NOVEL APPROACHES FOR EVALUATING BRASSICA GERMPLASM FOR INSECT RESISTANCE

by

Garima P Sharma

A thesis submitted to The University of Birmingham for the degree of

DOCTOR OF PHILOSOPY

Supervisors: Prof. Brian Ford-Lloyd and Dr. Jeremy Pritchard

School of Biosciences

The University of Birmingham

September, 2015

ABSTRACT

Brassica crops are grown worldwide for food and oil due to their nutritional, medicinal, bio-industrial and crop rotation properties. They suffer from insect pests which cause large yield and economic losses. Application of insecticides is the preferred way of dealing with insect problems, but this way of pest control is not only hazardous to the environment, it also affects humans as the chemicals easily get incorporated into the food chain. As a result, new more resistant varieties are urgently needed to meet the demand of growing populations.

Phenotyping of germplasm is often the first step in breeding new insect resistant varieties. A set of 200 accessions were classified as resistant (non-preferred) or susceptible (preferred) in response to cabbage aphid *(Brevicoryne brassicae)* feeding based on aphid performance in the field. Fifteen accessions were further assessed to identify the level and location of resistance factors by investigating feeding behaviour of cabbage aphid using the Electrical Penetration Graph (EPG) technique. This study aimed to characterize and compare feeding behaviour of cabbage aphid for resistance screening in Brassica germplasm and classify the germplasm into susceptible or resistant groups based on the acceptability of feeding on each genotype. The feeding behaviour assessment revealed the presence of both interspecific and intraspecific variation. The results also indicated that cabbage aphid encountered plant resistance at multiple levels i.e. at the surface of leaves, epidermis/mesophyll and sieve elements in order to feed.

The gene expression of these accessions was also investigated under induced (presence of aphid feeding for 24h) and non-induced (absence of aphid feeding) conditions. The transcriptional response after 24 h showed that gene expression is highly regulated in

response to aphid feeding. The gene expression analysis and gene ontology (GO) enrichment study helped identify candidate genes which may contribute to cabbage aphid resistance. The gene probes NPC6, PDCB3, At1g10155; At3g56240; At5g09650 were identified as strong candidates for surface based and phloem based resistance. It is strongly recommended that identified candidate genes should be further researched and studied to confirm their association with aphid feeding in future. There are a large number of genes which are still not annotated but were found significant, leaving a future possibility to investigate their function in cabbage aphid resistance in Brassica plants. In addition to this, the gene expression differences between crop wild relatives and landraces indicated adaptations of landraces during the process of domestication. Lastly, the Gene expression data were also successfully used to develop models to predict insect resistance status. . The findings from this study will contribute to current and future research in investigating cabbage aphid resistance in Brassica for comprehensive and durable insect resistant cultivars.

In conclusion, the combination of phenomics, physiology and transcriptomics provides an opportunity to assess Brassica germplasm for further research into defence mechanisms of cabbage aphids and for inclusion in the breeding programme.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Professor Brian Ford-Lloyd and Dr. Jeremy Pritchard sharing their expertise and providing guidance and support throughout the last four years. Particular thanks also to Dr. Nigel Maxted and Shelagh Kell for their advice and understanding throughout the project. I would also like to thank Dr Scott Hayward, who has given me good advice and feedback over the years.

I would like to acknowledge PGR Secure for providing the financial support for this research. My thanks extend to all colleagues involved with PGR Secure project who helped me throughout my work.

A particular thank you Laura Vickers, Rashed, Hannah, Holly, Maria, Arami, Anushree for all the fun, help and support that really made my PhD experience exciting and scientifically stimulating. I would also like to thank Fay Hughes, Lorraine Wallace, Karen Staples and Susan Bradshaw for their help with every step of my lab works.

Special thanks to my family (Papaji, Ma, Sunaindi, Jija and Mondi) for their constant support and unwavering belief in me (and for the survival kits!). Most importantly a huge thanks to my husband Prashant and daughter Shambhavi who have emotionally supported me throughout. Thank you both for listening to my moaning and rants on things you never understood, as well as for giving me lifts in and out of the labs when I was working unsociable hours.

Finally I am thankful to my parents and B2, for they have given me their unequivocal support throughout my life.

TABLE OF CONTENTS

Table of Contents.....i

List of Tables.....vii

List of Figures.....xii

List of Supplementary Tables.....xv

List of Abbreviation.....xx

***CHAPTER 1:* Review of literature & general introduction.....1**

1.1 The Experimental Plant: Brassicas.....2

1.1.1 Importance of Brassica: economic, social and medicinal.....2

1.1.2 Brassica genetics.....5

1.1.3 Crop Wild Relatives (CWR) and Loss of Resistance during Domestication.....7

1.2 Insect pests.....9

1.2.1 Importance of insect pests in agriculture.....9

1.2.2 Cabbage aphid (*Brevicoryne brassicae*).....10

1.2.3 Life cycle and reproduction.....11

1.2.4 Location and acceptance of host plant.....12

1.2.5 Feeding mechanism and avoiding host defences.....12

1.3 Aphid-plant interactions: plant defence strategies.....14

1.3.1 Insect resistance in plants.....14

1.4 Electrical Penetration Graph (EPG) technique: monitoring aphid feeding behaviour..... 16

1.4.1 EPG Waveforms.....18

1.4.2 EPG waveform description.....18

1.5 The 'omics' revolution in plant research.....19

1.5.1 Microarrays: Tool for whole genome expression study.....23

1.5.2 Microarray: Applications.....24

1.6 PGR secure project.....26

1.6.1 Research aims and thesis outline.....28

***CHAPTER 2*: Field screening Brassica germplasm for feeding suitability to cabbage aphid and whitefly.....31**

2.1 Abstract.....32

2.2 Introduction.....33

2.2.1 Phenotyping in field conditions vs. controlled environment.....34

2.2.2 Chapter Aim.....36

2.3 Materials and methods.....36

2.3.1 Plant material.....36

2.3.2 Field experiment.....38

2.3.3 Cultivation of Brassica accessions for field experiment.....38

2.3.4 Statistical analysis.....40

2.4 Results.....41

2.4.1 Field experiment.....41

2.4.2 Observational scoring and statistical analysis.....43

2.5 Discussion.....46

2.6 Conclusion.....48

List of Supplementary table.....48

CHAPTER 3: Investigating the differences in feeding behaviour of cabbage aphid on Brassica genotypes using Electrical Penetration Graph (EPG).....51

3.1 Abstract.....52

3.2 Introduction.....54

3.2.1 Aphid feeding.....56

3.2.2 Electrical Penetration Graph (EPG).....57

3.2.3 Chapter Aim.....60

3.3 Material and methods.....61

3.3.1 Plant material.....61

3.3.2 Plant culturing.....61

3.3.3 Aphid species and infestation conditions.....62

3.3.4 Electrical Penetration Technique.....63

3.3.5 Statistical Analysis.....64

3.4 Results.....65

3.4.1 Analysis of feeding behaviour of cabbage aphid (*B.brassicae*) on different species of Brassica.....65

3.4.2 Comparison of interspecies variation in performance of cabbage aphid on Brassica species.....66

3.4.3 Phloem location and acceptance.....70

3.5 Discussion.....72

3.6 Conclusions.....77

CHAPTER 4: Differential gene expression profiling of diverse Brassica genotypes in response to cabbage aphid feeding using microarray.....78

4.1 Abstract.....79

4.2 Introduction.....81

4.2.1 Chapter Aim.....85

4.3 Material and methods.....86

4.3.1 Brassica plants and cabbage aphid.....86

4.3.2 Sample RNA Preparation.....87

4.3.3 Minimum Information about a Microarray Experiment (MIAME).....88

4.3.4 Arabidopsis vs. Brassica array.....89

4.3.5 Microarray Hybridization.....90

4.3.6 Data Analysis.....92

4.4 Results.....93

4.4.1 Comparison of gene expression between resistant and susceptible Brassica accessions (as determined in field trial) of Brassica in the absence and presence of cabbage aphid feeding.....93

4.4.2 Differences in transcriptional response to feeding by cabbage aphid (*B.brassicae*) comparing resistant accessions with susceptible ones based on different parameters of EPG assessment.....98

4.4.2.1 Gene expression analysis between resistant and susceptible accessions based on the non- probing parameter of EPG.....98

4.4.2.2 Gene expression analysis between resistant and susceptible accessions based on the pathway parameter of EPG.....105

4.4.2.3 Gene expression analysis between resistant and susceptible accessions based on the E2 parameter of EPG.....110

4.4.3 Comparison of transcriptional profiles in response to presence or absence of aphid feeding in distinct groups as determined by field assessment and EPG parameters.....117

4.5 Discussion.....118

4.5.1 Arabidopsis gene 1.0 ST array microarrays are applicable to Brassica studies.....119

4.5.2 Transcriptional profile differ in Brassica genotypes in response to aphid feeding depending on grouping criteria.....120

4.6 Conclusion.....127

List of Supplementary tables.....129

CHAPTER 5: Transcriptional response of Brassica wild relatives and landraces to infestation by cabbage aphid (B.brassicae).....132

5.1 Abstract.....133

5.2 Introduction.....134

5.2.1 Chapter Aim.....136

5.3 Material and methods 136

5.4 Results.....136

5.4.1 Determination of significantly differentially expressed genes between CWR and LR using ANOVA.....136

5.4.2 Comparison of gene expression between CWR and LR accessions of Brassica in the absence of cabbage aphid.....137
5.4.2.1 Genes involved in Transport.....138
5.4.2.2 Response to Stress and response to stimulus.....138
5.4.2.3 Genes involved in Signal Transduction and transcription.....139
5.4.3 Comparison of gene expression between resistant CWR and resistant LR accessions of Brassica after 24 hours of Cabbage Aphid feeding.....144
5.4.3.1 Genes involved in Signal Transduction.....156
5.4.3.2 Genes in response to stress.....157
5.4.3.3 Genes in response to stimulus.....157
5.4.4 Comparison of gene expression of CWR and LR in response to presence or absence of aphid feeding.....158

5.5 Discussion.....176

5.5.1 Transcriptional responses of Brassica wild relatives and Landraces in absence and presence of aphid feeding.....176
5.5.2 Crop Wild Relatives and their potential for crop improvement.....178

5.6 Conclusion.....181

List of Supplementary Table.....183

CHAPTER 6: Gene expression arrays as a tool for prediction of insect resistance in Brassica germplasm.....184

6.1 Abstract.....185

6.2 Introduction.....187

6.2.1 Chapter Aim.....190
6.3 Material and methods.....190

6.3.1 Microarray Data.....190
6.3.2 Classification Method.....191

6.4 Results.....194

6.4.1 Building of classification model.....194

6.4.2 Testing and prediction of unknown samples using classification model with 100 selected gene.....195

6.5 Discussion.....212

6.6 Conclusion.....216

List of supplementary tables.....217

Chapter 7: General discussion.....218

7.1 Introduction.....219

7.2 Screening of Brassica germplasm based on field assessment to identify insect resistance.....221

7.3 Intraspecific and interspecific variation in feeding performance of Cabbage Aphid in Brassicas.....222

7.4 Gene expression study: an approach to identify genes of interest.....224

7.5 Impact of domestication: variation in gene expression patterns in crop wild relative (CWR) and landrace (LR) genotypes.....227

7.6 Limitations and future perspective.....229

***CHAPTER 8: References*.....232**

***APPENDICES*.....260**

Appendix I: List of accessions used in the field experiment to explore the level of resistance in the Brassica germplasm..... .260

Appendix II: MIAME/Plant..... 271

List of Tables

Table 2.1: List of Brassica species used in field trial…….**37**

Table 2.2: Analysis of Variance using adjusted SS for aphid infestation (General Linear Model: Aphid infestation versus Accession, block)……**43**

Table 2.3: Analysis of Variance using adjusted SS for whitefly infestation (General Linear Model: white infestation versus Accession, block)……..**44**

Table 2.4: Chi-Square Test of association between aphid occurrence and genotype location…….**45**

Table 2.5: Chi-Square Test of association between aphid and white fly occurrence….46

Table 2.6: List of Brassica CWR and LRs selected from field experiment for further assessment of feeding behavior and changes in gene expression in response to cabbage aphid infestation…….**49**

Table 3.1: List of Brassica CWR and LR genotypes used in EPG study……...**62**

Table 3.2: Comparison of feeding behaviour (mean ± SE of EPG parameters) of Cabbage aphid (*B.brassicae*) 16 different Brassica varieties (Time in Seconds): Mean amount of time recorded via the Electrical Penetration Graph (EPG) technique that aphids spent in different parameters on *Brassica* genotypes (a) NP (non-penetration), (b) Pathway, (c) Xylem ingestion), (d) Potential drop (Pd) and (e) Derailed stylet (F) mechanics over period of 6 hours……**67**

Table 4.1: List of Brassica genotypes: The grouping into resistant (non-preferred) or susceptible (preferred) in response to aphid feeding is based on the previous field experiment (chapter 2) and feeding behavior assessment using EPG study (Chapter 3). NP (non-penetration), Pathway, and E2 (phloem feeding) are Electrical penetration Graph parameters used in grouping. The accession 502202452 was not analyzed for feeding behavior analysis due to insufficient number of replicates…….**87**

Table 4.2: Number of genes showing significant differential gene expression patterns at p-value cut off 0.05 and fold change of ≥ 1.5 (up regulated) or ≤ -1.5 (down regulated) for comparison of accessions assessed as resistant (non-preferred) and susceptible (preferred) in the field trial and feeding behaviour experiments. The gene expression data was recorded for both presence and absence of aphid feeding for 24h. Each accession had 4 replicates……**93**

Table 4.3: List of differentially regulated significant probes in non-induced (absence of aphid feeding) plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the field assessment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation. Each genotype had 4 replicates……..**96**

Table 4.4: List of differentially regulated significant probes in Induced plants (presence of aphid feeding for 24h) when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the field assessment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation. Each genotype had 4 replicates……**97**

Table 4.5: Complete list of differentially regulated significant probes in non-induced (absence of aphid feeding) plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the Non-penetration parameter of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation……**100**

Table 4.6 List of differentially regulated significant probes in Induced (presence of aphid feeding for 24h) plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the Non-penetration parameter of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene

assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation……..**101**

Table 4.7: List of differentially regulated significant probes in Non-Induced (absence of aphid feeding) plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the Pathway parameter of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation……..……...**107**

Table 4.8: List of differentially regulated significant probes in Induced (presence of aphid feeding for 24h) plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the Pathway parameter of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation……**108**

Table 4.9: List of differentially regulated significant probes in Non-Induced (absence of aphid feeding) plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the E2 parameter of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation……...**112**

Table 4.10: List of differentially regulated significant probes in Induced (presence of aphid feeding for 24h) plants when resistant (non-preferred) and susceptible

(preferred) genotypes in response to aphid feeding in the E2 parameter of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation……**114**

Table 5.1 Differentially Expressed genes: Number of genes with significant different in expression at FDR corrected p- Value < 0.05 and fold change ≥1.5 to ≤ -1.5 for comparison of resistant Brassica Crop Wild relative (CWR) and Landraces (LR) in absence of aphid feeding……...**140**

Table 5.2 Differentially Expressed genes: Number of genes with significant different in expression at FDR corrected p- Value < 0.05 and fold change ≥1.5 to ≤ -1.5 for comparison of resistant Brassica Crop Wild relative (CWR) and Landraces (LR) in presence of aphid feeding……**145**

Table 5.3: List of 25 genes common between induced (presence of aphid feeding) and non -induced (absence of aphid feeding) treatment: List of common genes between two comparisons with significant differences in expression at FDR corrected p-value <0.05 and fold change ≥1.5 to ≤-1.5. A positive FC indicates up regulation; negative FC indicates down regulation……..**160**

Table 5.4: List of 236 uniquely expressed induced (presence of aphid feeding) treatment: List of unique genes in presence of aphid feeding with significant differences in expression at FDR corrected p-value <0.05 and fold change ≥1.5 to ≤-1.5. A positive FC indicates up regulation; negative FC indicates down regulation……...**162**

Table 5.5: List of 50 uniquely expressed gene in non -induced (absence of aphid feeding) treatment: List of unique genes with significant differences in expression at FDR corrected p-value <0.05 and fold change ≥1.5 to ≤-1.5. A positive FC indicates up regulation; negative FC indicates down regulation……...**174**

Table 6.1: Details of best Model with the classifiers for prediction resistance status of different Brassica species……...**192**

Table 6.2: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.incana* species using the Partek model selection tool……..**196**

Table 6.3: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.montana* species using the Partek model selection tool……..**199**

Table 6.4: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.villosa* species using the Partek model selection tool……..**201**

Table 6.5: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.fruticulosa* species using the Partek model selection tool……..**206**

Table 6.6: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.oleracea acephala* species using the Partek model selection tool……...207

Table 6.7: Details of Model (individual model for each species) run with the classifiers for prediction resistance status of different Brassica species……**213**

List of figures

Figure1.1: Rich morphotypes of *Brassica* plants. The picture shows different morphotypes of (a) *B.rapa*; (b) *B.oleracea*. The figure was adapted from (Cheng et al. 2014)……**4**

Figure 1.2: The triangle of U: diagram showing the genetic relationship between the six species of the genus *Brassica*. Chromosomes from each of the genomes A, B and C are represented by different colours (Diagram adapted from Wikipedia)...**6**

Figure 1.3: Cabbage aphids, *Brevicoryne brassicae* Linnaeus: A) on cabbage leaves. B) Cabbage aphid colony on a cabbage leaf. C) Cabbage aphid on a cabbage stem Photograph adapted from (Gill et.al.2013).....**10**

Figure 1.4: General life cycle of Aphid: During the spring and summer reproduction is by parthenogenesis. Asexually-produced embryos develop in approximately 10 days. In the fall there is a single generation of sexually-reproducing males and females which produce diapausing eggs. These eggs do not hatch until the following spring, more than 100 days later. The image is adapted from. http://www.ludwigsroses.co.za/rose-care/aphids-2.....**11**

Figure 1.5: Diagram of aphid mouthparts showing A: lateral view of head with labarum detached from front of stylet. B: Cross-section of stylet. (Image from The University of Arizona, Centre for Insect Science Education Outreach http://insected.arizona.edu/gg/resource/internal.html). C: Shows diagrammatic representation of aphid feeding and phloem location.....**13**

Figure 1.6: EPG technique and stylet movement across the plant tissue (A) Represents EPG setup where aphid and plant are incorporated in an electrical circuit. As soon as the aphid starts feeding the waveforms are generated and recorded. (B) Represents the probing behaviour, the movement of stylet inside tissue in order to locate phloem sap. [the figure is adapted from (Schwarzkopf et al. 2013).....**17**

Figure 1.7: Diagram showing the EPG waveforms and correlations to aphid feeding behaviours, as defined in the 'Probe 3.4 Manual: Software manual for EPG acquisition and analysis in Windows' distributed by EPG systems, Wageningen the Netherlands.....**21**

Figure 2.1: The position of each accession in the field. Each cell represents 4 plants of the same accession; X refers to accessions missing on day of scoring.....**39**

Figure 2.2: Field trial in Stratton Audley near Bicester. (A) Shows plants at 8 weeks in the field, dead plants as a result of flea beetle attack (B-C).....**40**

Figure 2.3: Trial field in Stratton Audley near Bicester (A), 15 week old cabbage (B), whitefly infestation (C) and plant highly infested by cabbage aphid (D).....**41**

Figure2.4: Geographical location of each accession in the field. Each cell represents 4 plants of same accession; Different colours of each cell represent infestation by insect pests. Green cells represent aphid infestation; yellow represents whitefly infestation and red represent plants where both aphid and whitefly were present.....**42**

Figure 3.1: Schematic overview of the plant-aphid interface: Aphid mouthparts penetrate the leaf surface upon encountering preformed defences such as trichomes and waxes. The figure was adapted from (Jaouannet et al. 2014).....**56**

Figure 3.2: Diagram represents a typical EPG setup: The aphid is incorporated into an electrical circuit, with one electrode being attached to the aphid, while the other is inserted into the soil of its host plant's pot. When the aphid penetrates the leaf with its stylets, it completes the circuit. Output EPG waveforms are recorded and vary with the aphid's behaviour and the position of the stylet tip (Tjallingii & Gabryś 1999)……**64**

