences were assigned along with the original annotations from the TIGR gene indices. The ten microarray probes with the highest signal

intensity in each of the two samples can be compared (Tables 4.2 and 4.3). Five of these are common to the top ten of both malting barley time points. These include nucleic acid binding proteins, metallothionein and lipid transfer proteins, all of which are known to be associated with germination. It appears that those transcripts expressed highly early on remain high throughout germination. This pattern was also observed in the LongSAGE data (White *et al*., 2006)

Table 4.2 Ten probes with the highest Microarray signal intensity in malting barley at 12 hours post-steeping

Probe sequence	TC number	Putative function	12 hours post-steeping signal
CCCTATGTAGTATGGCTACCAGTTATGCCGAGTTTATGC	X60292.1 [a]	Lipid transfer protein	22704.75
TGGAACCATACGCTTGTGTACCTACTTGAATCGAATGT	TC130715	Gamma hordothionin	17054.01
CCCATGGAACCATACGCTTGTGTACCTACTTGAATCGA	TC130715	Gamma hordothionin	16871.29
CGCCCAGTTATCCATCCATCCATCTATCGTCTACC	Z48624.1 [a]	Glycine-rich RNA binding protein	16685.38
CTTTAGTTCCGTCCTCGTGTGCCTTCCCCGTCTTATCTAC	BM816569	Metallothionein	15581.15
TACCGTCTTCGTCAGCAAGTTATCTATCTTCGGTGTGT	TC138719	Nucleic acid binding protein	15277.74
TTTGCCCTCCTATGTGGTTACCAGTGTGGATCCAC	TC139387	Oleosin	14919.17
CGTGAGAGTTGAGAGTTTGTACCATCGGCACTCCC	X60292.1 [a]	Lipid transfer protein	14408.44
TACCTACACCCACAGAAGCCAATGGAGTTCAAGAC	U22951.1 [a]	Gamma-thionin	13491.41
CATGTGTCCCCTGTCGTGTCGTCGGTCTAAATAAG	BQ760074	Unclassified	13373.96

[a] NCBI accession number.

Table 4.3 Ten probes with the highest Microarray signal intensity in malting barley at 48 hours post-steeping

Probe sequence	TC number	Putative function	48 hours post-steeping signal
CTTTAGTTCCGTCCTCGTGTGCCTTCCCCGTCTTATCTAC	BM816569	Metallothionein	24640.66
TTCATCTCCGAAGCTCCAACCGAACCGAGGATGTC	Z34269.1 [a]	Beta-ketoacyl ACP synthase	18072.14
TACCGTCTTCGTCAGCAAGTTATCTATCTTCGGTGTGT	TC138719	Nucleic acid binding protein	17633.50
CGCCCAGTTATCCATCCATCCATCTATCGTCTACC	Z48624.1 [a]	Glycine-rich RNA binding protein	17429.87
CCCTATGTAGTATGGCTACCAGTTATGCCGAGTTTATGC	X60292.1 [a]	Lipid transfer protein	17341.20
GCTACCGTTAGATGTGACGCTACCGTTCCGTGCTA	U49482.1 [a]	Low temperature responsive	15364.16

		RNA binding protein	
GCTTGGGTATCTTGTGTCTGCAGTATCTATCTTTGGCCA	TC131422	Acidic ribosomal protein	15034.94
CATGTGTCCCCTGTCGTGTCGTCGGTCTAAATAAG	BQ760074	Unclassified	14708.79
TAACCAGTGGCTGAGCTTCAGGTTTGCACGACTAT	TC139140	Metallothionein	14544.19
AAACCTGTGAATTCGTCCGCCCGAGTAATTTCAAG	TC139069	Gamma-TIP-like protein	14044.78

[a] NCBI accession number.

Figure 4.3 shows a comparison between the two samples of all probes following normalisation. There is no strong bias towards expression at either time point with this array design. Most expression signals show a less than two fold deviation from the diagonal. 1 822 probes show a greater than two fold difference, 470 probes show a greater than four fold difference and 113 probes show a greater than ten fold difference between the samples. To focus on genes with the highest fold changes in expression from 12 to 48 hours post-steeping, those probes with signal intensities differing by a factor of ten or more were examined (Tables 4.4 and 4.5). Analysis of the probes with a greater than ten fold change in signal intensity revealed 64 probes preferentially expressed at 12 hours post-steeping (Table 4.4) and 49 probes preferentially expressed at 48 hours post-steeping (Table 4.5). The probes which appear to decrease in expression from 12 to 48 hours post-steeping include several transcription factors and stress-related genes. This pattern of expression could reflect several things; that the transcription factor is no longer required after 12 hours post-steeping, that the encoded protein is relatively stable or that these are transcripts remaining from grain fill and maturation. RNA transcripts vary greatly in their stability and time taken to degrade, with mRNA half-lives in higher plants of up to 24 hours (Aalen *et al*., 2001; Narsai *et al*., 2007).