Figure 3.3: Mean amount of time (*hour*s) recorded via the Electrical Penetration Graph (EPG) technique that aphids spent in different parameters on *Brassica* genotypes (a) NP (non-penetration), (b) Pathway, (c) Xylem ingestion), (d) Potential drop (Pd) and (e) Derailed stylet (F) mechanics over period of 6 hours. Aphids species used for study was Brevicoryne brassicae over a six hour period. Plants were 12 weeks old. Aphids were removed from culture and wired straight up to the EPG system. Statistical analysis of this data can be found in table 3.2. N = 10 per genotype. Mean time spent is represented on vertical axis and Brassica genotypes are represented on horizontal axis. The data was recorded from first connection of aphid stylet with plant till end of recording period of 6hours. The data is provided in table 3.2. The Kruskal-Wallis showed significant difference ($P \leq 0.001$) among genotypes for these parameters.....**69**

Figure 3.4: Comparison of EPG Parameters: Mean time spent in each EPG parameter for 5 Brassica species over period of 6 hours. Plants were 12 weeks old. Aphids were removed from culture and wired straight up to the EPG system. Statistical analysis of this data can be found in table 3.2. N = 10 per genotype. Mean time spent is represented on vertical axis and Brassica genotypes are represented on horizontal axis. The data was recorded from first connection of aphid stylet with plant till end of recording period of 6hours. The data is provided in table 3.2. The Mann-Whitney U test showed significant difference ($P \leq 0.001$) among species for these parameters.....**70**

Figure 3.5: Mean time in E2 and time to E2: Mean duration of time taken by cabbage aphid to reach phloem and average time spent feeding on different genotypes. The genotypes vary significantly in response to aphid feeding on E2 and the time take to reach phloem ($P \leq 0.001$). Plants were 12 weeks old. Aphids were removed from culture and wired straight up to the EPG system. Statistical analysis of this data can be found in table 3.2. N = 10 per genotype. Mean time spent is represented on vertical axis and Brassica genotypes are represented on horizontal axis. The data was recorded from first connection of aphid stylet with plant till end of recording period of 6hours.....**71**

Figure 4.1: Schematic workflow for Affymetrix GeneChip one colour Microarray based gene expression experiment.....**90**

Figure 4.2: Comparison of number of genes, in response to aphid feeding in different groups when resistant (non-preferred) Brassica genotypes were compared with susceptible (preferred) genotypes. (A) Venn diagram representing the distribution of genes in absence of aphid feeding. (B) Venn diagram representing the distribution of genes in presence of aphid feeding for 24 h. The numbers in the overlapping area indicate the shared number of genes in the comparisons and include genes with an average expression ratio ≥ 1.5 -fold and a P value < 0.05 in both experiments. Numbers outside the overlapping area represent genes only in that group.....**117**

Figure 5.1: Venn diagram: Venn diagram shows the number of significantly differentially expressed among CWR and LR in presence and absence of aphid feeding.....**158**

List of supplementary table

Supplementary data is available on the CD included at the back of this thesis

Table S2.1: Scoring Data for the aphid and white fly infestation. The value 0 to 4 (where 0 for highly resistant and 4 for susceptible) represents the level of resistance present in Brassica plant against insect pest. X represents missing plants. The accessions used in further study are highlighted in yellow

Table S4.1: List of differentially regulated significant probes in **non-induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **field assessment** were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation. Each genotype had 4 replicates

Table S4.2: Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **field assessment** in **non-induced** (absence of aphid feeding) plants.

Table S4.3 List of GO terms under biological process for **non-induced**, **field assessment** for comparison between resistant and susceptible accessions using TAIR's Gene Ontology annotation search and functional categorization web tool analysis.

Table S4.4: List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Non-penetration parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

Table S4.5 Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **Non-penetration parameter** of EPG assessment in **induced** (presence of aphid feeding).

Table S4.6 List of GO terms under biological process for **induced**, **non-penetration parameter of EPG** for comparison between resistant and susceptible accessions using TAIR's Gene Ontology annotation search and functional categorization web tool analysis.

Table S4.7: List of differentially regulated significant probes in **Non-Induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Pathway parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

Table S4.8 Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **Pathway parameter** of EPG assessment in **non-induced** (absence of aphid feeding) plants.

Table S4.9: GO Enrichment analysis of significant differentially expressed genes between resistant and susceptible accession based on **pathway parameter** of EPG assessment in **non-induced** (absence of aphid feeding) plants

Table S4.10: List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Pathway parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation

Table S4.11: Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **Pathway parameter** of EPG assessment in **induced** (presence of aphid feeding)

Table S4.12: GO Enrichment analysis of significant differentially expressed genes between resistant and susceptible accession based on **pathway parameter** of EPG assessment in induced (presence of aphid feeding) plants.

Table S4.13: List of differentially regulated significant probes in **Non-Induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **E2 parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

Table S4.14 Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **E2 parameter** of EPG assessment in **non-induced** (absence of aphid feeding) plants.

Table S4.15: GO Enrichment analysis of significant differentially expressed genes between resistant and susceptible accession based on **E2 parameter** of EPG assessment in **non-induced** (absence of aphid feeding) plants.

Table S4.16: List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **E2 parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

Table S4.17 Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **E2 parameter** of EPG assessment in **induced** (presence of aphid feeding) plants.

Table S4.18: GO Enrichment analysis of significant differentially expressed genes between resistant and susceptible accession based on **E2 parameter** of EPG assessment in **induced** (presence of aphid feeding) plants.

Table S5.1: Differentially Expressed genes: Number of genes with significant different in expression at FDR corrected p- Value < 0.05 and fold change ≥1.5 to ≤ -1.5 for comparison of resistant Brassica Crop Wild relative (CWR) and Landraces (LR) in absence of aphid feeding.

Table S5.2: Functional characterization of significant differentially expressed genes between Brassica CWR and LR in absence of aphid feeding.

Table S5.3: List of GO terms under biological process for constitutive differentially expressed significant genes for comparison between CWR and LR using TAIR's Gene Ontology annotation search and functional categorization web tool analysis.

Table S5.4: Differentially Expressed genes: Number of genes with significant different in expression at FDR corrected p- Value < 0.05 and fold change ≥1.5 to ≤ -1.5 for comparison of resistant Brassica Crop Wild relative (CWR) and Landraces (LR) in presence of aphid feeding.

Table S5.5: Functional characterization of significant differentially expressed genes between resistant CWR and resistant LR of Brassica after 24 hours of aphid feeding.

Table S5.6: List of GO terms under biological process for induced (after 24 hours of aphid feeding) differentially expressed significant genes for comparison between CWR and LR using TAIR's Gene Ontology annotation search and functional categorization web tool analysis.

Table S6.1: List of 100 selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of B.incana species using the Partek model selection tool.

Table S6.2: List of 100 selected genes used to build the prediction model o test suitability to aphid feeding in two genotypes of B.montana species using the Partek model selection tool.

Table S6.3: List of 100 selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of B.villosa species using the Partek model selection tool.

Table S6.4: List of 100 selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of B.fruticulosa species using the Partek model selection tool.

Table S6.5: List of 100 selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of B.oleracea acephala species using the Partek model selection tool.

List of abbreviations

AA - Amino Acid

ABA - Abscisic Acid

ANOVA - Analysis of Variance

ATP - Adenosine Triphosphate

BAC - Bacterial Artificial Chromosome

BLAST - Basic Local Alignment Search Tool

CWR - Crop wild relative

E2 - phloem phase

EPG - Electrical Penetration Graph

ER - Endoplasmic Reticulum

EST - Expressed Sequence Tag

ET - Ethylene

FC - Fold Change

GO analysis – Gene Ontology Analysis

JA - Jasmonic Acid

LR - Landrace

LRR - Leucine Rich Repeat

MAPK - Mitogen Activated Protein Kinase

MIAME - Minimum Information about a Microarray Experiment

MGED - Microarray Gene Expression Database group

NBS - Nucleotide Binding Site

NCAP - Non-Cell-Autonomous Protein

NCBI – National Center for biotechnology Information

NGS - Next-Generation Sequencing

NO - Nitric oxide

NP- Non-penetration

PCA - Principle Component Analysis

PD - Plasmodesma (ta)

PFI - Phloem Feeding Insect

PM - Plasma Membrane

RIN – RNA integrity number

CHAPTER 1

Review of literature

&

general introduction

1.1 The experimental plant: Brassicas

1.1.1 Importance of Brassica: economic, social and medicinal

The Genus *Brassica* belongs to the mustard family, formerly known as Cruciferae or crucifer family. This group includes a wide range of plants including Broccoli, Brussels sprouts, Cabbage, Cauliflower, turnips, Kale, Mustard etc (Rich 1991)

Taxonomy of Brassica

Kingdom	*Plantae*
Subkingdom	*Tracheobionta*
Division	Angiosperms
Class	Dicotyledons
Order	*Brassicales*
Family	*Brassicaceae* (Mustard Family)
Genus	*Brassica*

The crop plants included in this family have a wide range of uses and a large amount of interspecific and intraspecific morphological phenotypic variation related to high levels of genetic variation (Lowe et al. 2004). While this genus contains many crop species, it also contains numerous related crop wild relatives (CWR). CWR are a plant genetic resource (PGR) whose largely untapped pool of genetic diversity can be utilised by the plant breeding community to breed novel varieties, with improved biotic and abiotic

stress tolerance, improved yield and reliance on fewer resource inputs, into modern crop varieties (Maxted and Kell, 2008; FAO, 2009). Nearly all crops have related crop wild species, which are a rich source of genes for the creation of new cultivars (Westman & Kresovich 1999). While introgression of genes from CWR has been occurring from the very beginnings of agriculture, the commercial use of CWR genes by plant breeders can be traced back only as far as the late 19th century (Hajjar & Hodgkin 2007). By the 1980s, the use of CWR as sources of novel traits in commercial breeding and the need to conserve the genetic variation they represent was widely recognised. In a review regarding the contribution of CWR, it was estimated that their contribution to commercial breeding in 1988 attributable to worth ~350 million US dollars a year (Prescott-Allen & Prescott-Allen 1988). This was due to the numerous successful breeding programmes that used CWRs to introduce new and valuable traits into existing cultivars. The most common traits introduced were and still are those that confer some kind of pest or disease resistance.

The mustard family is included among the ten most important crops to humans and especially genus *Brassica* contains most economically important species (Westman & Kresovich 1999). In 2006 approximately 29 million tonnes of *B.napus*, 69 million tonnes of cabbage and 80 million tonnes of cauliflower were produced (FAO, 2008). The members of the Brassica family like cabbage, cauliflower, broccoli, brussels sprouts are important vegetables while *B.napus* and *B.nigra* are important sources of edible oil. In addition, *B.napus* and *B.carinata* have also been explored for their non-food use as biofuels (Cardone et al. 2003).

Brassica crops also include a major group of vegetables, which make them internationally important. *Brassica* plants have rich diversity with respect to both

speciation and the abundant morphotypes in each *Brassica* species (Figure 1.1). It is a staple food in many countries. *Brassica* is also socially and medicinally important. It is used in national dishes such as sauerkraut (Germany) and kimchi (Korea) and also Brussels sprouts at Christmas time in the UK.

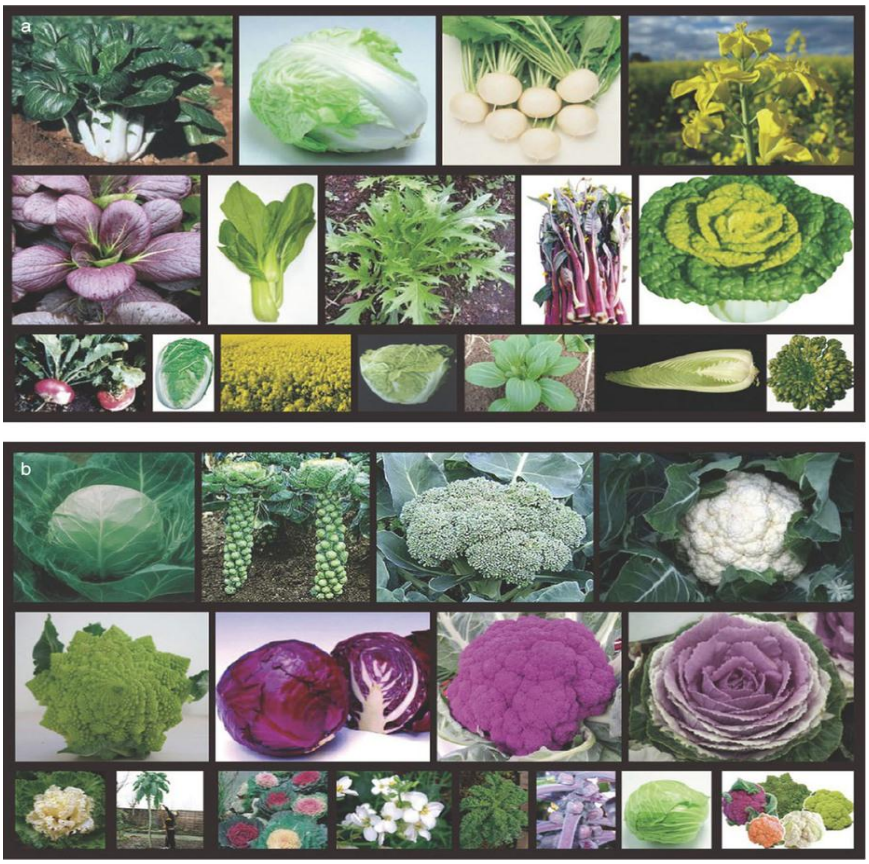


Figure 1.1: Rich morphotypes of *Brassica* plants. The picture shows different morphotypes of (**a**) *B. rapa*; (b) *B. Oleracea*. The figure was adapted from (Cheng et al. 2014)

There have been studies where the medicinal properties of Brassicas have been listed, particularly in protection against heart disease and cancer (van Poppel et al. 1999; Finley 2003; Willcox et al. 2003). Diets rich in broccoli have been associated with a reduction in risk of progression of prostate cancer progression (Traka et al. 2014). A

number of epidemiological studies have identified an inverse association between consumption of Brassica vegetables and the risk of gastric cancer (Ekström et al. 2011; Moy et al. 2009), lung cancer(London et al. 2000), bladder cancer(Zhao et al. 2007) and myocardial infarction (Cornelis et al. 2007). The genes from wild *B.oleracea* have been successfully used to breed increased levels of anti-cancer compounds into modern broccoli cultivars (Hajjar & Hodgkin 2007). In addition to this, it is also reported to be effective against declines associated with ageing and associated disorders like Alzheimer's and cataracts (Granado et al. 2003).

1.1.2 *Brassica* genetics

Brassicaceae is one of the most diverse plant families, comprising 49 tribes, 321 genera, and over 3660 species (Al-Shehbaz 2012). The *Brassica* genus, owing to its remarkable species, genetic, and physiological diversity as well as its significant economic potential, has become a model for polyploidy and evolutionary studies (Kagale et al. 2014). The genus *Brassica* is typical of many crop species in having a larger and more complex genome than other model plants. The genomic relationships are well characterised, as shown in the 'triangle of U' (Figure 1.2). The genomes have been denoted as A, B and C with three monogenomic diploid species, *B.rapa* (AA, n=10), *B.nigra* (BB, n=8) and *B.oleracea* (CC,n=9) and there are three allotetraploid species derived from each pair of the three diploid species, *B. juncea* (AABB, *n*=18), *B. napus (AACC,n=19)* and *B. carinata*, BBCC, *n*=17). The genetic relationships of these species were identified and confirmed by extensive experimental crosses between tetraploid and/or diploid plants as well as karyotyping or microscopic inspection at the synapsis stage of meiosis in these crosses (Nagaharu U 1935).

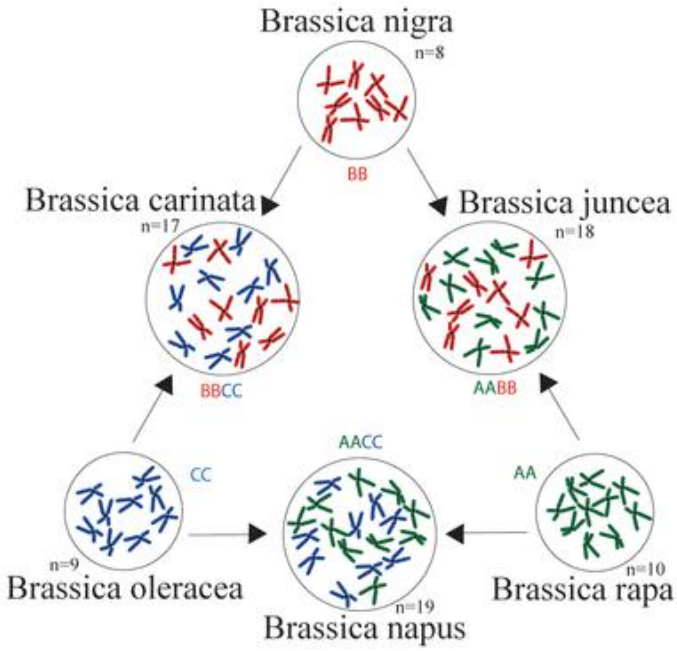


Figure 1.2: **The triangle of U**: diagram showing the genetic relationship between the six species of the genus *Brassica*. Chromosomes from each of the genomes A, B and C are represented by different colours (Diagram adapted from Wikipedia)

These species have been exploited to understand the basis of chromosome evolution since divergence from a common progenitor shared with *Arabidopsis*. This has allowed the *Brassica* genus to be a valuable tool for the study of chromosome evolution, in particular that associated with speciation and polyploidization (King 2005; Kagale et al. 2014).

Comparative physical mapping studies have confirmed genome triplication in a common ancestor of *B. oleracea* (O'Neill & Bancroft 2000) and *B.rapa* (Park et al. 2005) since its divergence from the *A. thaliana* lineage at least 13–17 MYA. The availability of genetic maps has allowed comparative mapping of three genomes (Lagercrantz & Lydiate 1996) to show that the genetic content of the species is highly conserved despite differences in chromosome number. In *Brassica*, early attempts (Hu

et al. 1998) to align linkage maps derived from different *Brassica* populations were based on very low numbers of shared markers, and suffered from a lack of resolution with respect to distinguishing between paralogous loci. More recent efforts have been successful in generating aligned maps for the *Brassica* A genome that integrate marker information using a common set of SSRs scored in *B.rapa* and *B.napu*s. Comparison with the closely related species *Arabidopsis thaliana*, whose genome has been sequenced, provides many possibilities for identification of homologous genes between *Arabidopsis* and *Brassica* spp. (Sadowski et al. 1996; Cheng et al. 2014). Genome datasets of the *Brassica* species are maintained and continuously updated within the *Brassica* database (http://brassicadb.org) (Cheng et al. 2011). The genome sequences of *B.oleracea* (Liu et al. 2014), *B.napus* (Chalhoub et al. 2014) and *B.rapa* (Wang et al. 2011) are now available online which will facilitate further research of Brassica crops and breed new improved Brassica crop varieties. Functional traits of many *Brassica* CWR have been identified and many have already been transferred to commercially cultivated varieties. In particular resistance to black leg fungus (*Leptosphaeria maculans)* was introduced into *B. napus* via hybridisation with *B. juncea*, *B.carinata* and *B. nigra* (Zhu & Struss 1991; Starzycki et al. 1996).

1.1.3 Crop Wild Relatives and loss of resistance during domestication

Crop wild relatives (CWR) are species which include crop progenitors and related species to them and may contribute to beneficial traits for crop improvement (Maxted et al. 2008). They are closely related to the domesticated plants and may possess traits like drought tolerance, salt tolerance and pest resistance.

It is a well-established fact that CWR can contribute to crop improvement with examples dating back more than 60 years (Hajjar & Hodgkin 2007). Prescott-Allen & Prescott-Allen (1988) reviewed the importance and contribution of CWR genes to improve cultivars. The genus *Brassica* includes many important vegetable species in addition to CWR like *B.villosa*, *B.incana* and *B.fruticulosa* which are thought to possess important genes for insect resistance. During the process of domestication of wild species into a cultivated crop as a result of selection process traits like fruit size, high yield, reduced dormancy, were preferred (Evans LT 1996). Domestication is believed to have reduced herbivore resistance when compared to wild species for many non-exclusive reasons (Turcotte et al. 2014). There are several reasons responsible for this reduction in resistance, for example reduction in levels of alkaloids in domesticated sweet lupin (*L.albus*) which suffers from severe herbivory as compared to wild lupins, which contain high levels of toxic alkaloids (Wink 1988), and similarly improvement in nutritive quality of certain crops resulted in promotion of herbivore growth and fitness on these crops (García-Palacios et al. 2013). The predicted resource allocation tradeoffs between growth and defence are also considered as a reason for decreased resistance against herbivores (Turcotte et al. 2014). In addition, it is also reported that the process of inbreeding can result in the decline in the plant's defence (Portman et al. 2015). Alteration in levels of insect resistance and host plant quality are reported in yellow monkey flower (*Mimulus guttatus*) (Carr & Eubanks 2002). Wild maize species (teosinte) have greater resistance to insect pests (de Lange et al. 2014). The domesticated perennial accessions of chickpeas are reported to have suppressed levels of resistance against larvae of cotton bollworm (*Helicoverpa armigera)* as compared to its wild relatives *C.microphyllum, C.canariense* (Chaudhary 2013). There is no doubt

about the fact that domestication of plants is one of the most important advances in human history (Turcotte et al. 2014) but at the same time it resulted in making our crops more susceptible to insects. This also hints towards the presence of insect resistance genes in CWR which may be used by breeders to develop more resistant varieties.

1.2 Insect pests

1.2.1 Importance of Insect Pests in agriculture

Insect pests cause huge damage to agricultural crops. The total crop production loss due to insect pest damage is reported to be 14-25 % (Devilliers & Hoisington 2011). The damage caused by insects to crop is directly by feeding on them or indirectly by transmission of plant viruses. It is not just the yield reduction due to insect pests; there are other related factors as well which affect the overall production cost like the use of pesticides (including insecticides, herbicides) for crop protection against the insect pest. Worldwide about 3 billion kg of pesticides is used every year costing around $40 billion per year ((Pimentel 2005)) In addition, (Pimentel 2005) reported $9.6 billion used in major economic and environmental losses due to the application of pesticides. In general, pesticides are considered profitable in agriculture; their use does not always decrease crop loss. Although the application of insecticides is the preferred way of dealing with the insect problem, but this way of pest control is not only hazardous to the environment, it also affects humans as the chemicals easily get incorporated into the food chain. Additionally, constant use of insecticides can result in the development of resistance or insensitivity against them. The green peach aphid, *Myzus persicae*, has developed resistance to at least seventy different synthetic compounds, and different insecticide resistance mechanisms have been reported worldwide (Silva et al. 2012). In a case study from Pakistan, its reported that melon aphid *Aphis gossypii* Glover

(Hemiptera: Aphididae) has developed a broad-spectrum resistance to insecticides (Ahmad et al. 2007) and also complaints of chemical control failure against *B. Brassicae* (Ahmad & Akhtar 2013) . Due to all these factors, it has raised the need to find the alternative methods to protect our crops. A possible, environmentally friendly solution is the use of host plant resistance and development of new more resistant crop varieties.

The major insect pests of *Brassica* crops are aphids, whiteflies, beetles, caterpillars, and grasshoppers. Beetles, bugs and caterpillars usually have chewing mouth parts and damage is seen as chewed leaves and stems. Other insects like aphids and whiteflies have developed more specialised feeding mechanisms. They feed with the help of piercing mouthparts, namely stylets to feed on plants (Pollard 1973; Tjallingii & Esch 1993; Walling 2000).

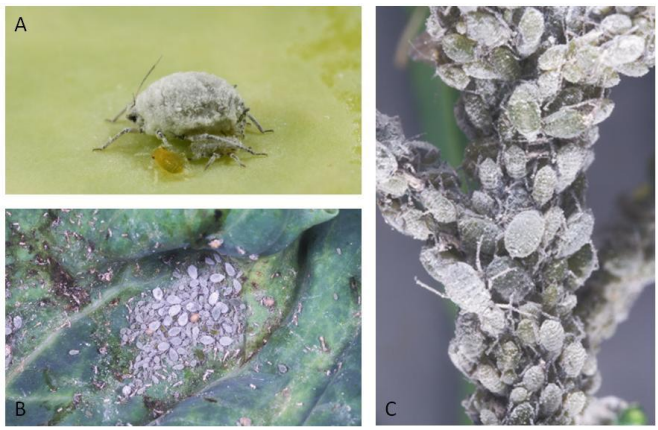


Figure 1.3: Cabbage aphids, *Brevicoryne brassicae* Linnaeus: A) on cabbage leaves. B) Cabbage aphid colony on a cabbage leaf. C) Cabbage aphid on a cabbage stem Photograph adapted from (Gill et.al.2013).

1.2.2 Cabbage aphid (*Brevicoryne brassicae*)

The cabbage aphid is a member of the genus *Brevicoryne*. It is native to Europe with worldwide distribution (Kessing JLM & RFL 1991) Cabbage aphid is greyish-white in

appearance with a waxy covering. They have short tube-like structures (cornicles) at the tip of the abdomen and feed on the stem and underside of leaves (Hines & Hutchison 2013). Adults are present in both wingless and winged form. However, wingless females producing live young (nymphs) are the most common. It is a *Brassica* specialist species which causes significant damage to Brassica crops like cabbage, cauliflower, broccoli, mustard etc. Gabryś & Pawluk (1999) reported the restricted host range of cabbage aphid due to the requirement of *sinigrin* (a glucosinolate in *Brassica* plants) to initiate feeding and spending the whole life cycle on *Brassicacea* family.

1.2.3 Life cycle and reproduction

Aphids have a highly adaptable life cycle. They can reproduce both sexually and asexually depending upon the environmental conditions. In warm climates, nymphs are produced by females without mating. The whole colony population consists of females only (Gill et al. 2013).

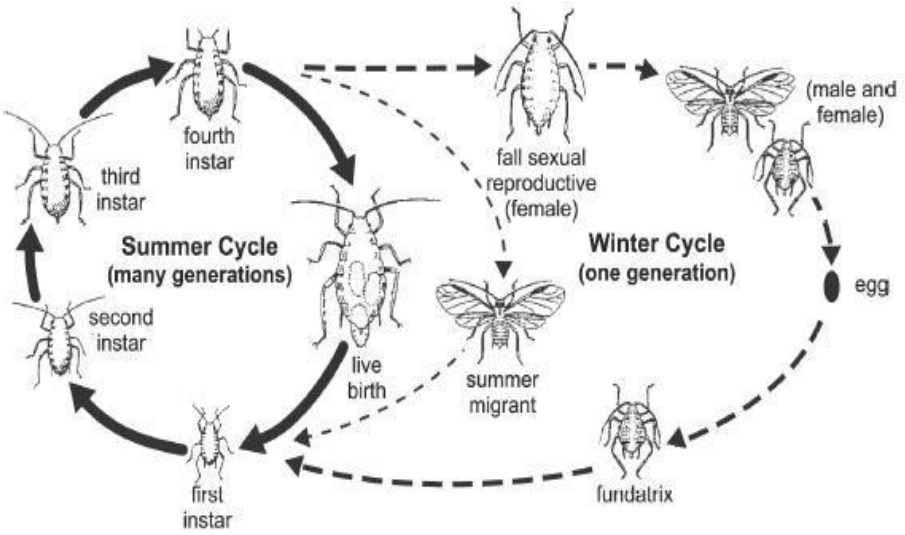


Figure 1.4: General life cycle of Aphid: During the spring and summer reproduction is by parthenogenesis. Asexually-produced embryos develop in approximately 10 days. In the fall there is a single generation of sexually-reproducing males and females which produce diapausing eggs. These eggs do not hatch until the following spring, more than 100 days later. The image is adapted from. http://www.ludwigsroses.co.za/rose-care/aphids-2

However prior to winter season the mode of reproduction changes in response to low temperature and decreased photoperiod and males are also produced (Blackman & Eastop 2000). Eggs are laid after mating. *B.brassicae* overwinters as eggs on the Brassica crops and hatch during spring (February to April).. The young alate (winged) forms migrate to new host plants when the plant condition deteriorates as a result of large population growth. Aphids have an evolutionary advantage of rapidly increasing population size and as a result, they can have up to 15 generations during a crop season (Hines & Hutchison 2013).

1.2.4 Location and acceptance of host plant

In order to locate new host plants, the winged aphids (alates) use visual cues to locate plants, with most species displaying a preference to land on yellow surfaces (Prokopy & Owens 1983) whilst the polyphagous species *R. padi* and *A. fabae*, have been found to be more responsive to green wavelengths (Hardie 1989; Nottingham et al. 1991). In addition to the visual clues they also use chemical and mechanical signals to decide whether to stay. Aphids can detect plant volatiles such as nitriles, green leaf volatiles, benzaldehydes, isothiocyanates and monoterpenes (Visser & Piron 1997). Mechanical information about a plant, such as leaf waxiness, hairiness, glandular trichome absence or presence, and epidermal thickness also be used to deciding whether to attempt probing (Rehman & Powell 2010; Powell et al. 2006; Dinant et al. 2010). After landing on the potential host, they begin to probe with the stylets, providing information for host acceptance or rejection(Powell & Hardie 2000).

1.2.5 Feeding mechanism and avoiding host defences

Aphids have developed highly evolved feeding strategies. They have special elongated mouthparts called stylet (figure 1.5) to pierce through the plant tissue and reach the

phloem sap (Pollard 1973). During feeding, the aphid produces sheath saliva which is rapidly gelling, and composed of proteins, carbohydrates and phospholipids (Will et al. 2012).

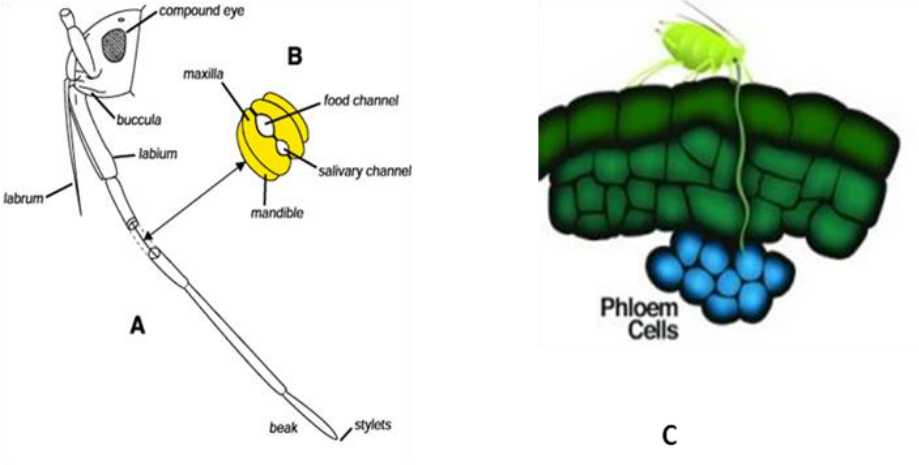


Figure 1.5**: Diagram of aphid mouthparts showing** A: lateral view of head with labarum detached from front of stylet. B: Cross-section of stylet. Image from The University of Arizona, Centre for Insect Science Education Outreach (http://insected.arizona.edu/gg/resource/internal.html). C: Shows diagrammatic representation of aphid feeding and phloem location.

The sheath saliva rapidly gels around the stylet at the feeding site and acts as a protective barrier from the plant defence response (Bhatia et al. 2011). In addition to sheath saliva aphids also introduce watery digestive saliva with a complex mixture of enzymes and other compounds which in the presence of oxidases inactivate defensive phytochemicals released in response to damage and wounding by plants (Miles 1999). The probing of tissue for locating food is often influenced by changes in the chemical content of sap and physiological changes induced by introduction of aphid saliva (Will et al. 2012). The piercing stylet causes a minimal damage to tissue and presence of calcium-binding proteins in aphid saliva make it difficult for the plant to perceive the

damage and can result in a very delayed plant defence response (Will et al. 2007). The aphids usually deceive the host and suppress its effective defences which help them to colonize and successfully feed on host plants (Walling 2008).

1.3 Aphid-plant interaction: plant defence strategies

1.3.1 Insect resistance in plants

In nature, plants are constantly exposed to biotic and abiotic stress. Plants are challenged by insects, pathogens and diseases but they still manage to survive. The reason behind plant success is presence of a broad array of plant defences, ranging from constitutively expressed to induced defences (War et al. 2012) Plants have evolved different defence strategies with some adapting to escape in space and time by growing in inaccessible locations or adapting growing seasons to avoid herbivores (Howe & Jander 2008). For plants where this is not possible, tolerance is the defensive tool. The adaptation strategy of tolerance or avoidance acts on plants as opposed to herbivores whereas antibiosis and antixenosis are plant defence strategies which directly affect herbivores (Goggin et al. 2015; Smith and Clement 2012). If the plant has antibiosis resistance it means that it will affect the herbivore survival by reducing longevity, fecundity and increased mortality. In comparison, antixenosis resistance plants own characters usually act as a deterrent towards a preference for that plant.

Both antibiosis and antixenosis are widespread defence strategies used against aphid against aphid herbivores (Walling 2008). These can be broadly classified as constitutive defence, which is always present in the plant even before the herbivore attack or damage, or induced defence, which is triggered in response to the herbivore attack (Kessler & Baldwin 2002; Schoonhoven et al. 2005).The constitutive defence can be due to the physical or chemical barriers present in plants. Physical barriers include

thorns, trichomes, thick cuticles and waxes covering leaves etc. They act as the first line of defence against any herbivore attack. In addition to these physical barriers, plants also produce certain chemical substances like latex, oil, resins, sticky compounds which are produced and stored in plants during growth and development process. These chemical substances function in response to membrane disruptions, signal transduction, and wound response (Wittstock & Gershenzon 2002; Bennett & Wallsgrove 1994).

The induced defence is triggered primarily after the herbivorous insect starts feeding. It can be localised at the feeding site or systemic i.e. in the distal undamaged tissue of plant (Santamaria et al. 2013). Induced defence activates various cell signalling pathways and results in transcriptional and metabolic changes and production of toxic or deterrent compound (Kuśnierczyk et al. 2008). In addition, signal transduction pathways like salicylic acid (SA), jasmonic acid (JA) and ethylene are activated in response to *insect –plant* interaction (Thompson & Goggin 2006; Moran & Thompson 2001).

Brassicas have a highly developed glucosinolate-myrosinase defence mechanism which protects them against the insect pests like aphids (War et al. 2012; Ahuja et al. 2009). When aphid feeding is initiated and cells are disrupted, glucosinolates are hydrolyzed by myrosinases and bioactive compounds like isothiocyanates, thiocyanates, nitriles, epithionitriles are formed (Elzinga et al. 2014; De Vos et al. 2007; Bones & Rossiter 2006). But some specialist insects like whitefly and diamond black moth have evolved specialist enzyme systems which detoxify the effect of glucosinolates. The *Brassica* specialist cabbage aphid hydrolyse glucosinolates with the help of its myrosinase and use the active product as a defence against natural enemies (Jones et al. 2002; Kazana et al. 2007) and also accumulate glucosinolates for self-defence (Schoonhoven et al.

2005). Plant-aphid interaction is a highly studied topic (Jaouannet et al. 2014; Broekgaarden et al. 2008) but still a very limited knowledge is available. So it becomes very important to further explore the defence mechanisms to get a better understanding. In addition, all different defence mechanisms present the plants to defend themselves against insect pest, plants also offer different levels of resistance by presenting variable interspecific and intraspecific variations which are among many other factors discussed in the previous section.

1.4 Electrical Penetration Graph (EPG) technique: monitoring aphid feeding behaviour

The monitoring of feeding behaviour is important as it is an effective way of assessing the level of feeding suitability in the plants. It is visually performed by using a microscope, with or without video recording (Kindt 2004). This mode of studying feeding behaviour is not very effective for studying sap sucking insects as we cannot get accurate information regarding activities happening inside the cell. To overcome this issue, Electrical Penetration Graph (EPG) technique was developed (Tjallingii 1985) which proved to be one of the most successful and is utilised in studying insect-plant interaction. This provides useful and vital information regarding the presence and location of resistance factors in plants (Tjallingii, W. F. and Mayoral 1992).

EPG is an excellent method to investigate the interaction present between insect and host during the process of feeding (Alvarez et al. 2006; Powell & Hardie 2000; Gabryś & Pawluk 1999). The EPG involves an aphid and the plant to be made part of an integrated electrical circuit (Tjallingii 1985) as shown in figure 1.2. When the aphid inserts its stylet into the plant tissue, the circuit is closed and current (DC) flows.

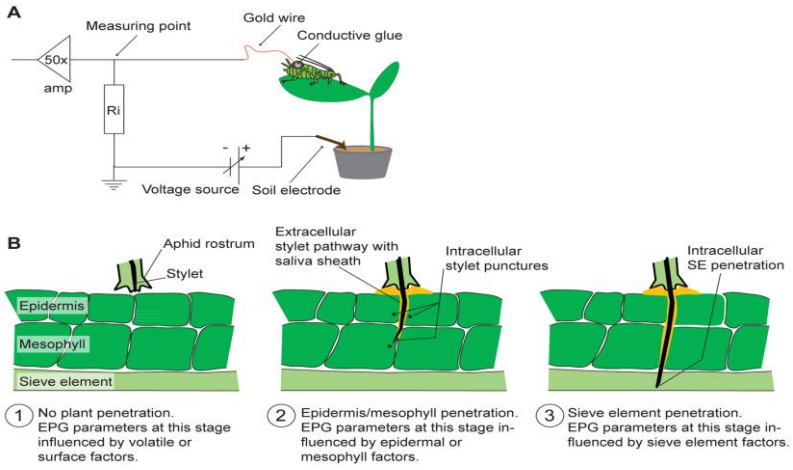


Figure 1.6: EPG technique and stylet movement across the plant tissue (A) Represents EPG setup where aphid and plant are incorporated in an electrical circuit. As soon as the aphid starts feeding the waveforms are generated and recorded. (B) Represents the probing behaviour, the movement of stylet inside tissue in order to locate phloem sap. [the figure is adapted from (Schwarzkopf et al. 2013).

A change in feeding behaviour generates the electric fluctuation, recorded as waveforms of specific patterns. The EPG technique has been extensively used to study the variation in resistance levels between different species (Cole 1994; Prado & Tjallingii 1999; Prado & Tjallingii 1997; Poelman et al. 2008).

Ghaffar et al. (2011) reported the variation in the resistance level against Brown Plant Hopper in rice germplasm. The genetic variability in resistance to aphid among peaches (*Prunus persica* (L.)Batsch) and nectarine (*Prunus persica* variety *nectarina*) has been evaluated using feeding behaviour assessment by EPG (Verdugo et al. 2012). Broekgaarden et al. (2012) reported the presence of phloem specific resistance in *B.oleracea* against whitefly *Aleyrodes proletella* based on feeding behaviour

assessment. There are countless examples where the use of EPG proved helpful in enhancing the insect-plant interaction knowledge.

1.4.1 EPG waveforms

EPG was first tested and used for studying the feeding of aphids. Then, it has been applied to a wide range of other piercing insects such as Brown plant hopper (Ghaffar et al. 2011), whiteflies (Janssen et al. 1989; Lei et al. 1998), thrips (Hunter et al.1993; Harrewinji et al. 1996), leaf –plant hoppers (Backus and Hunter, 1989; Lett et al., 2001) and also mealy bugs (Calatayud et al. 1994). Because the feeding behaviour of these insects is different, so will be the waveforms generated by them. For aphids, the behaviour associated with these waveforms has been already described (Tjallingii 1985). Broadly waveforms can be classified into 7 distinct patterns (fig1.6).