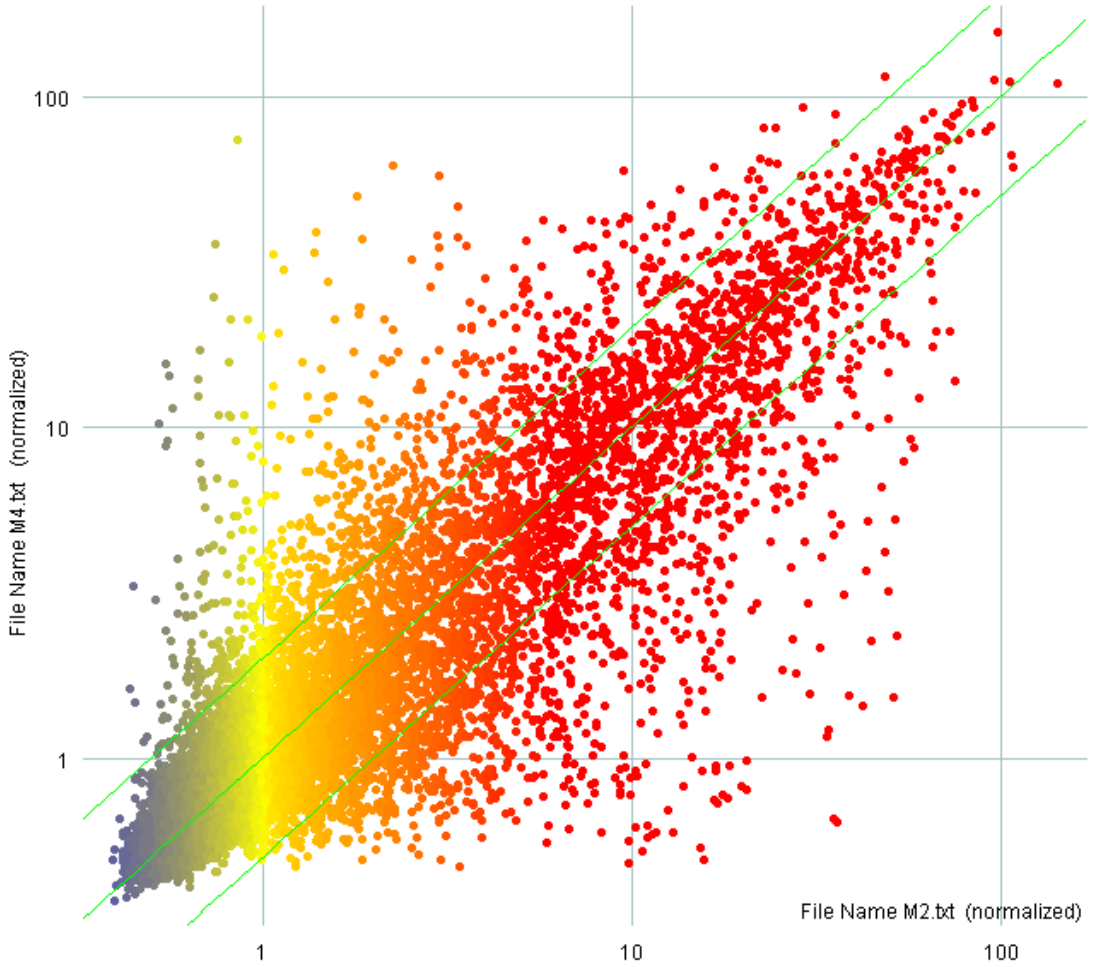


Figure 4.3 Comparison of the normalized signal intensities obtained from barley CombiMatrix ElectraSense™ 12K Custom microarrays. X axis represents normalised signal intensity at 12 hours post-steeping. Y axis represents normalised signal intensity at 48 hours post-steeping. Signals outside the diagonal lines differ by greater than a factor of two between the experiments.

In this study, the probes which show a preference for expression at 48 hours post-steeping are of the greatest interest. It can be assumed that genes with a significant increase in expression at this early stage of malting are significant to the underlying physiology of the malting process. Among the 49 probes with a ten fold or greater increase in expression from 12 to 48 hours post-steeping there are transcripts

associated with stress response, pathogen defence and metabolism. Stress response and pathogen defence transcripts are present as both a preventative and reactive measure to combat pathogen, osmotic, heat and other stresses imposed upon the germinating grain.

The probes with the highest and fifth highest fold change in signal from 12 to 48 hours post-steeping represent the stress response protein JIP-23 (23kDa Jasmonate-induced Protein). JIP-23 is one of several abundant jasmonate-induced proteins of different molecular masses and differing biological functions, the function of several of which is still unknown. The specific expression of JIP-23 mRNA has been noted in the coleoptile and phloem cells of the scutellum and is thought to be expressed in response to osmotic stress brought about by the transport of solutes from the endosperm to the germinating embryo (Hause *et al*., 1996). Watson *et al*. (2005) found that JIP-23 had a greater than ten fold increase in expression after 24 hours post-steeping using cDNA arrays. Glutathione-S-transferase, a protein also recognised as responsive to jasmonic acid, is up-regulated by approximately 13 fold at 48 hours post-steeping and has also been found to exhibit high expression in other studies (Sasaki *et al*., 2000; Potokina *et al*., 2002; Watson *et al*., 2005).

Other probes of interest with a ten fold or greater increase in expression at 48 hours post-steeping include Metallothionein, Serine/Threonine protein kinase, β-d-xylosidase, Glycolate oxidase, pZE40, Dehydroascorbate reductase, Xylose isomerase, α-methyltransferase and Aldose-1-epimerase. Dehydroascorbate reductase has been identified by statistical analysis as associated with the malt quality trait of mash viscosity (Potokina *et al*., 2004). Several of the genes found in this list match to

the 57 tags identified as significantly differentially expressed in the LongSAGE study of malting barley across time (White *et al*., 2006).

Table 4.4 Microarray probes with signal down-regulated more than 10 fold from 12 to 48 hours post-steeping.

Probe sequence	TC number	Putative function	M2 fluorescence signal	M4 fluorescence signal	Fold change
TCCAGAGTGTAAGGAGCTACGTGACCGATTTTGTC	TC140264	Calcium ion binding	5732.62	100.07	57.28
GCTACACCAGCTTCGATCGATGTGTGTGTCTTTCT	TC149834	Fruit weight QTL	5602.90	102.40	54.71
AGAGGCCATGCGGACGATAAGGTGGTTCAGAATTG	TC134605	Unclassified	8072.10	237.88	33.93
AGCTCGTGCAGCGAATCACACCTAGTTCAGAAGCG	TC149899	Unclassified	2493.08	77.09	32.34
CCGATGTCCAGTCACTCGCTACTCGCTAGTCCAGC	TC151637	Viral protein	6682.69	226.34	29.53
TACGATCAACTGGTCGCGGCCATCAAGCAAGAGAT	U01228.1 [a]	Oxygen binding	5344.84	182.46	29.29
CTTCAGTAGTACGTCCACTAACCCAACGATCGACG	TC136288	Unclassified	2438.20	83.40	29.23
AGCTGCCGGCTACGTATGCACCATGGATCTAGTAC	TC140981	Unclassified	5386.63	190.36	28.30
TGGTAGCTCATTCGGCTGACTCCATAGAACCGGAA	TC141920	Unclassified	3223.38	126.35	25.51
CCATGATCCTTAAGAGCCTAGAGATGTAGAGGTGTGTT	TC146947	Cytochrome P450	6140.34	241.61	25.41
TCCGTAGTTGGATTGTATTCGTCCCAATATAAGTGTCGA	TC143944	Putative esterase pir7b	3151.45	129.04	24.42
TTACCCGGATCTTGATATTCGCTTGTATGTGGCCG	TC147886	Serine/threonine protein kinase	2885.95	118.85	24.28
AGCAATGGCGTTAGACTTGATCGGTTGTAAAGACA	TC143368	Wound induced protein	2705.95	114.38	23.66
TCAGCGCACTCCTCTATCTACTTACTTACAGCATGT	TC139746	AP2 transcription factor	2780.00	121.21	22.94
GTAACCTGCTAGAGGAGCTAAGGGCATCTCAAGAC	TC147201	Fiber protein	8299.14	366.75	22.63
GCGAATGGTTTGTAATTGAGATATCTGCATACCGGGA	TC136788	Unclassified	2636.99	117.00	22.54
CAGCTCAGTTGAGTTCAACATCCGGCTTCTCGGTT	TC141008	Zinc finger protein	2047.63	92.84	22.06
CCTAGCAGCCAGCGCATGGATTAGCAGTAATCCCT	TC133816	EBV nuclear antigen 1	3252.92	153.29	21.22
CCTGAGTATGGACGACCTGCAGACTTGTTACATGA	AM500853.1 [a]	NAC transcription factor	1689.53	81.72	20.67
CCTCTGTTTCATTGCGGTAGCGGCTTGAAGCAGAG	TC152041	AP2 transcription factor	1545.31	75.17	20.56
ACTGGCCAAACCATTCAGTGCTCGTGCCCACTTAC	TC132315	Oxygen binding	7079.50	354.31	19.98
TTCAAGATGGATGTTGATCCTCTCGGCCTATTGCA	TC140302	Unclassified	2920.16	148.46	19.67
TGTCCGAGGAGCTGATGGCGTACGAGAGCTACATG	TC152456	Transcription factor	1830.35	93.14	19.65
TCGAGGACGAAGACGCCTTCAACTCATCGGTGGAC	TC136749	Calcium ion binding	1808.35	92.06	19.64
CCTAGCAGCCAGCGCATGGATTAGCACTAATCCCT	TC133815	EBV nuclear antigen 1	4801.65	245.78	19.54
TGAATTCTGATTGACTGATATCATCGGGCATGTTTGAGC	TC150059	AP2 transcription factor	1734.88	89.57	19.37
AGGACTACATGGGCGACATCGAAATGCTAAAGCTA	TC139585	Pyrophosphate-dependent 6-phosphofructose-1-kinase	2684.25	143.70	18.68
TGAGAATTCGGATGTATGAGGTGAACTGCAGTAGCT	TC141634	Reductase	1990.99	109.13	18.24
ATCCTAGGGTCCCGGTAGCTATAGCTGCTTACTTT	TC147875	CG1343-PA	1751.25	97.22	18.01
CCTACATGATCCGAGCTATTCCAAGTAATGCAGCG	TC140023	Pyrophosphate-dependent 6-phosphofructose-1-kinase	2680.86	148.91	18.00
CTGTGCACCAGACATTAAGCGAGTGAGGATAATGT	TC140265	Calcium ion binding	1551.44	87.72	17.69
TCCTCCTGCGGTGCAGATCTTGGTGGTAAGTAGCA	TC146486	Extensin	2444.57	143.34	17.05
TGACTATTCGGTTTAACATGGTAGCTAGCGACCCG	TC148895	BHLH transcription factor	2170.63	128.69	16.87
TACTCTGCATGCATTCTGAACATACCCGGGATGTC	TC139118	Putative early nodulin	1613.61	96.57	16.71
AGACGGTGACGGCAAGGACTAGCTCGTGGAAGAAG	TC134228	Unclassified	7840.21	500.13	15.68
ATCGCAACTTCGCAAGCCACTAGTACTGAGTCTCC					