1.4.2 EPG waveform description

All the waveforms used for characterizing aphid feeding were as described by Tjallingi (1985) and shown in figure 1.6. The patterns identified were:

1. Mean duration of non-penetration period (NP)
2. Period of pathway waveform pattern (C).
3. Period of potential drop (Pd)
4. Period of salivation into sieve element (E1).
5. Period of sustained phloem sap ingestion (E2).
6. Period of derailed stylet mechanics (F)
7. Period of ingestion of xylem sap (G)

For aphids, EPG waveforms have already been well defined (Prado & Tjallingii 1994) and commonly used as a reference for other sap feeding insects. The first type namely non-penetration phase (NP) is represented by a straight line indicating no probing

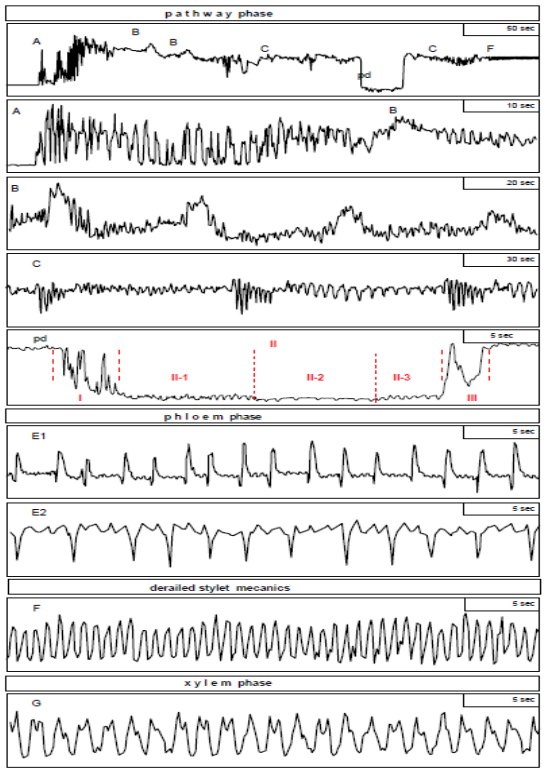
activities happening. Non-penetration waveform is followed by the pathway waveform (C), which represents stylet insertion into the plant until a sieve element is reached. Insects can spend several hours in this phase as the insect stylet moves through the epidermis, mesophyll and other tissues before reaching the sieve element. It is during this phase an insect has to find and to test a suitable site and food for feeding to take place. During this time, three other stylet activities have also appeared regularly such as xylem ingestion (G) as an indicator for insect drinking, derailed stylet mechanics (F) meaning the insect faces difficulty in penetration (Prado & Tjallingii 1994).

Waveforms E1 and E2 represent sieve element salivation and phloem ingestion respectively. Generally, E1 appears for a very short duration however it is an important critical stage for the insect to decide whether to continue to feed or not. If they proceed to phloem ingestion and remain longer, then the plant will be described as susceptible. The study of all these waveforms along with the total time taken to reach sustained E2 phase provide important indications regarding feeding suitability for insects.

1.5 The 'omics' revolution in plant research

Continuing advances in 'omics methodologies and instrumentation is enhancing the understanding of how plants cope with the dynamic nature of their growing environment. Characterization at the genome, transcript, protein and metabolite levels has illustrated the complexity of the cellular response to a whole series of environmental stresses, including nutrient deficiency, pathogen attack, heavy metal toxicity, cold acclimation, and excessive and sub-optimal irradiation (Witzel et al. 2015). Tailoring a crop cultivar to a specific environment is the central challenge for the plant breeder, and reflects the reality that genotype on its own will not generally be sufficient to support a biotechnology-driven crop improvement programme. The genome sequence of the

model plant *Arabidopsis thaliana*, has been known for nearly 15 years (Kaul et al. 2000).



EPG waveform[1]	relative[2] amplitude	rep rate	volt. level	el. origin	plant tissue	aphid activity	remarks
A	100	5-10	e	R	epidermis	cuticle penetration	first wave-form, elec trical stylet contacts
B	75	0.2-0.3	e	R	epidermis / mesophyll	sheath salivation	Waveforms overlap. Therefore, ABC mostly lumped as 'stylet pathway'
C	30	mixed	e	R	all tissues	many activities during pathway	activity (C) in EPG analysis
pd	.		i	emf	all living cells	stylet puncture	3 phases, II with 3 sub-phases:
II-1						salivation | ?	non-pers. virus inoculation
II-2						salivation | ?	non-pers. virus inoculation
II-3						ingestion	non-pers. virus acquisition
E1e	.	2-7	e	emf	mesophyll (?)	unknown	same activity as E1?
E1	.	2-7	i	emf	sieve elements	salivation	persistent virus inoculation
E2							
w	.	4-9	i	emf	sieve elements	watery salivation	persistent virus
p	5	0.5-4	i	R	sieve elements	(passive) ingestion	acquisition
F	5	11-19	e	R/emf	all tissues		derailed stylet 'penetration difficulties' mechanics
G							
w	.	4-9	e	emf	xylem	active ingestion	'drinking'
p	0-60	4-9	e	R	,,	unknown	only occasionally shown

[1] p: peaks, w: waves. [2] useless to provide for emf signals.

Figure 1.7 Diagram showing the EPG waveforms and correlations to aphid feeding behaviours, as defined in the 'Probe 3.4 Manual: Software manual for EPG acquisition and analysis in Windows' distributed by EPG systems, Wageningen the Netherlands.

The full genome sequences of both *B. rapa* (Wang et al. 2011) and B.oleracea (Ayele et al. 2005; Liu et al. 2014) have been published more recently. Substantial amounts of transcript-based data have been acquired for *B. oleracea* (Gao et al. 2014; Izzah et al. 2014; H. A. Kim et al. 2014).

Omics research has a great capability in speeding up breeding processes and several applications for crop improvement, through, e.g., marker-assisted selection and gene pyramiding. In case of brassicaceous vegetables, several populations have been generated to establish linkage maps using simple sequence repeat (SSR), amplified fragment length polymorphism (AFLP), sequence tag (EST) markers with the aim to genetically localize favourable traits by quantitative trait locus (QTL) analysis. A high-density linkage map was established for a *B. oleracea* population segregating for carotenoid concentration in florets and three carotenoid QTL were found (Brown et al. 2014). Recently, genomes of all three sequences *Brassicas* were compared to develop SSR markers and a total of 115,869, 185,662, and 356,522 primer pairs were designed from *B. rapa*, *B. oleracea*, and *B. napus*, respectively (Shi et al. 2014). Not only in genomics and transcriptomics, continuous advances are made in proteomics and metabolomics are facilitating the huge efforts in plant breeding and will trigger a major breakthrough in crop improvement by reduction of time and expenses of producing new stress resistant and highly nutritious crop varieties. Advances in next generation sequencing(NGS) technologies enabled surveying genotype-phenotype-relationships with the highest resolution to date (Wei et al. 2013; Varshney et al. 2014). NGS allows mass sequencing of genomes and transcriptomes, which is producing a vast array of genomic information. Genome-wide expression studies provide breeders with an understanding of the molecular basis of complex traits. In spite of recent developments

in omics technologies, microarray technology is still preferred and used in a number of research studies. This is mainly due to the ease of use and cost effectiveness. NGS and other sequencing are quite expensive at the moment which makes array technology most preferred choice for researchers. This current study will focus on the gene expression profiling and microarray studies to understand aphid and Brassica interactions.

1.5.1 Microarrays: Tool for whole genome expression study

The phenotype of a plant is always underpinned by its transcriptional profile, referred as the collection of genes which express themselves in response to physical or chemical changes in plants.

The past two decades witnessed the development and application of new techniques for transcriptional analysis of crops and model organisms. Microarrays are considered the standard molecular biology tools for assessing the whole genome transcriptional response of organisms (Thompson & Goggin 2006). The microarray has been used in almost all biological disciplines (Yauk & Berndt 2007). The main concept behind microarray technology is based on the southern blotting technique where a small fragment of DNA attached to the surface is probed with a known gene (Southern et al. 1992).This technique was first used to study the gene expression in normal tissue and mouse colon tumour (Augenlicht & Kobrin 1982).

The microarrays are basically microscope slides referred to as platforms which have thousands of known DNA segments corresponding to specific genes deposited at predefined positions (Dunwell et al. 2001). To study the gene expression, the mRNA from the experimental tissue is isolated, converted to single stranded form, fluorescently labelled, and hybridized to probes present on the chip. The intensity of hybridisation is scanned and quantified as gene expression values for all the probes present on the chip.

The microarray chip can vary in its design and make depending upon the manufacturer; Affymetrix produces silicon chips, Agilent probe the DNA segments on the glass slide whereas Illumina uses the microscopic beads. Depending upon the need and design of experiment selection can be made in regard to use of platform. Microarrays are capable of analyzing thousands of genes in parallel which has made them a very useful tool to study the whole genome gene expression of organisms. In plants, the use of microarrays has been substantial. They are used to compare the genetic variation present in species in response to both biotic and abiotic factors as discussed in previous sections. With advancement in the sequencing ability and rapid genome information available it has now become even more important to know the function of genes and microarrays to help in providing this vital information.

1.5.2 Microarrays: Applications

The expression of genes is the main determinant of many important factors like physical appearance, behaviour and plant resistance to biotic stress like insect resistance (Broekgaarden et al. 2008; Broekgaarden, Voorrips, et al. 2011; Pelgrom et al. 2014; Moran & Thompson 2001; Smith & Boyko 2007) and abiotic stress like salinity stress in rice (Walia et al. 2005). The regulation of gene expression is responsible for activation or repression of both constitutive and induced defence mechanisms present in plants.

Microarrays are excellent tools to monitor the whole genome expression of thousands of gene (Meyers et al. 2004; De Vos et al. 2005; Hammond et al. 2005; Duggan et al. 1999). Gene expression analysis has been successfully used to identify the genes involved in response to many biotic and abiotic factors (Walia et al. 2005; Mazzanti et al. 2004; Klingler et al. 1998; Moran & Thompson 2001; Hui et al. 2003; Smith &

Boyko 2007; Voelckel & Baldwin 2004). De Vos et al. (2005) noted a completely different transcriptional response of *A.thaliana* to feeding by phloem feeding *M.persicae* and chewing *P.rapae* larvae. Not only this, insect pests with similar feeding behaviours are known to activate different gene expression patterns in plants. For example, the feeding by aphid (*M.percicae)* and whitefly (*B.tabaci)* show variable transcription profiles in *A.thaliana* (Kempema et al. 2007). Variation in transcriptional response has also been noticed between members of the same genus and even differences between accessions of the same species (Kessler et al. 2004; Gao et al. 2008; Barah et al. 2013; Kuśnierczyk et al. 2008; Singh et al. 1994; Ellis et al. 2000).

Herbivores like aphids are known to regulate expression of genes involved in calcium-dependent signalling, cell wall modifications, signal transduction pathways and glucosinolate synthesis in plants (Thompson & Goggin 2006). The major signal transduction pathway involved in plant defence against herbivory is jasmonic acid (JA), which regulates a number of genes involved in plant defence mechanisms (Liechti & Farmer 2002; Reymond 2000; Schenk et al. 2000; Devoto et al. 2005).

More recently, with the advent of microarray technology becoming more affordable opens a new avenue. The transcriptomics data generated by different microarray experiments led to the development of different public data repositories. These data repositories and newly generated transcriptomics data can potentially be exploited to predict the tolerance or resistance status of genotypes of unknown status by computational model-based comparisons with the genotypes where tolerance or resistance status is known (Gavaghan et al. 2002; Cabrera-Bosquet et al. 2012)In recent years gene expression data is exploited for the class prediction classification (Tan et al. 2008). Survival prediction from gene expression data and other high-dimensional

genomic data have been subject to much research during the last years. The gene expression data provide more accurate information for sample classification and objective treatment strategies. The use of whole genome expression profiling is a highly accepted approach in classifying and predicting the disease in humans. This approach is very popular especially in the case of prediction of various human cancers and neurodegenerative diseases (Bucca et al. 2004; Reis-Filho & Pusztai 2011; Cooper-Knock et al. 2012). Trevino et al. (2011) successfully demonstrated the use of gene expression data and statistical modelling in the prediction of prostate cancer and the use of this approach as a diagnostic tool. The present study will attempt to use the gene expression data generated in the project to study class prediction for classification of Brassica germplasm into resistant or susceptible varieties.

1.6 PGR secure project

The goal of agrobiodiversity conservation, unlike other forms of conservation, is not only the conservation of species and intra-specific genetic diversity related to agriculture, but also to promote its sustainable use in facilitating agricultural production. PGR Secure was a collaborative project funded under the EU Seventh Framework Programme. The aim of PGR Secure was therefore to research novel characterization techniques and conservation strategies for European crop wild relative and landrace diversity, and further, to enhance crop improvement by breeders, as a means of underpinning European food security in the face of climate change. To achieve these goals, PGR Secure had four research themes:

1. Investigation of novel characterization techniques, including: Genomics, phenotyping and metabolomics, Transcriptomics, Focused Identification of Germplasm Strategy.
2. CWR and LR conservation, including: Europe-wide CWR inventory, Exemplar national CWR inventories, European CWR strategy, Europe-wide LR inventory, Exemplar national LR inventories, European LR strategy.
3. Facilitating breeders' CWR and LR use, including: Identifying breeders' needs, Meeting breeders' needs, Integration of conservation and user communities, Pre-breeding – channelling potential interesting germplasm into commercial breeding programmes.
4. Informatics development, including: CWR and LR inventory information web availability, Novel characterization information web availability, Inter-information system operability.

The current study was part of PGR SECURE, under theme 1 which aimed to develop and apply a novel high throughput method for phenotyping gene bank accessions of *Brassica* for resistance towards phloem feeding insects. Starting from a collection of about 3,700 *Brassica* accessions in BrasEDB a selection of approx. 350 were made for the phenotyping using the FIGS approach and literature data. From these, approx. 125 accessions were selected for further analysis using metabolomics techniques and a further subset of these were assessed in terms of resistance/susceptibility using the EPG (Electrical Penetration Graph) determining underlying mechanisms of the resistance measuring insect feeding behaviour. Based on the resistant and susceptible subsets that were identified, next generation sequencing technologies were used to access the total gene transcriptome content of approx. 15 accessions of *Brassica* crop

wild relatives and landraces which allowed the identification of novel genes (and allelic variation) in this plant material. At the same time, partly because of cost and the importance of biological replication, conventional transcriptomics was carried out using Affymetrix *Arabidopsis* GeneChip to assess gene expression variation across different accessions to determine within species variation and response to insect attack; also to compare the use of established and validated *Arabidopsis* arrays with the tailor made arrays from next generation sequencing.

1.6.1 Research aims and thesis outline

At present, it is not possible to feed the world population without the application of insecticides. Worldwide losses caused by insects would be at least 30% to 50% if no insecticides were used (Dedryver et al. 2010; Razaq et al. 2011; Isik & Görür 2009). However, the use of pesticides is hazardous to the environment and is usually not very durable as insects may develop resistance to pesticides very rapidly (Ahmad & Akhtar 2013). This study explores the plant-insect interaction by means of insect feeding and plant defence response to insect attack along with studying gene expression variation between different *Brassica* species and varieties and effects of insect feeding on them. The core objective of the project is to study feeding behaviour of Cabbage aphid (*Brevicoryne brassicae*) in *Brassica* germplasm comprising of CWR and cultivated varieties and investigate transcriptional response Brassica plants to aphid feeding. An attempt has been made to link the feeding behaviour assessment of cabbage aphid with the transcriptional responses of the plants.

Chapter 2 describes the germplasm screening conducted to identify the cabbage aphid resistant material out of the core collection of 200 varieties of *Brassica*. A field experiment was carried out with the help of Oxford Agriculture Trials Company in 2011

in Stratton Audley near Bicester (Oxford). The results from the experiment helped to attain the core aim of narrowing down the number of varieties to 15 from 200, which were then further studied for feeding behaviour assessment and transcriptomics.

Chapter 3 The main aim of this chapter was to investigate the variation in the feeding behaviour of cabbage aphid on different varieties of *Brassica*. The EPG technique was used to assess the feeding behaviour of aphids. Both interspecific and intraspecific variations were noted in response to aphid feeding in the *Brassica* germplasm, over a 6 hour recording time.

Chapter 4 Explores the changes in the gene expression in response to aphid feeding in 15 varieties of *Brassica* categorised as aphid resistant or susceptible based on field the experiment and feeding behaviour assessment. The changes in the gene expressions provide insight into the underlying mechanism adopted by the *Brassica* plants to counter the aphid feeding. These changes also provide vital information with regard to important genes involved in defence mechanisms and hence proposal of candidate genes for further assessment for breeding insect resistant Brassicas.

Chapter 5 Compared the transcription profile of CWR and LR of Brassica in response to cabbage aphid feeding and explored changes in gene expression during the process of domestication. This study provided vital clues regarding the loss of resistance during the process of agriculture making the cultivated species more susceptible to herbivory. It also provided indications for the presence of certain candidate genes for insect resistance in Brassicas.

Chapter 6 Explored the possibility of using gene expression arrays as a tool for prediction of insect resistance in Brassica germplasm. Prediction analysis was

performed using the model selection tool of Partek Genomics Suite version 6.6 (Partek Inc., St. Louis, MO, USA), software. The results from the study show that the model selection tool can be successfully used to develop and test the prediction model for the genotypes of unknown status

Finally**, Chapter** 7 summarizes all the important results from this study along with discussion with respect to available literature. Furthermore, the limitations of study along with the future perspectives are discussed in this chapter.

CHAPTER 2

Field screening Brassica germplasm for feeding suitability to cabbage aphid and whitefly

2.1 Abstract

Biotic stresses, including insect pests, cause severe yield losses that significantly impair crop production. The accurate prediction of crop performance in response to insect pests is one of the biggest challenges in the development of new resistant varieties of crops. The field assessment of available Brassica germplasm to assess suitability for insect feeding can provide vital clues to levels of resistance present in different species of Brassica. The core aim of the study was to explore the natural variation present in different species of Brassica to cabbage aphid (*Brevicoryne brassicae*) and Whitefly (*Aleyrodes proletella*) feeding and also between different accessions of the same species. In addition, selection of group of accessions with respect to contrasting aphid feeding response for further evaluation was aimed. Cabbage aphid and whitefly attack was assessed in terms of numbers of insects present on the plant. Presence of feeding preference by cabbage aphid and whitefly on some accessions more than others was observed in the Brassica germplasm. It was also noted that crop wild relative species were less preferred for feeding by insect pests as compared to landrace accessions. The result from the study confirms high levels of putative resistance in wild relative accessions while partial resistance was found in several other *Brassica* species. Although field assessment provided important information about the presence of insect pest preference/non-preference on Brassica species, additional studies are necessary to make conclusions about the resistance or susceptibility to cabbage aphid in Brassica germplasm. Sixteen accessions varying in levels of resistance to insect feeding were selected for aphid feeding behaviour and gene expression studies.

2.2 Introduction

Brassica crops are being attacked by many insect pests including sap feeding insects which cause serious damage to crops every year (Brown et al. 1999; Schoonhoven et al. 2005). To reduce the amount of damage caused, use of pesticides is most widely used method at the moment, but use of pesticides can be hazardous to the environment as well as to humans. This method is not very durable as insects can develop resistance against pesticides very rapidly (Metcalf 1989; Ahmad & Akhtar 2013). So we need crops which are resistant to these pests and can defend themselves. One ways is to know about the natural resistance factors present in the crop wild relatives (CWR) and the landraces (LR). The insect resistant crops can be a very effective alternative way to address both insect control and future use of harmful pesticides (Lewis et al. 1997). In nature there is considerable variation in different species and also within species which can make a species resistant or susceptible to a particular insect pest (Broekgaarden, Voorrips, et al. 2011; Schoonhoven et al. 2005). Ellis et al. (2000) reported variation in resistance to cabbage aphid between and within wild and cultivated Brassica species.

Resistance to insect pests may vary among individuals from a species because of differences in their genetic makeup. This genetic variation is thought to be maintained by trade-offs between the benefits of reduced herbivory and the costs of resistance and differential selection pressures (Schoonhoven et al. 2005). For example, when a population is frequently infested with a certain herbivore, the resistance traits that are effective against this herbivore will be selected for, whereas the lack of natural selection could result in a loss of the resistance from the population (Mooney et al. 2010). Besides variation among wild populations, variation can also be caused by artificial selection during domestication (Diamond 2002).

However, only very little of this natural variation has been exploited in agriculture. Exploring natural variation among wild relatives of crop plants, or even accessions of crop plants themselves, may yield resistant varieties. The limited availability of information regarding insect resistance in Brassicas makes it very important to characterise the available germplasm for resistance against insect pests. This will help in cataloguing the *Brassica* species according to resistance and susceptibility against insect pests which can help breeders to focus their interest on specific accessions of interest for future use (Varshney et al. 2006). So exploring the variation present among different CWR and LR can help us in recognising the insect resistant varieties. To be able to develop insect-resistant varieties, one first needs to identify resistance sources. Often, these are found in crop wild relatives (CWRs) or landraces (LRs). Once the accessions containing the resistance traits are known, the genes involved need to be located to facilitate efficient transfer to the crop species.