CGGGTTCAGATCTCGAGGTCGGTGGCATCAAAGTG	TC139332	Unclassified	4407.54	281.79	15.64
CTCCATCGAAGAGGACCGAGACTGCACGTACTACA	TC146946	Cytochrome P450	5115.10	339.35	15.07
TGTACATCTGCTAAGTGCCTTGTTATCAGGGCCTC	TC139310	Unclassified	3547.54	240.10	14.78
	TC145761	Methylmalonyl-CoA decarboxylase	2306.95	156.92	14.70
ATCTTGCCGAGGCATTCGTTTCGCCGTACTGTAAA	TC146271	Anthocyanin 5-O-glucosyltransferase	1304.16	2.70	14.07
TCCTTGTCGCCGGAATGACTCCTAGGACAGCGATA	TC134142	Sterol desaturase	4052.17	289.10	14.02
GAAGTGGGCTAGCAACTTGTGTACGTATGATGCAC	TC149529	Unclassified	1560.29	121.41	12.85
GATGGATCAGGAAGTGCTTGCATATGAGTACCCGA	TC130828	Ribosomal RNA	5941.16	490.92	12.10
TGGAATTCACCTACAGCGAGAAGCCGACACTTGTG	TC146301	Thioredoxin	1473.48	121.93	12.08
TCCCAACTAGTGCTAGCTTAGCGAGAGGTGTTTCT	TC139331	Unclassified	4321.03	358.63	12.05
ACGGTCACGTCTTCAAGGGCTTCCTCTACGACCAG	TC136821	Unclassified	1637.22	138.12	11.85
GATTTACATTGTTCGGGAGCCGGAGAGGTGCTTGC	TC135889	Phytase	1464.04	123.71	11.83
ACCTCCTTCCACTAATTAGACGCCGTGATTAGCTA	TC139125	Putative early nodulin	6804.36	580.48	11.72
TCAGCAACCACATCTACGACTTCATCTCCAAGTTCTGAA	TC132466	Epoxide hydrolase	7673.76	659.81	11.63
TGGTGCCTCATGCTGTAGCGTAGCCCAGTGTAACT	TC139111	Putative early nodulin	2500.29	216.25	11.56
ATCCATCCTAGGGTCCCGGTAGCTATAGCTGCTTA	TC147876	Unclassified	1731.40	150.47	11.51
ACCATCTTGCGTTTCTACGACCTATACTAGTGCTCTG	TC149137	Replication protein	1573.91	137.08	11.48
TCGACTCTCCATATATAATCAACGTTGTACACTGCTGCA	TC139116	Putative early nodulin	1468.80	129.74	11.32
CAAACATCATGGCGGGTACCGTTGACGTCACGACG	TC150072	Transcription factor	1624.03	143.74	11.30
TTCATATTATGTGCTCCTGAGTTGGTGTGCAGAAGC	TC146951	Hypoxia-responsive family protein-like	1070.75	95.74	11.18
GACCAATCATCTACTGGGATCACACGATTCACGGA	TC133180	Unclassified	2189.86	198.79	11.02
AGCCCATTCACGAGTTCTATCGGGTAACGCCAG	CV060467	Unclassified	933.45	86.54	10.79
TCTTCAGTCAAGGCTCTACACTAGGGCGGTACTAG	TC134778	Cold shock induced protein	1194.43	113.74	10.50
CAAGAGGAACGGAACGGAAGGCTGATTGGTTCGCC	TC140712	Wound induced protein	1424.28	137.38	10.37
GTCCACCAAGTCGGATGTGCATGGTTCTCGAGAAT	TC131454	Gamma-thionin	4753.73	462.53	10.28
AACAGCAGCTACGTCTACTTAGTGTTGTCCAGTCA	TC139296	Lipid-transfer protein	1511.10	147.46	10.25
CGTCACCACGGAGGAGATCATGAGACAGATCAAGCC	TC149659	Probable esterase pir7a	1255.43	123.07	10.20
TGATGTTGTAGACATGGATGCGTTGACTAGGGAAAAC	BQ663776	Unclassified	1476.68	144.83	10.20
GACCTCGTCCACTTCGGTATCACCTACTACCTGGT	TC134335	Glutathione transferase	2139.02	211.68	10.10

[a] NCBI accession number.

Table 4.5 Microarry probes with signal up-regulated more than 10 fold from 12 to 48 hours post-steeping.

Probe sequence	TC number	Putative function	M2 fluorescence signal	M4 fluorescence signal	Fold change
GGAAGCTGGGGTTATATCTATGGTGTGAAGCTTGC	TC138639	Jasmonate induced protein	135.00	11690.96	86.60
AGTCGACGATTACTTTACCTAATCATGCATCAACCGTACT	TC146793	Metacaspase	117.26	5610.08	47.84
CACATCGTAGCGCTAACCCAGAGCCATGAACATGA	TC146793	Metacaspase	115.54	3871.89	33.51
GCGGTTGGATCGATCGAACACGATAAACCTGCTGC	TC132616	Beta-D-xylosidase	168.22	5240.73	31.15
CCATCTAAACTAGGTGGAACAGTACGTACACGGCA	TC138652	Jasmonate induced protein	85.77	2458.42	28.66
TTCGTTTGCCAAGCGTGTACATGGTGTCGATATAA	TC132532	Pathogenesis-related protein	284.82	7910.11	27.77
CCATGTCCTATTGGGTCCGCACGTATACACTCATC	TC139061	Glycolate oxidase	219.65	6094.99	27.75