As a first step to identifying resistant material, a collection of 200 accessions, including wild material and landraces of *B.oleracea*, as well as other CWRs, were screened for cabbage aphid and whitefly resistance in the field. The selection of the accessions for the field screening was done in order to keep uniformity of accessions used for phenomics, transcriptomics and other metabolomics studies as part of the bigger PGR secure project. The field screening was undertaken as it is a rapid, simple and reliable technique for evaluating core collections of Brassica germplasm for resistance against insect pests.

2.2.1 Phenotyping in field conditions vs. controlled environment

The phenotype is referred to as the observable characteristics of organisms which results from the interaction between genotype and the environment. Phenotyping the genotypes

for resistance against or susceptibility to insect pests can be performed with help of field assessments as well as laboratory based experiments. Both field and lab based experiments have the 'pros' and 'cons'.

Working in controlled environments like green house or growth room has the advantage of having more control over conditions like temperature, relative humidity, photoperiod, soil, water and nutrients (Saint Pierre 2012). In addition laboratory based experiments can be planned even for the out of season crops without any difficulty. But the major disadvantage of greenhouse or laboratory based experiments is the plant and its interaction with natural environment is lacking which plays a vital role in performance of plants in nature. In contrast, field based experiments provide more realistic information regarding plant response biotic and abiotic factors as the plant is present in its natural habitat, interacting with environmental factors (Saint Pierre 2012). Field screening provides greater ecological validity than laboratory experiments as they are less sample bias as compared to laboratory plants in factors like soil temperature, uniformity in moisture in pots, proximity to ventilators, method of water and air circulation (Brien et al. 2013; Cox & Cochran 1946; Townend & Dickinson 1995). In addition to this field screening takes into account the insect pest variability also interaction of insect pests with other herbivores which may affect the insect pest performance in nature. Although field experiments have its positives but they are difficult to replicate as there is limited control over physical factors like temperature, weather and insect pest species which vary every season. Nevertheless, they are the vital tool used in phenotyping crop species response to biotic and abiotic stresses. The field screening in the current study was conducted to evaluate the natural variation in these

plants and also, due to the large scale of experiment which couldn't be accommodated in the glasshouse.

2.2.2 Chapter aim

The aim of the study is to explore the natural variation in feeding response present in different species and within species of Brassicas. The field experiment to characterize the feeding preferences of cabbage aphid (*Brevicoryne brassicae*) and Whitefly (*Aleyrodes proletella*) in a core collection of 200 *Brassica* accessions representing 18 Brassica species (Table 2.1) was conducted. This study focussed on the natural infestation of cabbage aphid and whitefly on plants in the field. The scientific questions to be answered by this study were firstly, is there any difference between wild relatives and landraces of *Brassica* species included in the core collection for resistance to cabbage aphid and whitefly feeding. Secondly, the most important output expected from the study was identification of the most preferred and non-preferred accessions by cabbage aphid for feeding, which could be used further for insect feeding behaviour investigations and whole genome transcriptomics.

2.3 Materials and methods

2.3.1 Plant material

Phenotyping plants for resistance to cabbage aphid and whitefly in all known CWR and LR would be very time and space consuming and not possible in the time frame of this project so a core collection of 200 Brassica accessions were selected (Table 2.1). This collection included both CWR and LR originating in European Atlantic coast to North Sea, Mediterranean basin and Russia. The wild relatives of Brassica species included in the core collection were *B.balerica, B.bovoniana, B.bourgeaui, B.cretica, B.frituculosa,*

B.hilarionis,B.incana, B.insularis, B.marcocarpa, B.maurorum B.montana, B.robertiana, B.rupestris, B.spinescens, B.villosa and also LR accessions of B.oleracea were included in study.

Table 2.1: List of Brassica species used in field trial

S.no	CWR/ LR	Genus	Species	Genome type	No. of accessions used in field trial
1	CWR	*Brassica*	*balerica*	C	2
2	CWR	*Brassica*	*bivoniana*	C	1
3	CWR	*Brassica*	*bourgeaui*	C	2
4	CWR	*Brassica*	*cretica*	C	12
5	CWR	*Brassica*	*fruticulosa*	B	16
6	CWR	*Brassica*	*hilarionis*	C	1
7	CWR	*Brassica*	*incana*	C	10
8	CWR	*Brassica*	*insularis*	C	2
9	CWR	*Brassica*	*marcocarpa*	C	3
10	CWR	*Brassica*	*maurorum*	B	2
11	CWR	*Brassica*	*montana*	C	7
12	CWR	*Brassica*	*oleracea*	C	45
13	CWR	*Brassica*	*robertiana*	C	1
14	CWR	*Brassica*	*rupestris*	C	3
15	CWR	*Brassica*	*spinescens*	B	1
16	CWR	*Brassica*	*villosa*	C	7
17	LR	*Brassica*	*oleracea*	C	85
Total					200

All these species have genome C with exception of *B.fruticulosa, B.maurorum and B.spinescens* with genome B. All the seed material (200 accessions) was provided by Wageningen UR Plant Breeding, Wageningen University and Research Centre, The Netherlands. The detailed list of accessions included in the study is provided in supplementary appendix I.

2.3.2 Field Experiment

The field experiment was carried out with help of Oxford Agriculture Trials Company in June 2011 to September 2011 in Stratton Audley near Bicester (Oxford). The trial plot was divided into 3 blocks with 200 subplots and each subplot with 4 replicates of the same accession. A randomised complete block design was used. All blocks were labelled blind (the accession name was removed and a number was assigned as an identifier) and the MS Excel random number generator function was used for randomisation. The plot design is shown in figure 2.1 and figure 2.2 shows the Brassica plants at week 8.

2.3.3 Cultivation of Brassica accessions for Field Experiment

The seeds for the 200 accessions selected were planted individually in 6cm pots containing 6 parts peat based compost (Humax multipurpose) to 1 part Silvaperl under 16:8 hour light: dark regime at 24±3 °C with 70% humidity for 5 weeks in the glasshouse at the University of Birmingham. The plants were transferred to the field in Bicester where they were monitored for growth and natural infestation (choice test) of cabbage aphids and whiteflies. The accessions were planted in 24 x 40 m grid. The plants were placed at a distance of 50 cm^2 apart and were placed in 2 x 2 positions. So in total there were 2400 individual plants to screen for aphid and whitefly infestation. The whole plot was surrounded by a set of guard plants (*B.nigra*). Weeds were handpicked and no herbicides, pesticides or fungicides were used. The staff at Oxford Agriculture monitored the plants over the summer and reported a flea beetle attack which resulted in loss of some of the accessions (indicated as X in figure 2.1). On day of scoring, difference in plants size, presence of aphids and whitefly on plants and presence of lady bird over some of the plants was noticed. For the current study, data

was collected for level of aphid infestation and presence or absence of whitefly on the plants.

	Block 1					Block 2					Block 3					
G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
G	135	450	404	368	469	469	460	379	131	X	451	X	409	67	X	G
G	X	353	384	261	332	416	41	392	468	X	421	332	323	150	X	G
G	199	330	131	276	326	18	409	429	395	X	127	26	430	59	452	G
G	97	21	61	373	343	15	188	259	336	167	189	325	X	454	392	G
G	397	27	227	466	259	227	39	X	410	3	450	329	X	169	363	G
G	453	417	144	336	350	346	13	391	465	330	371	X	345	X	30	G
G	2	392	363	X	260	135	354	456	269	21	X	174	438	X	346	G
G	357	167	174	203	41	144	67	413	394	397	456	X	275	9	326	G
G	344	37	342	329	382	26	125	466	368	66	393	25	13	23	128	G
G	371	461	460	20	14	327	8	467	61	451	311	X	X	334	351	G
G	201	39	29	310	313	417	203	235	260	53	413	229	434	115	310	G
G	8	316	12	116	6	217	59	311	427	10	322	457	404	453	272	G
G	307	32	401	449	31	X	X	326	4	145	459	364	22	232	152	G
G	5	434	429	394	325	X	137	X	320	458	416	39	132	461	399	G
G	322	X	34	387	451	421	349	450	266	352	X	X	396	100	X	G
G	334	463	16	100	19	24	396	276	401	17	16	17	250	348	463	G
G	190	266	311	410	137	232	100	462	329	316	400	220	331	X	237	G
G	18	150	53	297	26	371	19	22	323	6	199	24	10	357	190	G
G	237	X	346	127	400	328	169	23	322	152	260	313	382	119	166	G
G	3	229	25	X	24	324	X	332	166	200	466	41	394	297	146	G
G	323	128	395	418	66	250	119	220	398	97	387	11	145	458	402	G
G	67	232	464	17	349	28	453	310	31	237	462	464	320	235	X	G
G	452	52	146	347	391	382	454	393	404	115	467	54	336	276	61	G
G	265	455	250	198	468	116	261	32	X	345	34	324	356	308	135	G
G	427	399	324	X	348	X	127	351	X	63	97	66	18	321	327	G
G	56	15	458	345	335	331	357	254	X	384	217	347	X	330	469	G
G	33	456	396	352	393	362	38	387	321	449	42	144	200	265	125	G
G	169	409	362	35	234	399	X	265	325	128	333	137	X	29	352	G
G	235	245	36	454	364	335	25	229	34	132	245	395	188	X	131	G
G	308	465	337	188	320	348	29	430	201	56	X	335	349	21	X	G
G	416	217	13	356	413	33	342	459	334	464	167	203	201	386	40	G
G	X	430	220	354	X	234	297	275	37	190	401	384	427	X	398	G
G	321	38	275	11	327	36	363	333	174	35	6	53	X	27	350	G
G	398	40	459	X	421	27	386	198	461	54	397	368	X	261	266	G
G	10	59	331	119	9	400	307	337	364	245	7	8	28	465	373	G
G	132	269	328	200	145	452	402	40	434	42	307	234	429	337	X	G
G	7	63	272	457	23	347	16	455	308	457	449	254	116	3	328	G
G	115	333	351	254	166	5	150	313	30	350	468	269	343	32	X	G
G	X	462	54	42	28	272	X	146	373	343	4	X	X	362	417	G
G	125	30	4	467	402	7	463	9	356	353	259	38	37	56	X	G
G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G

Figure 2.1: The position of each accession in the field. Each cell represents 4 plants of the same accession; X refers to accessions missing on day of scoring

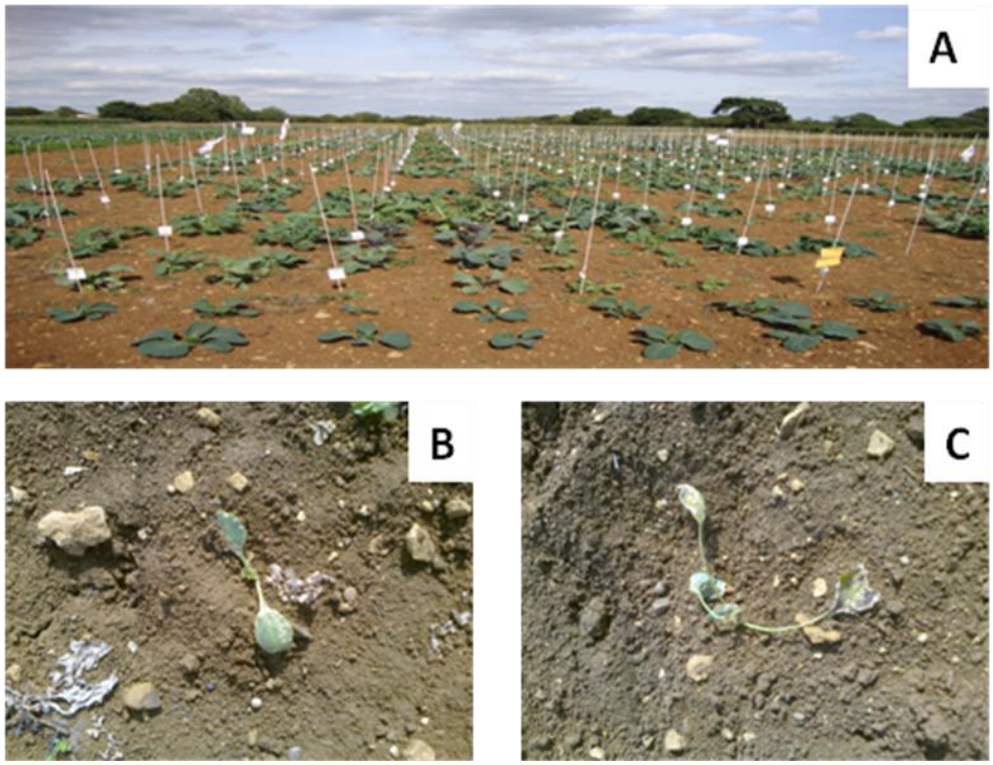


Figure 2.2: Field trial in Stratton Audley near Bicester. (A) Shows plants at 8 weeks in the field, dead plants as a result of flea beetle attack (B-C)

2.3.4 Statistical Analysis

Data analysis was done with MINITAB 15 using the Anderson-Darling Normality Test and Analysis of Variance (ANOVA) for aphid infestation on each genotype across 3 blocks and utilized the Univariate General Linear Model (GLM) to allow for the unbalanced number of replicates per block due to missing plants on day of scoring. In

addition the Chi square test of association was used to look for any significant effect of presence of accessions next to each other in the field.

2.4 Results

2.4.1 Field experiment

As a result of unusual climatic conditions (initial cold summer) or other environmental effects, the normal natural aphid and whitefly infestation was delayed in year 2011 when the study was conducted. So the final scoring was delayed by 3 weeks. The 15 week old plants in the field were scored by University of Birmingham staff for presence of aphids and whiteflies according to method established by (Mamun et al. 2010). A total of 2400 plants were screened and scored. Aphids were scored on a scale of 0-4 where 0 is complete absence and 4 is highly infested, and whiteflies were scored 0-1 for absence or presence. Data is presented in supplementary table S2.1. The accessions not accounted for were those that either did not germinate or were lost due to flea beetle attack. Figure 2.4 the spread of aphid and whitefly infestation in the experimental plot

Figure 2.3: Trial field in Stratton Audley near Bicester (A), 15 week old cabbage (B), whitefly infestation (C) and plant highly infested by cabbage aphid (D)

	Block 1					Block 2					Block 3					
G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
G	135	450	404	368	469	469	460	379	131		451		409	67		G
G		353	384	261	332	416	41	392	468		421	332	323	150		G
G	199	330	131	276	326	18	409	429	395		127	26	430	59	452	G
G	97	21	61	373	343	15	188	259	336	167	189	325		454	392	G
G	397	27	227	466	259	227	39		410	3	450	329		169	363	G
G	453	417	144	336	350	346	13	391	465	330	371		345		30	G
G	2	392	363		260	135	354	456	269	21		174	438		346	G
G	357	167	174	203	41	144	67	413	394	397	456		275	9	326	G
G	344	37	342	329	382	26	125	466	368	66	393	25	13	23	128	G
G	371	461	460	20	14	327	8	467	61	451	311			334	351	G
G	201	39	29	310	313	417	203	235	260	53	413	229	434	115	310	G
G	8	316	12	116	6	217	59	311	427	10	322	457	404	453	272	G
G	307	32	401	449	31			326	4	145	459	364	22	232	152	G
G	5	434	429	394	325		137		320	458	416	39	132	461	399	G
G	322		34	387	451	421	349	450	266	352			396	100		G
G	334	463	16	100	19	24	396	276	401	17	16	17	250	348	463	G
G	190	266	311	410	137	232	100	462	329	316	400	220	331		237	G
G	18	150	53	297	26	371	19	22	323	6	199	24	10	357	190	G
G	237		346	127	400	328	169	23	322	152	260	313	382	119	166	G
G	3	229	25		24	324		332	166	200	466	41	394	297	146	G
G	323	128	395	418	66	250	119	220	398	97	387	11	145	458	402	G
G	67	232	464	17	349	28	453	310	31	237	462	464	320	235		G
G	452	52	146	347	391	382	454	393	404	115	467	54	336	276	61	G
G	265	455	250	198	468	116	261	32		345	34	324	356	308	135	G
G	427	399	324		348		127	351		63	97	66	18	321	327	G
G	56	15	458	345	335	331	357	254		384	217	347		330	469	G
G	33	456	396	352	393	362	38	387	321	449	42	144	200	265	125	G
G	169	409	362	35	234	399		265	325	128	333	137		29	352	G
G	235	245	36	454	364	335	25	229	34	132	245	395	188		131	G
G	308	465	337	188	320	348	29	430	201	56		335	349	21		G
G	416	217	13	356	413	33	342	459	334	464	167	203	201	386	40	G
G		430	220	354		234	297	275	37	190	401	384	427		398	G
G	321	38	275	11	327	36	363	333	174	35	6	53		27	350	G
G	398	40	459		421	27	386	198	461	54	397	368		261	266	G
G	10	59	331	119	9	400	307	337	364	245	7	8	28	465	373	G
G	132	269	328	200	145	452	402	40	434	42	307	234	429	337		G
G	7	63	272	457	23	347	16	455	308	457	449	254	116	3	328	G
G	115	333	351	254	166	5	150	313	30	350	468	269	343	32		G
G		462	54	42	28	272		146	373	343	4			362	417	G
G	125	30	4	467	402	7	463	9	356	353	259	38	37	56		G
G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G

Figure2.4: Geographical location of each accession in the field. Each cell represents 4 plants of same accession; Different colours of each cell represent infestation by insect pests. Green cells represent aphid infestation; yellow represents whitefly infestation and red represent plants where both aphid and whitefly were present.

2.4.2 Observational scoring and statistical analysis

The data collected were found to be normally distributed with P-Value < 0.005 using Anderson-Darling Normality Test. However due to the unbalanced nature of the data due to presence of missing data, the Univariate General Linear Model (GLM) was used to check the level of aphid infestation on different accessions. . In our experiment our dependant variable was aphid infestation and independent variable were block and accession. ANOVA results are shown table 2.2. Similarly the effect of white fly infestation on each accession present in the field was assessed using the GLM model which is shown below in table 2.3.

Table 2.2: Analysis of Variance using adjusted SS for aphid infestation (General Linear Model: Aphid infestation versus Accession, block)

Factor	DF	SS	MS	F-ratio	P-value
Accession	190	608.355	3.202	6.59	0.000
Block	2	0.292	0.146	0.3	0.741
Error	285	138.375	0.486		
Total	477	748.326			
S = 0.696797 R-Sq = 81.51% R-Sq (adj) = 69.05% SS= Type III Sums of Squares					

The ANOVA results highlighted that there is no significant difference for the aphid infestation in all 3 blocks ($F_{2,285} = 0.30$, $P > 0.1$) but also indicated that aphid infestation varied significantly between the accessions ($F_{190,285} = 6.59$, $P \leq 0.001$). This was in contrast to whitefly infestation where both accession ($F_{190,284} = 2.178$, $P \leq 0.001$) and

block ($F_{2,284}$ = 5.793, P ≥0.003) have significant effects on extent of whitefly infestation in field.

Table 2.3: Analysis of Variance using adjusted SS for whitefly infestation (General Linear Model: white infestation versus Accession, block)

Factor	DF	SS	MS	F-ratio	P-value
Accession	190	63.226	0.333	2.178	0.00
Block	2	1.7703	0.885	5.793	0.003
Error	284	43.396	0.153		
Total	476	108.392			

S = 0.390902 R-Sq = 60.23% R-Sq (adj) = 33.35% SS= Type III Sums of Squares

A Chi Square Test of association was also conducted using MINITAB 15 on the data obtained from the field to look for any association between aphid infestations on the different genotypes present in the field due to their spacial proximity in the field. The Chi Square (X2) test is the most important and most used member of the nonparametric family of statistical tests. Chi Square is employed to test the difference between an actual sample and another hypothetical or previously established distribution such as that which may be expected due to chance or probability. Chi Square can also be used to test differences between two or more actual samples. To get the expected frequency the row total is multiplied by column total, and then divide by the overall total.

Basic Computational Equation

$$X^2 = \sum \frac{(\text{Observed frequency} - \text{Expected frequency})^2}{\text{Expected frequency}}$$

The data were placed into three categories: aphid absent in both adjacent plots; aphid present or absent in any one of the adjacent plots and aphid present in both adjacent plots. The plots where there were missing data because of absences of plants were excluded from the analysis. The null hypothesis for this test was that all three categories are independent, i.e. the value assigned to one plot is independent and has no effect on the value assigned to the adjacent plot. The Chi-Square results confirmed that there is no association between the aphid infestations between 3 blocks due to spacial presence of genotypes in the blocks as the P-value is more than 0.05 so the null hypothesis is accepted i.e. aphid infestation is independent of location of genotypes in the field.

Table 2.4: Chi-Square Test of association between aphid occurrence and genotype location

Block No.	count value	Aphid	Whitefly	Both	Total
1	Observed count	38	19	29	86
	Expected count	26.41	23.55	36.04	
	Chi-Square contribution	5.08	0.8797	1.37	
2	Observed count	23	19	35	77
	Expected count	23.64	21.09	32.27	
	Chi-Square contribution	0.017	0.20	0.23	
3	Observed count	13	28	37	78
	Expected count	23.95	21.36	32.69	
	Chi-Square contribution	5.0065	2.063	0.56	
Total		74	66	101	241
Chi-Sq = 15.439, DF = 4, P-Value = 0.004					

Similarly, another Chi-Square test of association was conducted to see if presence of aphid or whitefly on one accession affects the presence of the other. The null hypothesis for this test was that presence of aphid and whitefly is independent of each other and has no effect on presence or absence of each other's presence. The Chi-Square test shows a degree of association between presence of aphid and whitefly on each accession (Chi-Sq = 15.439, DF=4, P-value < 0.05), so the null hypothesis is rejected. The aphid data set

was then arranged, based on the mean score from smallest to highest number and grouped as resistant and susceptible accessions

Table 2.5: Chi-Square Test of association between aphid and white fly occurrence

Block No.	count value	Aphid absent in both plots	Aphids present in one plot and absent in other	Aphid present in both plots	Total
1	Observed count	62	67	17	146
	Expected count	70.25	58.48	17.27	
	Chi-Square contribution	0.9695	1.2418	0.0042	
2	Observed count	59	39	16	114
	Expected count	54.85	45.66	13.48	
	Chi-Square contribution	0.3132	0.9718	0.4695	
3	Observed count	58	43	11	112
	Expected count	53.89	44.83	13.25	
	Chi-Square contribution	0.3131	0.0771	0.3812	
Total		179	149	44	372
Chi-Sq = 4.759, DF = 4, P-Value = 0.313					

. The resistant group had a mean value of 0 whereas the susceptible group had a mean value of 3 or above. The selection of accessions for further study from both the groups was done in accordance with the other work packages as part of bigger PGR secure project, so as to keep the uniformity in the end results of project as a whole. The selected accessions represent both CWR and LR species from Brassica germplasm which were either preferred for feeding by aphid (scored 3 or 4 in all 3 blocks) or non-preferred (scored 0 in all 3 blocks) in the field trial. The details of the selected genotypes are given in table 2.5.

2.5 Discussion

Field screening during June-September 2011 suggest the presence of different levels of putative resistance towards insect pests in *Brassica* species. Using this approach it was possible to screen 200 accessions in a cost efficient way and in a relatively short period.

This study explored the natural variation present in *Brassica* germplasm towards sap feeding cabbage aphid and whitefly. The results from the study suggest that CWR are less preferred as compared to the LR species by both cabbage aphid and whitefly. The CWR *B.incana, B.fruticulosa, B.villosa* have previously been reported to possess resistance to cabbage aphid (Ellis et al. 2000).The accessions from these species were found to possess putative resistance in this study, as well as the level of insect infestation was lower on them as compared to LR *B.oleracea* species. Singh et al. (1994) reported presence of resistance to cabbage aphid through antibiosis in wild species, *B.fruticulosa*, *B.insularis* and *B.villosa*. Although no consistent antixenosis was found in field screening, lower numbers of aphids were observed on the CWR. Brassicaceae are characterized by the presence of constitutive glucosinolates and induced volatile secondary defences, such as isothiocyanates and nitriles that are derived from the hydrolysis of glucosinolates upon damage by herbivores. This characteristic quality may aid in resistance via both antixenosis and antibiosis. Interspecies as well as intra-species variation in level of resistance against aphid feeding in B*rassica* species are reported in the literature (Verdugo et al. 2012; Cole 1994; Broekgaarden et al. 2007).The field results showed variation in the level of feeding preferences of cabbage aphid and whitefly infestation both between and within Brassica species. It is often reported that presence of more than one pest can affect the performance of the other pest (Brien et al. 2013; Broekgaarden, Snoeren, et al. 2011; Broekgaarden et al. 2008). In this study the presence of association between aphid and whitefly was noted. ANOVA analysis showed that blocks within the field did not have any effect on aphid presence but it is highly affected spread of whitefly.

Although field screening provided vital information with respect to aphid feeding preference on different Brassica species, there were some limitations to this study. First and foremost, field screening was conducted only once due to the cost and time constrains within the project. There are many important factors like soil, pest and species heterogeneity and plant neighbourhood which was not accounted for due to the reason that the field experiment was designed just to categorize the Brassica accessions into putative resistant and putative susceptible categories in order to reduce the number of accessions for further experiments.

As this study was part of the PGR secure project, Brassica accessions showing contrasting characters which were suitable for feeding behaviour assessment, transcriptomics and metabolomics studies were selected. By doing so, the core aim of the study to evaluate 200 accessions for aphid feeding preference and the selection of accessions with different levels of putative resistance to aphid feeding was achieved. Table 2.5 shows the list of accessions which were further explored to look at variation in feeding behaviour and gene expression changes in response to cabbage aphid infestation. From this point onward in the thesis, the term resistance will be used for the non-preferred and susceptible will be used for preferred accessions with respect to aphid feeding.

2.6 Conclusion

To conclude, field screening assessment was conducted to evaluate the feeding preference of insect pests, cabbage aphid and whitefly in a core collection of 200 *Brassic*a accessions which included wild relatives and LR of *Brassica* species. Results

obtained from the field reflected that CWR possess higher levels of the putative resistance than LR accessions.

Table 2.6: List of Brassica CWR and LRs selected from field experiment for further assessment of feeding behaviour and changes in gene expression in response to cabbage aphid infestation

S.no	Seed no.	Accession no/ cultivar name	Species	subtaxon	Material	Suitability in Field
1	24	BRA 2856	*B.incana*		CWR	Non-Preferred
2	26	K 10373	*B.incana*		CWR	Non-Preferred
3	37	K 9404	*B.montana*		CWR	Preferred
4	38	BRA 1644	*B.montana*		CWR	Non-Preferred
5	199	BRA 2401	*B.oleracea*	capitata	CWR	Preferred
6	321	57071	*B.oleracea*		CWR	Non-Preferred
7	325	70432	*B.oleracea*		CWR	Non-Preferred
8	397	BRA 2923	*B.villosa*		CWR	Non-Preferred
9	398	K 6926	*B.villosa*		CWR	Preferred
10	401	K 10259	*B.villosa*		CWR	Non-Preferred
11	453	Bol2009-0080	*B.fruticulosa*		CWR	Non-Preferred
12	454	Bol2009-0081	*B.fruticulosa*		CWR	Preferred
13	54	BRS-0103	*B.oleracea*	acephala	LR	Non-Preferred
14	116	CGN18468	*B.oleracea*	acephala	LR	Non-Preferred
15	229	HRIGRU 6568	*B.oleracea*	capitata	LR	Non-Preferred
16	260	BRA 915	*B.oleracea*	capitata	LR	Preferred

Significant differences were found between and within species in respect to presence of aphid and whitefly. The core aim of grouping the whole core collection into two distinct groups - aphid preferred and non-preferred accessions was achieved. Additionally, this

study helped in selecting 16 *Brassica* accessions which were further used in the feeding behaviour assessment and gene expression studies. Although field screening did not provide any definitive confirmation with respect to resistance/susceptible status of these accessions, but together with results from feeding behaviour experiments (EPG) and gene expression studies will result in categorising these more firmly into resistant or susceptible accessions.

List of supplementary table

Table S2.1: Scoring Data for the aphid and white fly infestation. The value 0 to 4 (where 0 for highly resistant and 4 for susceptible) represents the level of resistance present in Brassica plant against insect pest. X represents missing plants. The accessions used in further study are highlighted in yellow

CHAPTER 3

Investigating the differences in feeding behaviour of cabbage aphid on Brassica genotypes using Electrical Penetration Graph (EPG)

3.1 Abstract

Plants have developed defence mechanisms to deal with attack from herbivorous insects. Brassica crops are frequently attacked by the cabbage aphid, with significant negative impact on yield and crop production. The variability in feeding behaviour of cabbage aphid on different CWR and LR accessions was evaluated in this study. It also aimed to characterize and compare feeding behaviour of *B.brassicae* using EPG technique for resistance screening in *Brassica* germplasm and classify the germplasm into susceptible or resistant groups based on the acceptability of feeding on each genotype.

EPG revealed the presence of waveforms Non-Penetration (NP), Pathway (C), E1 (Phloem salivation), E2 (phloem ingestion), Xylem, Derailed stylet (DS) and potential drop. The waveforms obtained during a 6h recording time were comparable to already defined waveform patterns. The main criteria applied to classify genotypes into susceptible or resistant was time spent in NP, Pathway , E2 and time to locate phloem. A Kruskal-Wallis non- parametric analysis indicated that cabbage aphid exhibited significantly different feeding patterns on different genotypes. The most susceptible genotypes based on EPG analysis were BRA2923, BRA915 as time spent to locate phloem was shorter and additionally aphid had sustained phloem feeding on these genotypes. In contrast, genotypes BRA 2856, K6926, K10259 were among highly resistant varieties. The time spent in NP and pathway was significantly higher and aphids could not maintain a sustained phloem feeding. The results of this study suggest that the deterrent agents vary in activity and may hinder stylet activity at different levels (epidermis, parenchymatous tissue or phloem) depending on plant species. This study provides vital information about the level of

resistance against cabbage aphid feeding present in Brassica germplasm and classification of germplasm as susceptible or resistant which will be used for gene expression analysis studies.

3.2 Introduction

Aphids (Hemiptera, Aphididae) are one of the most serious pests of vegetable *Brassica* crops (Dedryver et al. 2010). It is reported that around 30% to 80% *Brassica* crop yield is lost annually due to aphid infestation in developed and developing countries respectively (Dedryver et al. 2010; Razaq et al. 2011; Isik & Görür 2009). The cabbage aphid, is a severe pest of *Brassica* crops throughout temperate regions of the world (Brown et al. 1999). Aphids are extremely efficient at colonising new plants and due to parthenogenic reproduction can increase their population rapidly. It is known for its highly specific feeding behaviour which involves consumption of phloem sap from *Brassica* or closely related plant species (Cole 1997a). However not much is a known about aphid-plant interactions at the molecular level resulting in lack of information about the level of resistance present in *Brassica* germplasm. Plant resistance in response to aphid feeding is considered as a major component in insect-plant interaction (Guerrieri & Digilio 2008). To explore plant and aphid interactions in *Brassica*, Cabbage aphid and *Brassica* species provide an ideal system to study feeding behaviour patterns and preferences of the aphid.

Aphids are primarily phloem feeders and occasionally consuming xylem sap to overcome dehydration after a period of starvation (Spiller et al. 1990). Aphids receive the majority of their nutrients from plant sap with help of piercing and sucking mouthparts called stylets. The stylets are used for exploring and probing plant tissue to locate sieve elements (Dixon 1998). A stylet penetration process enables aphids to puncture the symplast and exploit intracellular compartments without wounding and reach phloem sap (Powell et al. 2006). Phloem sap is rich in sugars and free amino acids therefore providing a high source of

carbon and nitrogen and is often free from toxins and feeding deterrents (Dinant et al. 2010). Aphid feeding not only reduces crop yield but also has the potential to transmit viruses which is another cause of crop loss (Dedryver et at. 2010). Aphid transmitted viruses are responsible for diseases in *Brassica* (Matthews, 1991) such as *Turnip mosaic virus* (TuMV), spread by *B.brassicae* in *Brassica* (Walsh et al. 1999; Harvey et al. 2007; Suehiro et al. 2004), and *Cauliflower mosaic virus* (CaMV*)* transmitted by *B.brassicae and M.persicae* (Palacios et al. 2002; Martinière et al. 2009). They have a short life cycle and the ability to produce their offspring though both asexual and sexual reproduction (Matis et al. 2007). Therefore, an outbreak phenomenon could easily happen in a short time when the conditions suit them. At present aphid management in Brassica crops is heavily reliant on insecticides and aphicides that account to 39% of all insecticides application (Garthwaite 2012). They are used to control the aphid outbreaks which is harmful to environment and also poses danger of getting incorporated into the human diet (Bhatia et al. 2011) making development of insect resistant varieties a priority. This will enhance economic, ecological and environmental benefits of crop plants (Teetes 2014). The knowledge of plant-aphid interactions, feeding behaviour assessments and gene expression studies will help in better understanding of resistance mechanisms. The feeding behaviour study will provide an insight into the location and path followed by aphid inside plant tissue to feed. This information along with the gene expression studies can provide vital help in discovering genes involved in response to feeding. The knowledge of resistance factors can then be exploited to develop new more resistant crop varieties.

3.2.1 Aphid feeding

Aphids are able to pierce the plant cuticle and the cell wall with their needle-like stylets (Schoonhoven et.al 1998). The stylets are made of two outer mandibles and two inner maxillae, together forming a food canal and a salivary canal (figure 1.5). As the stylet penetrates, a salivary sheath made of lipoproteins is secreted. The sheath encases the stylets and makes a rigid tube to facilitate penetration and direction of probe in search of the feeding site (Pollard 1973; Schoonhoven et.al 1998). It has also been hypothesised to protect aphids against plant defence (Miles 1999).

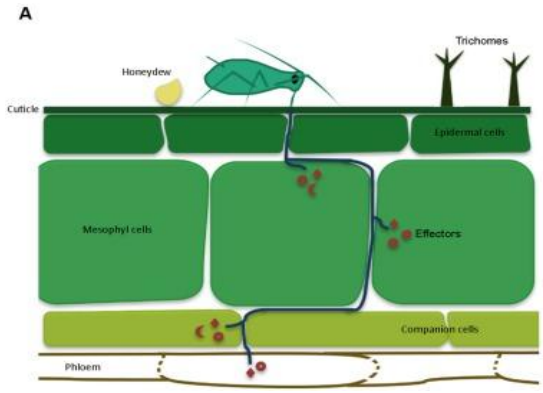


Figure 3.1: Schematic overview of the plant-aphid interface: Aphid mouthparts penetrate the leaf surface upon encountering preformed defences such as trichomes and waxes. The figure was adapted from (Jaouannet et al. 2014).

The path followed by aphid to reach phloem sap is mostly extracellular (Jaouannet et al. 2014; Tjallingii & Esch 1993) from the cuticle to the vascular element of the plant, stylet penetrates between mesophyll cells (Figure 3.1). Most cells along the stylet-pathway are punctured, including the phloem cells. Saliva, containing effectors is secreted into the different cell types as well as the apoplast. Aphids are also known to use chemical gradients

to locate the feeding site (Klingler et al. 1998; Powell & Hardie 2000; Schoonhoven et.al 1998). Probing and feeding behaviour during compatible plant–aphid interactions has been well documented and is thought to involve cues such as pH, sucrose and amino acid content (Hewer et al. 2011; Will & Vilcinskas 2013).

When a suitable feeding site is located, saliva is injected into the sieve element to overcome any phloem based defences (Will et al. 2007). Sieve element sap moves though the stylet food canal under the often large pressures found in the phloem (Will et al. 2013; Tjallingii 2006) The cibarial pump in the head regulates this flow ((Ponsen 1987). The saliva contains many enzymes which help in overcoming the plant defence by detoxifying various allelochemicals and other plant defences (Tjallingii & Esch 1993; Miles 1999) . Under most conditions aphids ingest more sugar than they can assimilate, so that their faeces consist carbon; low nitrogen sap (honey dew) that is passed out of the anus (Wilkinson et al. 1997). Aphids secrete honeydew on the leaf surface, which may also contain molecules that alter plant defence responses (Jaouannet et al. 2014). Researchers are now able to investigate what happens as an aphid is piercing and then feeding from a plant by using a technique called the electrical penetration graph (EPG) technique which is described in next section.

3.2.2 Electrical Penetration Graph (EPG)

The EPG technique is used to monitor the aphid feeding behaviour and can give indications about the source of plant resistance towards insects. Some quantitative information about insect behaviour studies like settling down, population growth has been done visually by monitoring aphid feeding using microscope or with video recording (Kindt, 2004). However it is impossible to predict what is happening inside the plant tissue. The EPG

technique provides detailed information about insect feeding and stylet movement inside the plant, while aphid locates the phloem. The ability of insects to reach the phloem sap and commence feeding is different on different host plant species and depends on many factors. In resistant plants, the insect commonly takes longer to access phloem and have sustained feeding following the initial stylet insertion as compared to susceptible plants (Prado & Tjallingii 1994; Prado & Tjallingii 1997; Tjallingii & Esch 1993).

The EPG technique was first developed in 1960s and a significant modification was made in 1978 by substitution of AC(Alternating Current) circuitry with DC (Direct Current) circuitry (Tjallingii 1985). The basic principle of EPG is the integration of a plant and an insect in an electric circuit. When the aphid inserts its stylet into plant tissue, the circuit is completed and EPG waveform will be recoded and visualized (Janssen et al. 1989; Prado & Tjallingii 1994; Huang et al. 2012; Tjallingii 1985). For aphids, the behaviour associated with these waveforms has been already described (Tjallingii 1985). Broadly waveforms can be classified into 7 distinct patterns as described below and also shown in Figure 1.7.

The main patterns identified are:

1. Mean duration of non-penetration period (NP)
2. Period of pathway waveform pattern (C).
3. Period of potential drop (Pd)
4. Period of salivation into sieve element (E1).
5. Period of sustained phloem sap ingestion (E2).
6. Period of derailed stylet mechanics (F)
7. Period of ingestion of xylem sap (G)

The EPG waveforms are also used to group the aphid resistance mechanisms into three categories mainly surface based resistance, pathway based resistance and phloem based resistance. The surface based resistance is represented by non penetration waveform. When aphids land on the plant leaf, the first resistance barrier is the leaf morphology. The leaf surface and the cell wall incorporate different physical defences such as a waxy coating on the leaves, presence of thorns or hairy tissues and also chemical defences that are rapidly activated when presence of insect is detected by the plant (Bruce 2015a; Walling 2008; Kessler & Halitschke 2007). The pathway based resistance comes into play when the aphid is able to penetrate the cell wall and looks for the phloem. The waveforms C, PD, Derailed stylet and xylem all fall under this category. During this activity aphids can encounter both primary and secondary metabolites, phenolic substances, proteins and enzymes which can act as resistance factors and make aphid feeding difficult (Tjallingii & Esch 1993; Louis et al. 2012). The phloem based resistance is represented by waveforms E1and E2. The active feeding period (E2) of EPG is considered important in terms of resistance against aphids as it is the phase where an aphid obtains its nutrition. Phloem based resistance is often taken into account when deciding about suitability of genotypes to insect feeding (Pegadaraju et al. 2007; Dinant et al. 2010; Broekgaarden et al. 2012).

Electrical Penetration Graph technique was first tested and used for studying aphids (Tjallingii 1985; Janssen et al. 1989). Since then, it is one of the common tools used to characterize and identify, in detail, plant resistance factors against insects with piercing mouthparts by monitoring their probing behaviours (Calatayud et al. 1994; van Helden & Tjallingii 1993; Spiller et al. 1990; Schwarzkopf et al. 2013; Diaz-Montano et al. 2007). It has been applied to a wide range of other piercing insects as such Brown plant hopper (Ghaffar et al. 2011),

whiteflies (Janssen et al. 1989; Lei et al. 1998), thrips (Hunter et al.1993; Harrewinji et al. 1996), leaf –plant hoppers (Backus and Hunter, 1989; Lett et al., 2001) and also mealy bugs (Calatayud et al. 1994). EPG has been used to study the impact of mutation Arabidopsis genes on aphid feeding behaviour and thus contribution of individual genes and mechanisms to different aspects of Arabidopsis defence and susceptibility against different aphid species (Pegadaraju et al. 2007; Kempema et al. 2007; Louis & Shah 2015; Gao et al. 2008; Nalam et al. 2012). Khan et al. (2015) reported presence of significant difference in probing and non-probing phase in aphid resistance wheat and barley genotypes in response to Russian wheat aphid. Recently, EPG technique is used as a tool to monitor the early stages of aphid resistance to insecticides (Garzo et al. 2015). Not only this, EPG is being used for acquiring and measuring electrical signals in phloem sieve elements from the wounded to unwounded area in plants (Salvador-Recatalà & Tjallingii 2015). In this study, the EPG technique is used to characterise the feeding behaviour of cabbage aphid across 16 Brassica genotypes and provide valuable information on stylet activities.