CCTCCATGACATGCCTTAGTTGATCATTAACCATGCTC	TC147022	Unclassified	354.86	9768.19	27.53
TGTGAAAGTTAAGAGGTGCCGGTCTACGTACCCAC	AF181457.1[a]	Dehydrin	179.91	4722.58	26.25
CTCCGTCGGGAGCTCCTACAACCTCATAACCTTCA	TC150722	Proline-rich APG like protein	129.37	3324.17	25.69
GGATGCATGACCAACACTACCGGAAACCTGAATTG	BQ763715	Unclassified	87.97	2236.94	25.43
TTCATCTGCGAAGCTCCAACCGAACCTATGATGTC	BI947407	Metallothionein	106.46	2686.88	25.24
ACTTGAAACTTGAGGTGCCGGTCTGGCTACCTTGC	TC133121	Dehydrin	216.98	5336.53	24.59
CATGACGGGTGAGTGCATGTCATTCAGCGAATAAA	TC140393	Thaumatin-like protein	142.89	3319.96	23.23
TGAGACCGAGATGTGAAGCGTGTGTAAATCCTACC	TC132146	Nucleolar protein family A	104.70	2184.80	20.87
GCCATTGTTCGTATCAGGCATAGCTTCGGTGTTTTG	TC145699	Beta-1,3-glucanase	89.71	1795.47	20.01
ATGCACCTTCGAATTCCTAACTAGCTAGAGGGACC	X74940.1[a]	Pathogenesis-related protein	293.19	5831.04	19.89
CTTCTCAGTTGGAGACAAGCTCGTGTTCAACTACG	BU968132	Unclassified	82.56	1611.67	19.52
TATCGTCCGGCCGTCATGTCAATGTAATGGAGGGT	AJ250662.1[a]	Putative protein	126.90	2462.45	19.40
AGATCGATCACTGTTCCTACTTCGTAGCACCACAA	X98326.1[a]	Dehydrin	474.16	9025.92	19.04
ACCATGTCCTACTCCAACACCGACGAGTGCTTCGA	TC147142	Bundle sheath protein	155.66	2960.13	19.02
CGAGGAACAACTTCACTTGCTCGATCGGCCAAAGC	TC147020	Glycine-rich protein	105.41	1952.73	18.52
GGATTAACCCAAGTCCATTCGCGAGGCACTTCTCC	BM817394	Unclassified	236.54	4309.85	18.22
CAGCTCCAACCAGTCTAGCCGATCAGACACGCTCT	TC130737	Unclassified	174.58	3025.99	17.33
ATTCAAGCTTCCGGTTAATCTCGCATGCAGTTGTA	U01963.1[a]	Germin	86.50	1424.72	16.47
TGTTCAACAATCTTGATACCATCACCAGATCGAAGCAA	M19048.1[a]	Leaf specific thionin	85.66	1378.30	16.09
CACCATGCCCATCCTCGTAAAGGGAGTCATCACCG	TC139061	Glycolate oxidase	119.42	1705.17	14.28
GCTGTTAATGGTCTGCAGTCAGGTCGTTAAGTAATCA	TC136137	Serine/threonine-protein kinase	106.03	1475.20	13.91
GTCTACGTAAGGCGCGACAATGGAGTACCAGGGAC	AF181452.1[a]	Dehydrin	532.59	7305.82	13.72
GGAACTGCTTGCGTTCAGCGATGGGTTTAATAGTGC	TC139406	Glutathione transferase	247.46	3329.47	13.45
ACCTTGCTCGAAAGATCCCTGATACGCTGTTTGGA	TC134034	Arabinoxylan arabinofuranohydrolase isoenzyme	127.26	1705.43	13.40
CTTGGGATTAACACACGTACCATCTCGAGCAGCAA	TC141237	O-methyltransferase	128.12	1691.89	13.21
TCCCAGTAGGAGATGCCGTGCCACTTGACTTCAAC	TC139613	Subtilisin-chymotrypsin inhibitor	397.41	5075.95	12.77
ACGTACGTTCTCGTGTGTTTGCTACCACCTATGAATA	TC146251	Lipid transfer protein	470.56	5967.69	12.68
ATCGCCATCATATTATCAGGTAGCCGCTGCATGCG	BI960632	pZE40	288.20	3624.92	12.58
CAATTGACAGGGCAGTCCAACTACGATCTCGCCGG	AJ276226.1[a]	Chitinase	167.80	2103.03	12.53
GGCGGTTCGTCTACAGTGCTATTAGCTGGTAGATA	AB331641.1[a]	HvMATE gene for aluminum activated citrate transporter	289.70	3603.44	12.44
CCTACGAGATGAAGCTCATCGACGTCAGCAACAA	TC139245	Dehydroascorbate reductase	202.55	2482.96	12.26
CATTGCCAAGATCAACTATGAGGGTCCTACCAGCA	X95257.1[a]	Xylose isomerase	140.01	1713.99	12.24
GCTTCGTGGTTGACTTCGCGAGAAGAATCAAGTAG	AY220734.1[a]	Pathogenesis-related protein	287.27	3471.74	12.09
GGAATCGTGCTTCCATCAGTCTGGTGATTCCAAGA	TC146761	Phosphoethanolamine methyltransferase	475.75	5532.68	11.63
CCCTCCTGCTAGATGGACGCATATATGATCGATCCA	TC132846	Lipid transfer protein	106.06	1223.33	11.53
GTCCCAGTAGGAGATGCCGTGCCACTTGACTTCAA	TC139613	Subtilisin-chymotrypsin inhibitor	532.27	5906.97	11.10
TGATGATGTGTGCATAGGATATCAGCATCGCACTCA	BF259687	Unclassified	165.99	1834.82	11.05
CCTGGCATGCAGTTCTATACCGCCAACTATGTGAA	TC147171	Aldose 1-epimerase	133.65	1431.91	10.71
AAGTTCGGCGTGCCTTGGACAACAAGAGCCCAATA	TC149510	Cytochrome P450	147.24	1536.69	10.44
GGTTCACGCTGGACGGGAAATCCTACCGCTTGTTT	TC140027	Aldose 1-epimerase	226.80	2360.45	10.41
CCATACCAGTGCACCTTCTCCAAGCAGTACGGGTG	TC146794	Metacaspase	473.22	4818.68	10.18
TGAATAACGAGCGTCTATGATCTGATGGAGTCGGG	CB866425	Metallothionein	107.54	1080.94	10.05

[a] NCBI accession number.

4.4.3 Microarray results compared to LongSAGE and real-time PCR results.

The LongSAGE study identified several genes as candidates of interest to the underlying physiology of malting barley (White *et al*., 2006). Several of these genes were analysed by real-time RT-PCR which validated the expression patterns identified by LongSAGE. These genes were represented by one or more different probes on the malting barley CombiMatrix 12K Custom array. For those genes represented by more than one probe the signal intensity from each of the samples was averaged. Fold changes observed between 12 and 48 hours post-steeping for the genes of interest in the microarray experiment, reinforce the fold changes observed in the LongSAGE and real-time RT-PCR experiments (Table 4.6). The cell wall hydrolysing enzyme (1-3,1-4)-β-D-glucanase is significantly up-regulated between 12 and 48 hours post-steeping and is observed to increase approximately seven fold by microarray, almost thirteen fold by LongSAGE and over eight fold by real-time RT-PCR. Alpha-amylase, a key enzyme responsible for degrading the starchy endosperm of the germinating barley grain, shows significant increases in gene expression at 48 hours post-steeping. It represents the highest fold change observed in the real-time RT-PCR experiment with an increase of almost 28 fold, and at nearly 50 fold shows one of the highest increases in expression in the LongSAGE experiment. However, the microarray results show just a three and a half fold increase in α-amylase expression between 12 and 48 hours post-steeping. The carbohydrate metabolising enzyme β-1,3-glucanase, the protein hydrolysing enzyme Cysteine Proteinase EPB, and the key regulator of the glyoxylate cycle, Isocitrate lyase, produced microarray probe signal intensity fold changes consistent with those obtained for the same genes using LongSAGE and real-time RT-PCR. These genes were observed to show

significant up-regulation of expression at 48 hours post-steeping in comparison to 12 hours post-steeping. Pyruvate kinase, which controls the rate of glycolysis, and Xyloglucan endotransglycosylase, an enzyme that cleaves and regulates xyloglucan polymers, show a decrease in expression between 12 and 48 hours post-steeping by approximately three to six fold as analysed by microarray, real-time RT-PCR and LongSAGE.