3.2.3 Chapter aim

The present study compared the feeding behaviour of *B.brassicae* across 16 *Brassica* genotypes which were selected to cover CWR and LR *Brassica* species. It also aimed to quantify both interspecies and intraspecies variations present in genotypes in response to aphid feeding on *Brassica* germplasm. The main aim was to characterize and compare feeding behaviour of *B.brassicae* using EPG technique for resistance screening in *Brassica* germplasm and determine if antibiosis mechanisms (i.e. factors in the host plant which make insect survival difficult) are important in determining differences in aphid feeding.

The results of classification from the study were used to provide molecular explanation for transcriptomics analysis in later chapters.

3.3 Material and methods

3.3.1 Plant material

The *Brassica* germplasm used in the study was obtained from Wageningen UR Plant Breeding, Wageningen University and Research Centre, The Netherlands. The germplasm collection of included both CWR and LR originating in Eastern Europe and Russia (as described in section 2.3.1). Sixteen genotypes were selected for the EPG study based on the feeding preference of cabbage aphid in field. The selection was done to cover both CWR and LR genotypes for which aphid showed high levels of preference and non-preference. As this study is part of the bigger PGR project, uniformity of genotypes used in EPG, transcriptomic and metabolomic study was maintained which also influenced the selection of genotypes. The list of crop species grown for the feeding behaviour study experiment is given in Table 3.1.

3.3.2 Plant culturing

The seeds for all 16 accessions selected from the field trial were planted individually. Ten plants (replicates) for each accession were potted individually in 10 cm pots containing 6 parts peat based compost (Humax multipurpose) to 1 part Silvaperl. The plants were grown under 16:8 hour light: dark regime at 24±3 °C with 70% humidity in the glasshouse at the University of Birmingham for 12 weeks with exception of *B.fruticulosa* which was grown for 3 weeks as it reached flowering age much earlier compared to other species. For the

EPG experiment, the aphids were placed on underside of the leaf to access the feeding behaviour of aphid.

Table 3.1: List of Brassica CWR and LR genotypes used in EPG study

S.no	Seed no.	Accession no/ cultivar name	Species	subtaxon	Material
1	24	BRA 2856	*B.incana*		CWR
2	26	K 10373	*B.incana*		CWR
3	37	K 9404	*B.montana*		CWR
4	38	BRA 1644	*B.montana*		CWR
5	199	BRA 2401	*B.oleracea*	capitata	CWR
6	321	57071	*B.oleracea*		CWR
7	325	70432	*B.oleracea*		CWR
8	397	BRA 2923	*B.villosa*		CWR
9	398	K 6926	*B.villosa*		CWR
10	401	K 10259	*B.villosa*		CWR
11	453	Bol2009-0080	*B.fruticulosa*		CWR
12	454	Bol2009-0081	*B.fruticulosa*		CWR
13	54	BRS-0103	*B.oleracea*	acephala	LR
14	116	CGN18468	*B.oleracea*	acephala	LR
15	229	HIGRU 6568	*B.oleracea*	capitata	LR
16	260	BRA 915	*B.oleracea*	capitata	LR

3.3.3 Aphid Species and infestation conditions

Plants were infested with adult *B.brassicae* aphids cultured from a clone kept at the University of Birmingham on B.nigra plants, under the conditions outlined above. *B.nigra*

plants were grown in an insect-free growth room before they were placed into Bugdorm tents for insect feeding. Tents were used to isolate aphid culture from mixing with other insects and also preventing aphid escape. The *B.nigra* plant was changed every 15 days in order to have a constant food supply and maintain aphid culture. Only apterous adult aphids were chosen for the feeding behaviour experiment.

3.3.4 Electrical Penetration Technique

The feeding behaviour of the active adult aphid selected from the culture was observed with EPG technique. As *B.brassicae* develops a waxy coating, it needs to be brushed before wiring the aphid. A gold wire 0.25mm diameter, 2-3cm long was attached to the dorsum of aphid using silver conductive paint (RS Components). The wired aphid was left to starve for about 30 minutes before being plugged into the Giga 8-CD EPG amplifier with$10^9\Omega$ input resistance. The circuit was completed by putting a copper electrode into the soil in which plant is growing. This electrode was also connected to the amplifier (Figure 3.2). All experiments were performed at room temperature (18°C) in the laboratory. The recording for 8 plants and 8 aphids were made on 8 channels simultaneously. The probing behaviour of aphid was recorded for 6 hours and the signals were recorded on the computer using EPG Stylet+d software and were subsequently analysed using EPG Stylet+ a analysis programme. For each *Brassica* genotype selected for feeding behaviour, EPG data was recorded for 10 plant and aphid combination making it 10 replicates per genotype. EPG experiments were run at the same time-points and under the same, carefully controlled conditions as the microarrays (in chapter 4), provided an opportunity to map the gene

expression data onto aphid behaviours on the plant, thus translating the molecular data collected into real world applications.

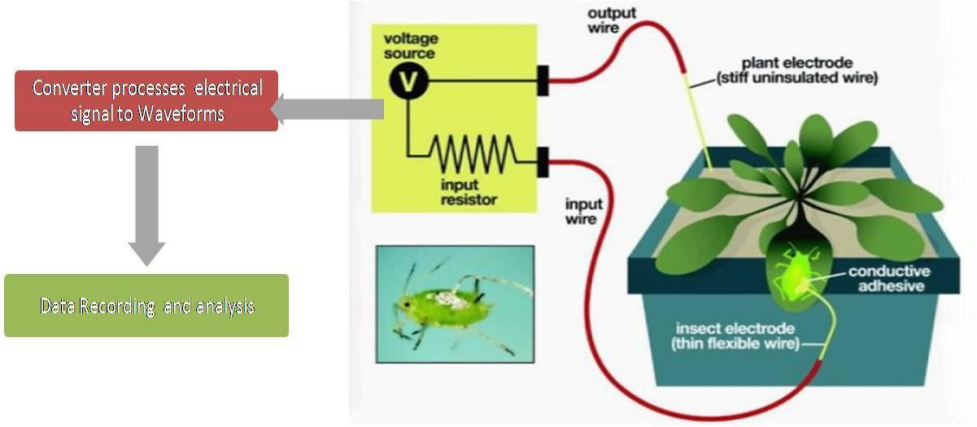


Figure 3.2 Diagram represents a typical EPG setup: The aphid is incorporated into an electrical circuit, with one electrode being attached to the aphid, while the other is inserted into the soil of its host plant's pot. When the aphid penetrates the leaf with its stylets, it completes the circuit. Output EPG waveforms are recorded and vary with the aphid's behaviour and the position of the stylet tip (Tjallingii & Gabryś 1999)

3.3.5 Statistical Analysis

The analysis of the EPG data was conducted using Microsoft Excel and Minitab software. The Anderson Darling test was performed on the data to determine if the data had a normal distribution. It is very important to carefully choose analysis method for this type of data as all parameters of EPG are not always independent of each other. One parameter can be negatively correlated to other not because they are associated but merely because there is less time for the other parameter to occur due to pre-defined time scale of recording (Cane, 1961; Sackett et.al 1978).

The EPG parameter recording was performed for 6 h as this time duration was found sufficient for aphid to commence feeding on majority of genotypes. Although cabbage aphid did not feed on accession BRA 2856 (B.incana) and K 6926 (*B.villosa*) at all during 6h time period, the possible reasons for this are discussed later in chapter. The EPG data from all 10 replicates for each genotype was averaged and data for all parameters, non-penetration, pathway phase and sap ingestion, xylem ingestion, derailed stylet and time to first E2 were analysed. The data for these parameters did not show a normal distribution, so they were analysed with a non-parametric Kruskal-Wallis analysis of variance to identify any significant differences in the behaviour observed on different *Brassica* genotypes.

3.4 Results

3.4.1 Analysis of feeding behaviour of cabbage aphid (*B.brassicae*) on different species of *Brassica*

EPG monitoring of feeding behaviour of cabbage aphid on 16 *Brassica* genotypes for period of 6 *hours* is shown in table 3.2. The EPG data analysis using Kruskal-Wallis, non parametric analysis showed that all EPG activities varied significantly between genotypes. The feeding activity of cabbage aphid on genotype BRA 2856 (B.incana) and K 6926 (B.villosa) were remarkably different when compared to rest of the genotypes. Cabbage aphid spent all 6 hours (100%) time in Non-penetration phase on these two genotypes, hence no feeding was observed. In contrast, cabbage aphids spent only 1.4 *hours (*23.6%) in non penetration when feeding on K 10259 (B.villosa) see figure 3.4. During 6 hour recording cabbage aphid spent around 25-60% time in the pathway phase on different

aphid-genotype combination. The average time spent in pathway phase was highest (3.6 *hours)* for genotype K10259, whereas it spent only 1.5 *hours* (24%) of average time when feeding on genotype BRA915 as shown in figure 3.4. The Kruskal-Wallis showed significant difference ($P \leq 0.001$) among genotypes for these parameters. Xylem feeding among different genotypes ranged from 0-10%. It was noted that cabbage aphid spent highest time (0.6 *hours*) feeding on xylem on 70432, whereas this activity was not observed in genotype BRA 2856, K6926, Bol2009-0080 and BRA 915 as shown in figure 3.4. The highest (0.5 *hours*) derailed stylet activity was observed in genotype 57071 while it was missing in genotypes BRA2856, BRA2401, 7043, K6926, BRA2923 and Bol-2009-0081. The data for all genotypes is presented in table 3.2. The phloem phase (E1, E2 and time to first (E2) will be discussed in detail in a latter section.

3.4.2 Comparison of interspecies variation in performance of B.brassicae on *Brassica* species.

EPG recordings conducted for the aphid-plant combinations revealed the differences between genotypes belonging to the same species (figure 3.4). The data from genotypes belonging to the same species were averaged to have overall picture for species, although differences were observed within the accessions as well. The Mann-Whitney U test was performed for the analysis. For all data each period of NP, Pathway and Phloem consumption, xylem all varied significantly ($P \leq 0.001$) between species.

Table 3.2: Comparison of feeding behaviour (mean ± SE of EPG parameters) of Cabbage aphid (*B.brassicae*) 16 different Brassica varieties. (Time in Seconds): Mean amount of time (*hour*s) recorded via the Electrical Penetration Graph (EPG) technique that aphids spent in different parameters on *Brassica* genotypes (a) NP (non-penetration), (b) Pathway, (c) Xylem ingestion), (d) Potential drop (Pd) and (e) Derailed stylet (F) mechanics over period of 6 hours.. N = 10 per genotype.

S.no		Genotype	NP	Pathway	E1	E2	DS	Xylem	PD	Time to first E2
1	24	BRA 2856	21600 (±0.00)	0 (±0.00)	0 (±0.00)	0 (±0.00)	0 (±0.00)	0 (±0.00)	0 (±0.00)	21600 (±0.00)
2	26	K 10373	7036.33 (±2174)	9267.11 (±1683)	148.11 (±5323)	1222.46 (±491)	1377.82 (±617)	1925.52 (±1057)	622.64 (±147)	10386 (±2950)
3	37	K 9404	8673.57 (±2083)	7824.7 (±1740)	438.3 (±105)	2658.56 (±595)	559.83 (±410)	1003.98 (±558)	441.07 (±106)	2861 (±978)
4	38	BRA 1644	13249.34 (±2745)	6991.62 (±2499)	5.7 (±5.70)	8.85 (±8.85)	390.88 (±381)	772.14 (±685)	181.47 (±78.2)	19340 (±2260)
5	54	BRS-0103	2307.88 (±505)	11909.92 (±1733)	262.64 (±54.4)	4170.62 (±1939)	1226.96 (±909)	787.47 (±553)	934.51 (±141)	1310 (±321)
6	116	CGN18468	7669.26 (±1928)	9432.38 (±1354)	74.02 (±26.7)	1369 (±686)	1439.2 (±605)	941.38 (±563)	674.75 (±121)	4404 (±1286)
7	199	BRA 2401	4357.14 (±1596)	9616.27 (±1622)	130.02 (±30.3)	5866.78 (±2145)	0 (±0.00)	881.2 (±853)	748.57 (±141)	3336 (±1758)
8	229	HIGRU 6568	8202.63 (±2053)	10038.9 (±1791)	67.82 (±32.8)	511.07 (±179)	561.67 (±435)	1948.33 (±1773)	269.55 (±102)	6333 (±2617)
9	260	BRA 915	10504.40 (±2258)	5366.60 (±1066)	319.76 (±134)	3864.53 (±1704)	965.67 (±966)	65.02 (±44.1)	514.01 (±178)	3168 (±1713)

10	321	57071	1025.26 (±662)	10248.14 (±2486)	36.75 (±9.40)	5982.18 (±3836)	1798.47 (±1137)	1738.76 (±1450)	770.44 (±241)	10510 (±3967)
11	325	70432	5757.07 (±3248)	10601.85 (±2515)	57.15 (±44.4)	2491.66 (±1357)	0 (±0.00)	2171.25 (±616)	520.99 (±194)	2566 (±1423)
12	397	BRA 2923	5962 (±1372)	8395.55 (±1376)	60.51 (±13.8)	6035.36 (±1947)	148.94 (±148)	213.64 (±212)	783.98 (±132)	2395 (±820)
13	398	K 6926	21600 (±0.00)	0 (±0.00)	0 (±0.00)	0 (±0.00)	0 (±0.00)	0 (±0.00)	0 (±0.00)	21600 (±0.00)
14	401	K 10259	5106.28 (±1993)	13041.27 (±2084)	44.43 (±11.0)	592.43 (±221)	306.84 (±291)	1613.79 (±810)	894.93 (±214)	6915 (±2599)
15	453	Bol2009-0080	3304.15 (±1772)	12115.63 (±1564)	53.55 (±12.7)	5162.38 (±1368)	153.05 (±153)	0 (±0.00)	811.23 (±172)	7808 (±2417)
16	454	Bol2009-0081	9864.8 (±2473)	6297.02 (±1811)	45.29 (±13.5)	3564.29 (±1652)	455.195 (±455)	905.35 (±905)	468.03 (±133)	7201 (±2969)

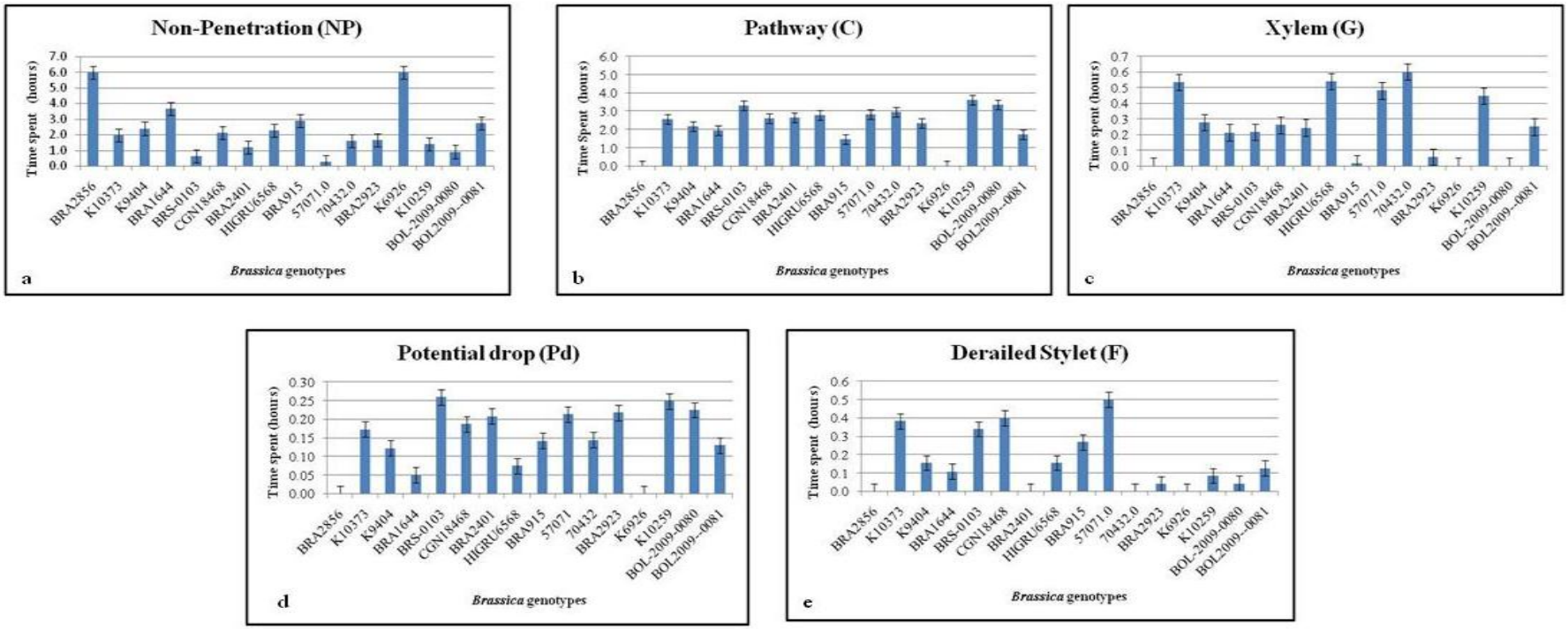


Figure 3.3: Mean amount of time (*hour*s) recorded via the Electrical Penetration Graph (EPG) technique that aphids spent in different parameters on *Brassica* genotypes (a) NP (non-penetration), (b) Pathway, (c) Xylem ingestion), (d) Potential drop (Pd) and (e) Derailed stylet (F) mechanics over period of 6 hours. Aphids species used for study was Brevicoryne brassicae over a six hour period. Plants were 12 weeks old. Aphids were removed from culture and wired straight up to the EPG system. Statistical analysis of this data can be found in table 3.2. N = 10 per genotype. Mean time spent is represented on vertical axis and Brassica genotypes are represented on horizontal axis. The data was recorded from first connection of aphid stylet with plant till end of recording period of 6hours. The data is provided in table 3.2. The Kruskal-Wallis showed significant difference ($P \leq 0.001$) among genotypes for these parameters.

Mean duration of phloem consumption was lowest (0.2 *hours)* in B.incana and highest (1.0 *hour*) in B.oleracea. The cabbage aphids spent the majority of time in NP phase in CWR species which included B.montana, B.villosa, and B.fruticulosa. This was consistent with the suggestion that CWR species are more resistant to cabbage aphid. Additionally, the time taken by aphids to reach the phloem and undertake sustained phloem consumption was highest in *Brassica oleracea* and lowest in *B.incana (*Figure 3.5).

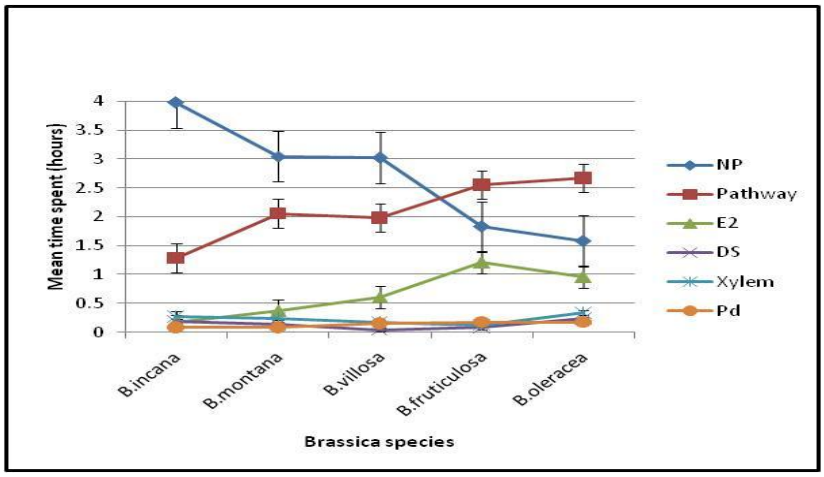


Figure 3.4: **Comparison of EPG Parameters**: Mean time spent in each EPG parameter for 5 Brassica species over period of 6 hours. Plants were 12 weeks old. Aphids were removed from culture and wired straight up to the EPG system. Statistical analysis of this data can be found in table 3.2. N = 10 per genotype. Mean time spent is represented on vertical axis and Brassica genotypes are represented on horizontal axis. The data was recorded from first connection of aphid stylet with plant till end of recording period of 6hours. The data is provided in table 3.2. The Mann-Whitney U test showed significant difference ($P \leq 0.001$) among species for these parameters

3.4.3 Phloem location and acceptance

Cabbage aphids took the shortest time (0.4 h) to locate phloem (Time to first E2) on genotype BRS-010 (Table 3.2.). It was noted that time to first E2 was similar in genotypes BRA2923 (0.7 h), 70432 (0.7 h) and K9404 (0.8 h), whereas aphids failed to

reach phloem during 6 h recording in accessions BRA2856 and K6926. The most sustained feeding on phloem was observed on genotypes BRA 2923 and 57071 (1.7 h) followed by BRA2401 (1.6 h) see figure 3.6. The cabbage aphid took 5.4 h to reach E2 on genotype BRA1644 but could not feed subsequently on phloem. In addition, cabbage aphids were unable maintain sustained feeding (E2) for more than 20 minutes on genotype HIGRU 6568 (9 min), K10259 (10 min) and K10373 (20 min). The details of time to first E2 and average duration of time spent by cabbage aphid are shown in figure 3.5.