Table 4.6 Malting barley microarray signal intensities of genes of interest and comparison to gene expression values of the same genes of interest using LongSAGE and Real-time RT-PCR.

Gene of interest	Microarray signal intensity		LongSAGE expression value		Real-time RT-PCR expression value	
Time	12 hours post-steeping	48 hours post-steeping	12hours post-steeping	48 hours post-steeping	12hours post-steeping	48 hours post-steeping
(1-3,1-4)-beta-D-glucanase	1584.91	11056.67	0.065	0.83	0.517	4.220
Fold change		*6.98*		*12.76*		*8.16*
Beta-1,3-glucanase	1912.23	11798.74	0.010	0.040	0.047	0.343
Fold change		*6.17*		*3.95*		*7.31*
Alpha amylase	229.13	798.91	0.010	0.494	0.153	4.240
Fold change		*3.49*		*49.38*		*27.66*
Cysteine proteinase EPB	3340.83	9403.65	0.070	0.494	0.768	4.496
Fold change		*2.81*		*7.07*		*5.85*
Isocitrate lyase	3012.36	7797.07	0.005	0.375	0.417	2.868
Fold change		*2.59*		*75.00*		*6.89*
Pyruvate kinase	9500.15	3289.57	0.130	0.010	1.006	0.373
Fold change		*2.70*[a]		*13.00* [a]		*2.89* [a]
Xyloglucan endotransglycosylase	1962.88	364.67	0.035	0.010	1.054	0.233
Fold change		*4.53* [a]		*3.54* [a]		*5.38* [a]

[a] NB. In these fold change ratios the signal intensity is higher at 12 hours post-steeping.

The majority of the microarray signal intensities produce relatively lower fold change ratios between 12 and 48 hours post-steeping than the LongSAGE and real-time RT-PCR technologies. This may be because of a number of contributing factors.

Microarray signal intensity is subject to background noise caused by cross hybridisation and the misreading of high signals adjacent to low signals. This generally leads to overestimation of the signal associated with very lowly expressing transcripts. This may be the case with α-amylase where a signal intensity of only 229 is present at 12 hours post-steeping which is within the range of background noise associated with the array hybridisation. Another point to note is that the fold change ratio of α-amylase as analysed by LongSAGE is considerably higher. This may be a result of more than one gene family member being identified by the one LongSAGE tag, ultimately overestimating the gene expression at 48 hours post-steeping. The oligonucleotide probe on the CombiMatrix array is approximately fifty percent longer than the SAGE tag, allowing for more specific hybridisation and differentiation of gene family members.

The microarray results for α-amylase are different to the LongSAGE and real-time RT-PCR results not only in terms of the fold change value but in the level of total signal intensity compared to the other types of gene transcripts analysed. In the LongSAGE experiment α-amylase has one of the highest total tag counts at 48 hours post-steeping and the relative expression value of α-amylase identified at the same time point by real-time RT-PCR is also quite high compared to the other gene transcripts studied. However the signal intensity of α-amylase detected in the microarray experiment is quite low at 799 compared to the other gene transcripts which reach over 11 000 units. This may be explained in part by the presence of other gene family members in the case of the LongSAGE experiment but not in the real-time RT-PCR experiment where the PCR primers are specific for one gene family member only.

The binding specificities of particular probe sequences have an effect on signal intensity. For example, if the probe sequence representing α-amylase did not have strong binding efficiency it may result in lower signal intensity. Probe sequence specificity may contribute to a general underestimation of microarray signal intensity due to the cross hybridisation of similar probe sequences especially if the binding efficiency of those sequences was not high. This may be the case for Cysteine Proteinase EPB which has a relatively low fold change value in the microarray experiment compared to the other methodologies, apparently due to a comparatively higher signal intensity detected at 12 hours post-steeping than would be expected (Table 4.5). This could have resulted from the cross hybridisation of other proteinase transcripts to the probe.

It is possible that the fold change values calculated from the microarray results for the genes of interest are underestimated due to the saturation of signal intensity at high levels. Some probes on the array are detected at signal intensity levels above 20 000 units (Table 4.2) compared to a maximum of 12 000 units amongst the genes of interest so this seems unlikely (Table 4.5). Conversely the fold change values detected by LongSAGE tag counts may be overestimated if the SAGE library has not been sequenced to a reasonable depth. For example, the fold change ratios for α-amylase and pyruvate kinase are considerably higher when analysed by LongSAGE than by microarray or real-time RT-PCR (Table 4.5). This is because a false value had to be used to estimate the relative expression of transcripts not detected by even one tag count. Those tags with a count of 0 were given a false count of 0.01 to enable comparisons with other libraries in which the tag was present (White *et al*., 2006). It

cannot be determined whether increasing the number of SAGE tags sequenced would reveal a count for those particular transcripts, and so the fold change is overestimated because the lower level of relative expression is unknown or undetected by the depth of sequencing undertaken in the experiment. The depth of sequencing chosen for SAGE experiments is a critical factor in the effectiveness of the technology for quantitative analysis of gene expression. A comparison of Affymetrix chips with SAGE libraries predicted the sensitivity of an Affymetrix chip to be approximately comparable to a SAGE library of 120 000 total tags (Lu *et al*., 2004).

SAGE is a relatively expensive method of gene expression analysis. To compare different samples; for example, different developmental time points, tissues or treated versus untreated samples; an entire SAGE library must be created for each sample of interest. The process of making a SAGE library takes approximately two days after which the clones have to be sequenced. Advancements in the SAGE protocol, to analyse longer tag concatemers per sequencing run (Crawford *et al*., 2005), and increased throughput and efficiency of modern sequencing platforms has reduced both the cost of sequencing and the time taken to generate large amounts of tag sequence data. The effectiveness of SAGE technology has been shown to rely on the depth of tag sequencing undertaken and in a recent extensive study it was found that approximately 650 000 tags were required to sample a transcriptome of approximately 56 000 (Velculescu *et al*., 1999).

SAGE can be limited in its ability to detect lowly expressed transcripts. The method was developed with the intention of reducing the amplification bias caused by transcripts of significantly high copy number that is observed in some other methods.

By reducing this bias SAGE claims to present a balanced global picture of gene expression (Margulies *et al*., 2001; Velculescu *et al*., 1995). However if the SAGE library is not sequenced to a significant depth, transcripts with low levels of expression may not be detected.

Other factors inherent to the SAGE technology which could result in under or over-estimation of gene expression include, antisense tags, gene family members and alternative splice variants. LongSAGE tags are 21bp long; theoretically long enough to differentiate approximately 4 x 10^{21} different transcripts. However very similar or exact sequences are likely to arise particularly from gene family members. If more than one gene family member shares the same tag sequence, their counts would be combined resulting in an overestimate of the frequency for that gene transcript. This may be the case for α-amylase as mentioned above.

SAGE results can also be underestimated when a transcript is represented by multiple tags. Alternative splicing, incomplete digestion with the anchoring enzyme, alternative polyadenylation and antisense tags are reported as sources of multiple tags (Pauws *et al*., 2001; Gibbings *et al*., 2003; Ibrahim *et al*., 2005; Wahl *et al*., 2005; White *et al*., 2006). These events are a common problem in SAGE studies. White *et al*. (2006) report an occurrence of alternative polyadenylation in a Gamma-thionin transcript. When the tag count of each of the two tags was combined the total relative abundance of the transcript was significantly higher than either of the individual tags indicated. Unless these events are identified and the count of the two or more corresponding tags added together, the transcript expression level of certain SAGE tags can potentially be significantly underestimated.

Despite its downfalls, SAGE is still a useful tool for gene expression analysis. It has the ability to capture expression data for all transcripts in a particular tissue, at a particular time or under specific conditions. This includes transcripts about which no sequence data has previously been obtained. This is a unique advantage of this high throughput gene expression technology.

CombiMatrix ElectraSense™ Microarray technology also has its pros and cons. As seen above, microarray technology provides a more accurate picture of gene expression over a wider range of expression values than SAGE. The CombiMatrix ElectraSense™ CustomArray is also a cost effective method. Once the initial chip has been produced, it is relatively inexpensive to produce replicate chips. The chips can also be stripped and reused up to three times, for a total of four hybridisations per chip. However gene expression values can be subject to cross hybridisation, probe specificity and saturation effects. If these things were corrected, particularly for particular genes of interest, for example for genes directly related to malting quality, then the CombiMatrix ElectraSense™ CustomArray would make an excellent tool for analysis of the transcriptome of barley. A chip such as this could be used to check and QC samples of barley for qualities such as shelf life and malt performance. Such assays could be performed with relatively small amounts of grains, in a short amount of time and at a reasonable cost. This assay would be a valuable tool to provide the malt barley industry with information about the impact of grain storage on shelf life and malt quality, the malting performance of different barley varieties, growing environments and much more.

4.5 Acknowledgements

The authors would like to thank Dana Pascovici and Mervyn Thomas from Emphron Informatics for advice and assistance on background correction and normalisation of microarray data.

4.6 References

Aalen, R.B., Salehian, Z. and Steinum, T.M. (2001) Stability of barley aleurone transcripts: dependence on protein synthesis, influence of the starchy endosperm and destbilization by GA3. *Physiol. Plant*., **112**, 403 – 413.

Close, T.J., Wanamaker, S.I., Caldo, R.A., Turner, S.M., Ashlock, D.A., Dickerson, J.A., Wing, R.A., Muehlbauer, G.J., Kleinhofs, A. and Wise, R.P. (2004) A New Resource for Cereal Genomics: 22K Barley GeneChip Comes of Age. *Plant Physiol*., **134**, 960 - 968.

Crawford, A.C., White, J.F., Bundock, P., Cordeiro, G., McIntosh, S., Pacey-Miller, T., Rooke, L. and Henry, R.J. (2005) A modified protocol for consistent production of cost-effective LongSAGE libraries. *Plant Molecular Biology Reporter*, **23**, 139 - 143.

Ghindilis, A.L., Smith, M.W., Schwarzkopf, K.R., Roth, K.M., Peyvan, K., Munro, S.B., Lodes, M.J., Stover, A.G., Bernards, K., Dill, K. et al. (2007) CombiMatrix oligonucleotide arrays: Genotyping and gene expression assays employing electrochemical detection. *Biosens. Bioelec*., **22**, 1853-1860.

Gibbings, J.G., Cook, B.P., Dufault, M.R., Madden, S.L., Khuri, S., Turnbull, C.J. and Dunwell, J.M. (2003) Global transcript analysis of rice leaf and seed using SAGE technology. *Plant Biotech. J.*, **1**, 271 - 285.

Hause B., Demus U., Teichmann C., Parthier B., Wasternack C. (1996) Developmental and tissue-specific expression of JIP-23, a jasmonate-inducible protein of barley. *Plant Cell Physiol*., **37**, 641 - 649.

Ibrahim, A.F.M., Hedley, P.E., Cardle, L., Kruger, W., Marshall, D.F., Muehlbauer, G.J. and Waugh, R. (2005) A comparative analysis of transcript abundance using SAGE and Affymetrix arrays. *Funct. Integrative Genomics*, **5**, 163 - 174.

Lu J., Lal A., Merriman B., Nelson S. and Riggins G. (2004) A comparison of gene expression profiles produced by SAGE, longSAGE and oligonucleotide chips. *Genomics*, **84**, 631 - 636.

Margulies E.H., Kardia S.L.R. and Innis J.W. (2001) A Comparative Molecular Analysis of Developing Mouse Forelimbs and Hindlimbs Using Serial Analysis of Gene Expression (SAGE). *Genome Res*., **11**, 1686 - 1698.

Narsai R., Howell K.A., Millar A.H., O'Toole N., Small I. and Whelan J. (2007) Genome-wide analysis of mRNA decay rates and their determinants in Arabidopsis thaliana. *Plant Cell*, **19**, 3418 - 3436.

Pauws, E., van Kampen, A.H.C., van de Graaf, S.A.R., de Vijlder, J.J.M. and Ris-Stalpers, C. (2001) Heterogeneity in polyadenylation cleavage sites in mammalian mRNA sequences: implications for SAGE analysis. *Nucleic Acids Research*, **29**, 1690 - 1694.

Potokina, E., Caspers, M., Prasad, M., Kota, R., Zhang, H., Sreenivasulu, N., Wang, M. and Graner, A. (2004) Functional association between malting quality trait components and cDNA array based expression patterns in barley (*Hordeum vulgare* L.). *Mol. Breeding*, **14**, 153 - 170.

Potokina, E., Sreenivasulu, N., Altschmied, L., Michalek, W. and Graner, A. (2002) Differential gene expression during seed germination in barley (*Hordeum vulgare* L.). *Funct. Integrative Genomics*, **2**, 28 - 39.

Roth, K.M., Peyvan, K., Schwarzkopf, K.R. and Ghindilis, A. (2006) Electrochemical detection of short DNA oligomer hybridization using the CombiMatrix ElectraSense Microarray reader. *Electroanalysis*, **18**, 1982-1988.

Sasaki, Y., Asamizu, E., Shibata, D., Nakamura, Y., Kaneko, T., Awai, K., Masuda, T., Shimada, H., Takamiya, K.I., Tabata, S. and Ohta, H. (2000) Genome-wide expression-monitoring of jasmonate-responsive genes of Arabidopsis using cDNA arrays. *Biochem. Soc. Trans*., **28**, 863 – 864.

Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative Monitoring of Gene-Expression Patterns with a Complementary-DNA Microarray. *Science*, **270**, 467 - 470.

Sreenivasulu, N., Altschmied, L., Panitz, R., Hähnel, U., Michalek, W., Weschke, W. and Wobus, U. (2002) Identification of genes specifically expressed in maternal and filial tissues of barley caryopses: a cDNA array analysis. *Mol. Genet. Genom*., **266**, 758 - 767.

Sreenivasulu, N., Altschmied, L., Radchuk, V., Gubatz, S., Wobus, U. and Weschke, W. (2004) Transcript profiles and deduced changes of metabolic pathways in maternal and filial tissues of developing barley grains. *Plant J*., **37**, 539 - 553.

Wahl, M.B., Heinzmann, U. and Imai, K. (2005) LongSAGE analysis significantly improves genome annotation: identifications of novel genes and alternative transcripts in the mouse. *Bioinformatics*, **21**, 1393 - 1400.

Watson, L. and Henry, R.J. (2005) Microarray analysis of gene expression in germinating barley embryos (*Hordeum vulgare* L.). *Funct. Integrative Genomics*, **5**, 155 - 162.

White, J., Pacey-Miller, T., Crawford, A., Cordeiro, G., Barbary, D., Bundock, P. and Henry, R. (2006) Abundant transcripts of malting barley identified by serial analysis of gene expression (SAGE). *Plant Biotech. J*., **4**, 289 - 301.

Velculescu V.E., Madden S.L., Zhang L.Z., Lash A.E., Yu J., Rago C., Lal A., Wang C.J., Beaudry G.A., Ciriello K.M., et al. (1999) Analysis of human transcriptomes. *Nature Genet*., **23**, 387 - 388.

Chapter Five

General Conclusions

5.1 Conclusions and summary

The genetic processes involved in barley seed germination, or malting, are poorly understood. Physiological and biochemical changes are occurring within the grain at all times and this study aimed to determine the genetic processes that occur during malting that are associated with these changes.

This study used two large scale gene expression analysis technologies to examine gene expression changes in malting barley over time. As technology has developed over the past ten years, LongSAGE and microarray have become two of the most commonly used methods for generating large amounts of gene expression data. In this study 155 000 LongSAGE tags, representing 41 909 unique tag sequences, were generated from eight tag libraries of malting barley. Analysis of the most abundant LongSAGE tags highlighted the functional groups of dominant mRNA transcripts which include transcripts coding for stress response and cell defence, ribosomal proteins, storage proteins, cell structure and plant growth, and metabolism related transcripts. Annotation of the LongSAGE tags revealed that multiple tags matched to the same transcript. The presence of alternative transcripts has been explored in several SAGE studies and appears to be a result of several variables including alternative splicing, incomplete digestion, alternative polyadenylation and anti-sense transcripts (Pauws *et al.*, 2001; Gibbings *et al.*, 2003; Robinson *et al.*, 2004; Ibrahim *et al.*, 2005; Wahl *et al.*, 2005). SAGE libraries can therefore also be used to examine

aspects of mRNA processing such as the presence of anti-sense transcripts and enzyme sites that enable alternative tags to be created.

Statistical analysis confidently identified 57 LongSAGE tags with significant changes in abundance across the eight malting time points. Three general patterns of expression were observed; tags with higher expression during the first 12 to 24 hours post-steeping, tags showing up-regulation from 24 hours post-steeping and tags with specific up-regulation at 24 hours post-steeping and 120 hours post-steeping. A sample of LongSAGE tags showing significant changes in gene expression across time were chosen for validation by real-time RT-PCR. Metabolism related transcripts, particularly carbohydrate and protein modifying transcripts, were selected for real-time RT-PCR analysis because of a more direct relationship to seed germination and malting.

Real-time RT-PCR experiments were performed on samples obtained at four time points corresponding to the LongSAGE study; pre-steeping, 12 hours post-steeping, 48 hours post-steeping and 96 hours post-steeping. The mRNA abundance levels observed by real-time RT-PCR were compared with the abundance levels from the LongSAGE experiment with similar patterns observed for the two methods. The RT-PCR results provide support for the precision of detection achievable with SAGE and provide greater confidence in the levels of expression detected in the study. Three transcripts, α-amylase, (1-3,1-4)-β-D-glucanase and cysteine proteinase EP-B, produced very similar patterns of expression. The transcripts for these enzymes were significantly up-regulated at 48 hours post-steeping and remained significantly up-regulated throughout malting. The co-ordinated expression of these enzymes provides

substantial support for the view that these enzymes have key roles as the major contributing enzymes in the depolymerisation of starch, the breakdown of cell walls and degradation of protein in the germinating grain.

Chapter four describes a microarray study of malting barley samples. A CombiMatrix 12K ElectraSense™ Custom array was designed on the basis of the unique LongSAGE tags from 24 hours post-steeping. Chips were hybridised with cRNA from 12 and 48 hours post-steeping. The reproducibility of results was examined and very high correlation co-efficients were seen for replicate scans and between chip comparisons. Re-use of the chips of up to three times did not appear to have adverse effects on the signal intensity results and this contributes to the economic advantages of the CombiMatrix ElectraSense™ array system.

There is always the risk in any study for experimental error. Several elements were incorporated into the experimental design to minimize error in this study. The same biological samples were used for each of the three experiments, the LongSAGE libraries, the real-time RT-PCR experiments and the ElectraSense™ microarray hybridisations. Multiple cDNA synthesis reactions and duplicate PCRs were performed in the real-time RT-PCR experiment. A reference gene was amplified in each experiment to normalise all of the real-time RT-PCR experiments and genes of interest were compared to multiple internal control genes. The microarray experiment included several chip swaps, stripping and re-hybridisations to test the reproducibility of results with very high correlation co-efficients. Spike-in controls were also included to control for bias across an individual chip.

Gene expression levels for eight genes of interest were compared from the three different technologies. In most cases fold change values between 12 and 48 hours post-steeping were not remarkably dissimilar, however differences were observed. Several variables contribute to these differences, some which can be readily controlled and others which cannot. Transcript abundances can be underestimated if a transcript is represented by more than one SAGE tag, however if these occurrences are identified the tag count of the individual tags could be added together to generate the frequency by which the transcript is actually represented in the library. The opposite problem can occur with gene family members, whereby more than one transcript can be represented by a unique tag. However with LongSAGE this is unlikely since a tag of 21 bp is likely to provide a unique tag signature for each transcript. The depth to which a library is sequenced will determine how many different transcripts are identified and genes with low expression may not be detected. Despite these limitations SAGE is a very useful tool for gene expression analysis. It provides the opportunity to capture expression data for all transcripts in a particular tissue, at a particular time or under particular conditions, including transcripts about which there is no known sequence data.

The CombiMatrix ElectraSense™ microarray is a very cost effective and accurate tool for large scale gene expression studies. With the ability to re-use chips and the sensitive electrodetection system it has several advantages over other oligonucleotide chip systems. This experiment examined a large number of transcripts and compared the expression at 12 and 48 hours post-steeping. The CombiMatrix ElectraSense™ array has the potential to be used in further studies to compare many different time

points, varieties and treatments of barley grains to generate valuable information for the malting and brewing communities.

This study has highlighted many areas of interest for potential further work. The data collected in this work could be used in further experiments to determine the distinction between transcript levels in the embryo versus the endosperm. The level of various proteins could be examined using assays or antibodies to examine if protein level correlates with transcript level. siRNA or microRNA experiments could be used to determine the level of effect of post-transcriptional silencing. There are also several new technologies available for whole transcriptome analysis including Next Generation (NG) sequencing or 'deep sequencing' techniques which include 454 sequencing, Illumina sequencing and the SOLiD system (Wall *et al.*, 2009). These technologies allow the analysis of the whole transcriptome or in many cases the whole genome of an organism in one run. The potential application for this technology is enormous. The same biological samples were used for these experiments which enabled the comparison of three technology platforms however it would also be useful to perform some of these experiments with different biological samples.

This study has provided valuable insight into the types of genes expressed during malting and the temporal patterns in which they are expressed. This information provides a useful understanding of the genetic control of the biological and physiological changes which occur in the embryo and endosperm of the germinating barley grain. Other important outcomes of this study include an extensive LongSAGE database of quantitative transcript data which covers the malting barley transcriptome of a common Australian malting variety. This resource has been validated by real-

time RT-PCR and microarray analyses. The LongSAGE database and the CombiMatrix ElectraSense™ microarray chip are valuable tools for further research of malting barley qualities.

5.2 References

Gibbings, J.G., Cook, B.P., Dufault, M.R., Madden, S.L., Khuri, S., Turnbull, C.J. and Dunwell, J.M. (2003) Global transcript analysis of rice leaf and seed using SAGE technology. *Plant Biotechnology Journal*, **1**, 271 - 285.

Ibrahim, A.F.M., Hedley, P.E., Cardle, L., Kruger, W., Marshall, D.F., Muehlbauer, G.J. and Waugh, R. (2005) A comparative analysis of transcript abundance using SAGE and Affymetrix arrays. *Functional & Integrative Genomics*, **5**, 163 - 174.

Pauws, E., van Kampen, A.H.C., van de Graaf, S.A.R., de Vijlder, J.J.M. and Ris-Stalpers, C. (2001) Heterogeneity in polyadenylation cleavage sites in mammalian mRNA sequences: implications for SAGE analysis. *Nucleic Acids Research*, **29**, 1690 - 1694.

Robinson, S.J., Cram, D.J., Lewis, C.T. and Parkin, I.A.P. (2004) Maximizing the Efficacy of SAGE Analysis Identifies Novel Transcripts in Arabidopsis. *Plant Physiol*, **136**, 3223 - 3233.

Wahl, M.B., Heinzmann, U. and Imai, K. (2005) LongSAGE analysis significantly improves genome annotation: identifications of novel genes and alternative transcripts in the mouse. *Bioinformatics*, **21**, 1393 - 1400.

Kerr Wall, P., Leebens-Mack, J., Chanderbali, A.S., Barakat, A., Wolcott, E., Liang, H., Landherr, L., Tomsho, L.P., Hu, Y., Carlson, J.E., Ma, H., Schuster, S.C., Soltis, D.E., Soltis, P.S., Altman, N. and dePamphilis, C.W. (2009) Comparison of next

generation sequencing technologies for transcriptome characterization. *BMC Genom.*, **10:**347

Appendix One

Acceptance letter

"Differential gene expression analysis in a barley seed germination time course and confirmation with relative real-time RT-PCR" in the journal, Plant Science

Date: Fri, 15 Aug 2008 09:49:24 +0100
From: german.spangenberg@dpi.vic.gov.au
Subject: Your Submission
To: robert.henry@scu.edu.au
X-IronPort-Anti-Spam-Filtered: true
X-IronPort-Anti-Spam-Result:
AqMBAKvfpEiRJNeKmmdsb2JhbACHL4NxhWOBBAEBAQEBCAUIBxGaRIZZg
VY
X-IronPort-AV: E=Sophos;i="4.32,214,1217772000"; d="scan'208";a="12893230"
X-SMTP-Vilter-Version: 1.2.4
X-SMTP-Vilter-Virus-Backend: clamd
X-SMTP-Vilter-Status: clean
X-SMTP-Vilter-clamd-Virus-Status: clean
Original-recipient: rfc822;robert.henry@muon.lis.scu.edu.au
X-OriginalArrivalTime: 15 Aug 2008 08:49:24.0992 (UTC)
FILETIME=[D1A52000:01C8FEB3]

Ms. Ref. No.: PSL-D-07-00648R2
Title: Differential LongSAGE tag abundance analysis in a barley seed germination time course and validation with relative real-time RT-PCR
Plant Science

Dear Prof. Robert Henry,

I am pleased to confirm that your paper "Differential LongSAGE tag abundance analysis in a barley seed germination time course and validation with relative real-time RT-PCR" has been accepted for publication in Plant Science.

Thank you for submitting your work to this journal.

With kind regards,

German Carlos Spangenberg, Ph.D DSc
Receiving Editor
Plant Science

Appendix Two

Copyright permission letters

Wiley-Blackwell publishing - Plant Biotechnology Journal

Subject RE: Copyright permission for publication in thesis.
From Journals Rights <jrights@wiley.com>
Date Tuesday, August 26, 2008 5:30 pm
To "jwhite@scu.edu.au" <jessica.white@scu.edu.au>

Dear Jessica White

Thank you for your email request. Permission is granted for you to use the material below for your thesis, subject to the usual acknowledgements and on the understanding that you will reapply for permission if you wish to distribute or publish your thesis commercially.

Good luck!

Yours sincerely

Sue Mattingley

Senior Rights Controller

From: jessica.white@scu.edu.au [mailto:jessica.white@scu.edu.au]
Sent: 25 August 2008 04:34
To: Journals Rights
Subject: Copyright permission for publication in thesis.

To whom it may concern,

I request permission to include the journal article

White J., Pacey-Miller T., Crawford A., Cordeiro G., Barbary D., Bundock P. and Henry R. J. (2006) Abundant transcripts of malting barley identified by serial analysis of gene expression (SAGE). Plant Biotechnology Journal 4:289-301

of which I am the primary author, in its entirety in my PhD thesis.

Regards,

Jessica White

Elsevier publishing - Plant Science.

Subject RE: Copyright permission for publication in thesis.
From "Gould, Laura (ELS-OXF)" <L.Gould@elsevier.com>
Date Tuesday, August 26, 2008 8:22 pm
To jessica.white@scu.edu.au

Dear Ms White

As per your letter dated 25 August 2008, we hereby grant you permission to reprint the below mentioned material at no charge **in your thesis** subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.

2. Suitable acknowledgment to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"This article was published in Publication title, Vol number, Author(s), Title of article, Page Nos, Copyright Elsevier (or appropriate Society name) (Year)."

3. Your thesis may be submitted to your institution in either print or electronic form.

4. Reproduction of this material is confined to the purpose for which permission is hereby given.

5. This permission is granted for non-exclusive world **English** rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form other than submission. Should you have a specific electronic project in mind please reapply for permission.

6. Should your thesis be published commercially, please reapply for permission.

Yours sincerely

Laura Gould

Rights Assistant

Your future requests will be handled more quickly if you complete the online form at www.elsevier.com/permissions

From: jessica.white@scu.edu.au [mailto:jessica.white@scu.edu.au]
Sent: 25 August 2008 04:38
To: Rights and Permissions (ELS)
Subject: Copyright permission for publication in thesis.

To whom it may concern,

I request permission to include the journal article

Ms. Ref. No.: PSL-D-07-00648R2
White J., Pacey-Miller T., Bundock P. and Henry R. J. (2008) Differential LongSAGE tag abundance analysis in a barley seed germination time course and validation with relative real-time RT-PCR.

of which I am the primary author, in its entirety, in my PhD thesis.

Regards,
Jessica White