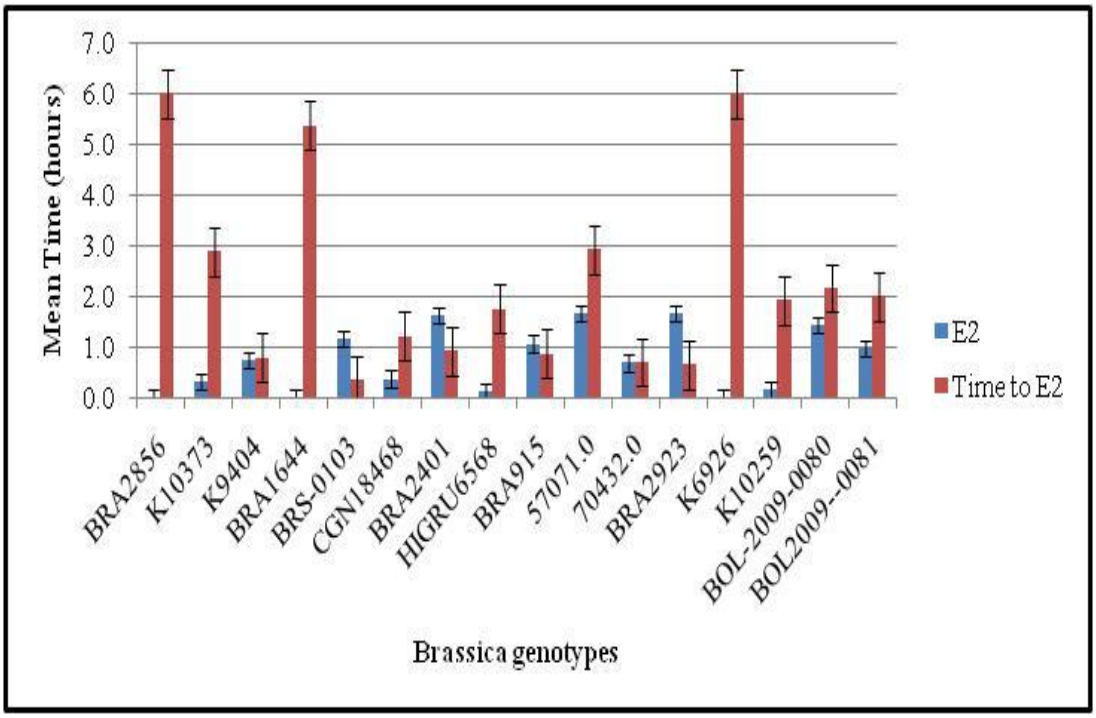


Figure 3.5: Mean time in E2 and time to E2: Mean duration of time taken by cabbage aphid to reach phloem and average time spent feeding on different genotypes. The genotypes vary significantly in response to aphid feeding on E2 and the time take to reach phloem ($P \leq 0.001$). Plants were 12 weeks old. Aphids were removed from culture and wired straight up to the EPG system. Statistical analysis of this data can be found in table 3.2. N = 10 per genotype. Mean time spent is represented on vertical axis and Brassica genotypes are represented on horizontal axis. The data was recorded from first connection of aphid stylet with plant till end of recording period of 6hours.

3.5 DISCUSSION

Understanding the resistance factors of *Brassica* crop species against cabbage aphid is crucial in plant breeding programme (Nishio 2014; Pérez-de-Castro et al. 2012). Field screening can provide information about the resistance or susceptible genotypes but lack in detailed information on how plans defend themselves against aphids (Soffan & Aldawood 2015). Therefore Electrical Penetration Graph (EPG) can be used as an alternative to study resistance factors present in plants (Will et al. 2007; Lei et al. 1998; Calatayud et al. 1994). In this study, it was noted that during 6h of feeding, differences in feeding behaviour were mostly related to duration of time the aphid spent in NP, Pathway, Phloem ingestion or Xylem phase (Figure 3.3). Additionally, time taken to reach phloem was an important criterion to determine resistance or susceptibility to aphid feeding (Figure 3.5). The results suggest that difference in feeding behaviour is highly dependent on the Brassica genotypes since the aphid population used was clonal.

The EPG technique provided important information about feeding behaviour of cabbage aphid on different genotypes of Brassica. This technique is also very useful in studying the resistance factors and their localisation in plant tissues (Tjallingii and Mayoral 1992). In this study, EGP was used detect levels of acceptability of cabbage aphid on different Brassica genotypes. Aphids come across different types of resistance during the process of feeding (Tobias Züst & Agrawal 2016). The wave forms observed during the recording time can also be grouped under three main categories i.e. surface based (NP), pathway based (C, Pd, F & G) and phloem based (E1 & E2) based upon the resistance mechanisms of the host plant.

Surface based resistance

Non penetration waveform represents the surface based resistance by indicating the level of ease/difficulty in penetrating the cell wall. The surface based resistance can be due to chemical or physical factors present in plant (Will & Van Bel 2006). Chemical defence usually involves compounds with antibiotic activity that is present on the leaf surface (Wagner et al. 2004). For instance, secondary metabolites present in trichomes of tomato prevent aphids from settling (Simmons et al. 2005). Similarly, a protein possessing lectin activity in *Arabidopsis thaliana* has an insecticidal effect towards aphids (Beneteau et al. 2010). The EPG showed some very interesting results. The genotype BRA2856 (*B.incana*) and K6926 (*B.villosa*) both showed no penetration at all during 6 h of recording period, suggesting high levels of resistance to cabbage aphid feeding at surface level in these two genotypes. This difficulty in penetrating the leaf surface of these genotypes can be accounted for by the fact that *B.villosa* (CWR) has a very dense coverage of trichomes (Palaniswamy & Bodnaryk 2012). It is reported to be resistant to insect pests due to presence of trichomes which act as a barrier to aphid settlement and feeding on plants (Nayidu et al. 2014). Interestingly, it was found to be susceptible in the field experiment (chapter 2, table 2.5). This contrast in the results may be due to the fact that EPG recording was conducted for 6 *hours* duration. Cabbage aphid may require more time to settle on this genotype, but once settled can feed on it. Another genotype BRA 2856 belonging to *B.incana* species was also previously found to be highly resistant to aphid feeding. (Ellis et al. 2000) reported that CWR *B.incana* posses a significant level of resistance against cabbage aphid. As first level of resistance, insect pests are exposed to surface chemicals imbedded in cuticular wax which may include volatile and non volatile compounds like monoterpenes and

glucosinolate-derived volatiles (Cole 1997b). The importance of glucosinolates in determining the feeding response of cabbage aphid is already a known fact (Cole 1994). In general, CWR accessions also showed comparatively higher levels of resistance as compared to the LR species which is consistent with earlier reports (Singh et al. 1994; Ellis et al. 2000; Ellis et al. 1998).

Pathway based resistance

The pathway based resistance provides the indication about the possible resistance factors present inside the plant tissue. The waveform C, Pd, derailed stylet and xylem all fall under this category as shown in figure 1.7. The results showed that cabbage aphid altered the proportion of time it spent in pathway phase over the period of 6 h feeding. It was noted that during initial settling phase more time was spent between probing and non- probing before finally having a consistent feeding activity (figure 3.3). During the settling on new host plants aphids often sample the chemical composition, and only after satisfactory test probes do they start a sustained period of pathway in order to locate the phloem (Tjallingii 1985; Powell et al. 2006). The repeated cell puncture detects the carbohydrate content and presence or absence of metabolites which help in decision to feed or not to feed on the plant (Muller & Riederer 2005). Analysis of pathway activity as shown in table 3.2 reveals variable levels of resistance between different *Brassica* genotypes. The occurrence of xylem phase is often linked with the difficulty to locate phloem. Cole (1994) noticed xylem ingestion behaviour in cabbage aphid when feeding on resistant Brassica species, where it was anticipated to play a role in osmoregulation (Spiller et al. 1990). It is during the pathway phase, the aphid also determines the level of constitutive and induced defences (Chen et al. 2013). The repeated stylet insertions cause damage to the plant and induce production of proteins,

metabolites, defence signalling and wound signalling pathways (De Ilarduya et al. 2003). After encountering these difficulties, aphid finally reaches sap of the plant which is the final destination for phloem feeding insects.

Phloem based resistance

The waveforms E1 and E2 account for time spend by the aphid in phloem feeding. Once the aphid locates phloem and starts ingestion, the composition of the phloem plays a vital role in resistance against aphid feeding (Kehr 2006; Dinant et al. 2010; Tjallingii 2006).The phloem sap is mainly made up of water, sugars, hormones and minerals. All these together determine the suitability of sap for feeding. The results indicate that the cabbage aphid has different sap feeding ability on Brassica genotypes with maximum sustained feeding genotypes BRA 2923 and 57071 (1.7 h) followed by BRA2401 (1.6 h) while failed to feed on genotype BRA 2856 and K6926 (figure 3.5). The accession 57071 was noted to be highly resistant in the field assessment (chapter 2), whereas it has found to have a sustained phloem feeding in this study. This discrepancy may be accounted by the fact that aphid may have induced the plant defence response like sap element occlusion reaction, plugging of sieve plates and callose deposition to inhibit aphid feeding as shown in other aphid-plant interaction studies (Will & van Bel 2006; Aidemark et al. 2009; Will et al. 2007).

The plant defence is a complex multi-level process. After reaching the host plant, aphid examine the surface (Ellis et al. 1996; Bruce 2015a; Walling 2000) followed by probing intercellular though plant tissue. They manipulate defence pathways activated in response to aphid feeding and wounding and finally feeding on phloem sap (Schwartzberg & Tumlinson 2014; Schwarzkopf et al. 2013; Moran & Thompson

2001). At the phloem level aphids manipulate the phloem resistance in order to sustain uninterrupted feeding. The understanding of underlying defence mechanisms has only recently begun and requires further research. EPG is an excellent tool to locate tissues most likely to play a role in resistance mechanisms (Helden & Tjallingii 1993). This technique has been extensively utilized to investigate the details of plant resistance to aphids, whiteflies and leafhoppers ((Broekgaarden et al. 2008; Janssen et al. 1989; Backus et al. 2013).

Results from this study indicate variable levels of resistance to cabbage feeding present in Brassica germplasm .Significant delay in first probe for accessions BRA 2856 and K6926 suggests presence of some deterrent resistance factor against cabbage aphid is located on plants surface for these genotypes. There was significant delay noticed in access to phloem in BRA1644, 57071, K10373 as compared to BRS-0103, BRA 2923, 70432. We can thus hypothesize that some resistance factor is located in either epidermis or mesophyll. However, increased phase of salivation not followed by phloem ingestion and reduced phloem consumption in HRIGRU 6568, K10259, K10373, and BRA1644 suggests that plant resistance is phloem based in these genotypes.

This study was designed to gain insight into the feeding behaviour of cabbage aphid and categorise the levels of resistance present in Brassica germplasm. The result from study are comparable to field assessment (chapter 2) except for accession K6926 which was susceptible (preferred) for aphid feeding, but was found highly resistant in EPG study. This difference in our study was thought to be due to short recording time (6 *hours*) as compared to that in field. The results clearly demonstrate the separation of genotypes into resistant and susceptible to aphid feeding on surface based, pathway based and

phloem based resistance mechanism. The three level approach adopted in this chapter will be further used in the transcriptomics study in chapter 4 to study the gene expression changes in theses genotypes in response to aphid feeding.

3.6 CONCLUSIONS

The feeding behaviour patterns of *B.brassicae* were studied using the EPG technique on various Brassica species and accession. The data presented demonstrate distinguishable patterns of feeding by cabbage aphid on Brassica germplasm. EPG measurement of feeding behaviour demonstrated significant differences in activity and acceptance of host plants by the cabbage aphid. Differences in feeding behaviour were observed both within accessions and also between species. This study indicates that various levels of resistance are present in the plants as part of their defence mechanism. The results showed that *B.incana, B.villosa, B.montana and B.fruticulosa* species were more resistant to aphid feeding as compared to *B.oleracea*. Another important observation was the variation in level of resistance among the accessions of the same species indicating that cabbage aphid clearly responds differently to different genotypes. The results also demonstrate that cabbage aphid applies different plant selection mechanisms depending on host plant. This result is well supported by previous reports(Powell & Hardie 2000; Goggin 2007). The results from this study strongly support and extend prior knowledge about aphid-plant interactions and successfully categorise the 16 genotypes into resistant and susceptible groups which form a firm basis for the gene expression study to look for genes playing roles in resistance against aphids in Brassicas.

CHAPTER 4

Differential gene expression profiling of diverse Brassica genotypes in response to cabbage aphid feeding using microarray analysis

4.1 ABSTRACT

Cabbage aphid (Brevicoryne brassicae) is a Brassica specialist phloem feeding insect. Insect feeding can induce substantial changes in the host plant transcriptome. Plants confront the insect with a variety of direct and indirect defence mechanisms. This study provides a never before attempted approach and scale of profiling the whole genome transcriptional response in the Brassica genotypes using Affymetrix Arabidopsis AraGene ST microarray in response to cabbage aphid feeding. The transcriptional changes in response to aphid feeding in 15 *Brassica* genotypes were compared in presence or absence of aphid feeding. *Brassica* genotypes were classified as resistant (non-preferred) and susceptible (preferred) based on response to aphid feeding in field assessment (Chapter 2) and Electrical penetration graph (Chapter 3) studies. The gene expression data analysis was done by using a novel approach to compare resistant and susceptible accessions based on field assessment, non-penetration phase (NP), Pathway phase (C), and Phloem phase (E2) parameters of Electrical Penetration Graph studies. It was noted that gene expression was highly regulated; as a different set of differentially expressing significant genes were found in all four groups in response to presence or absence of aphid feeding. The overrepresentation of gene probes responding to Gene Ontology terms such as response to stress, wounding, defensive proteins, protein kinases, cell wall modifications, glucosinolates and involvement in hormone signalling pathways indicated towards their crucial role in defence regulation in *Brassicas* during aphid attack. This approach provided an insight into the location of possible resistance factor and speculation about the role of candidate genes identified. Gene probes like NPC6 and PDCB3 are strong candidates to be tested for their specific role in surface resistance and PP2-A10 – At1g10155; At3g56240; At5g09650 for phloem-based

resistance in plants against aphid feeding. Among the important identified candidate genes, many were not previously reported and hence, form the set of novel targets which may improve our understanding of insect resistance mechanism in *Brassica* crops provided that their role is functionally verified which ultimately will be useful for *Brassica* crop improvement. To conclude, this study clearly indicates that Brassica plants respond to cabbage aphid attack at various levels and activates many defence related genes which can be specific to the location of resistance factor.

4.2 INTRODUCTION

Plants being sessile cannot escape biotic and abiotic stresses. Due to their immobility , they have evolved different defence mechanisms in response to changing environmental conditions like light, temperature, and herbivore attack (Barah et al. 2013). At the present time, due to ever increasing population, food shortage, and climatic changes, it is even more important than ever to secure our crops from biotic and abiotic factors and to develop new more resistant crops (Ahuja et al. 2010). In order to achieve this, it is crucial to understand the underlying defence mechanisms of crop plants in response to insect feeding. This has been studied for many decades (Hancock et al. 2015; Broekgaarden et al. 2007; Cole 1997a) but still a large knowledge gap is present, and a clear picture as to how the plants respond to these stresses does not exist, particularly in Brassicas.

Biotic stress like insects, bacteria, fungi, viruses are the major cause of crop damage (Gurr & Rushton 2005; Ahuja et al. 2010). Insect pests results in 30% to 80% crop yield loss worldwide (Dedryver et al. 2010; Razaq et al. 2011). Resistance to insect feeding is a complex trait involving both physical and chemical factors. The herbivory defences of plant may be present constitutively or they may be induced on insect attack (Broekgaarden, Voorrips, et al. 2011; Kessler & Baldwin 2002; Schoonhoven et al. 2005; Fürstenberg-Hägg et al. 2013). It is also known that different guilds of insects activate different defence responses in plants (Broekgaarden, Voorrips, et al. 2011; Walling 2000; Kerchev et al. 2012; Foyer et al. 2015). For example De Vos et al. (2005) have shown that phloem feeding Myzus persicae and chewing insect *Pieris rapae* larvae activate completely different transcriptional responses in *Arabidopsis thaliana*. In addition, variation in transcription response of *Gossypium hirstum* L. and *Arabidopsis*

thaliana to sap-sucking aphids and whitefly have been reported (Dubey et al. 2013; Kempema et al. 2007). This indicates that plants have highly evolved and specific defence mechanisms for specialist and generalist insects (Barthel et al. 2014; Ali & Agrawal 2012).

In general, the plant defence against aphids involves direct defence like structural barriers (trichomes, wax, and hairy leaves), toxic chemicals and secondary metabolites (alkaloids, phenolic compounds, oxidative enzymes, proteinase inhibitors), and indirect defence regulated by signalling pathways like salicylic acid (SA), jasmonic acid (JA) and ethylene which are activated in response to insect–plant interaction (Moran & Thompson 2001; Thompson & Goggin 2006; Fürstenberg-Hägg et al. 2013). The structural barriers inhibit settling of aphids on plants in the first instance, while other toxic chemicals and secondary metabolites can change aphid physiology and reduce plant suitability to aphid feeding (Bruce 2015b). A distinct glucosinolate myrosinase system is present in cruciferous plants like the model plant *Arabidopsis thaliana* as well as *Brassica* crops (Ishida et al. 2014; Del Carmen Martínez-Ballesta et al. 2013). It acts by production of toxic compounds like isothiocyanates, epithionitriles, thiocyanates, and nitriles due to hydrolysis of glucosinolates by myrosinases (Grubb & Abel 2006; Bones & Rossiter 2006; De Vos et al. 2007). But specialist aphids like *Brevicoryne brassicae,* which feed primarily on Brassica species, have developed their own protective myrosinase enzyme system to defend themselves against glucosinolates (Jones et al. 2002; Schoonhoven et al. 2005; Pratt et al. 2008; Kazana et al. 2007).

With the recent advances in the genomic technologies, it is now possible to analyse transcriptomics data and determine the underlying mechanisms and genes involved in response to insect attack. With the use of microarray technology, simultaneous

monitoring of gene expression changes of thousands of genes is possible. The changes in the gene expression of plants in response to aphid attack using transcriptional profiling have been studied (Appel et al. 2014; Thompson & Goggin 2006; Couldridge et al. 2007; Smith & Boyko 2007). A number of genes coding for cell wall modifications, oxidative stress, transcription factors, and proteins involved in defence mechanisms have been implicated (Delp et al. 2009; Hui et al. 2003; Korth 2003). In addition constitutive or induced gene expression levels of different genotypes under various stress conditions can be compared and analysed to get better understanding of plant and insect interaction (Becher et al. 2004; Wang et al. 2008; Walia et al. 2005).

Several studies using the model plants are now available which address plant and insect interaction at transcriptional level. The cell wall modification genes have been reported to have differential changes due to aphid feeding (Thompson & Goggin 2006). Not only this, salicylic acid (SA) and jasmonic acid (JA) regulated genes have been shown to have different gene expression when exposed to aphid feeding (Walling 2008). Foyer et al.(2015) reported systematic analysis of phloem-feeding insect-induced transcriptional reprogramming in *Arabidopsis* and highlighted common features that reveal distinct responses to specialist and generalist insects. Gene expression profiling of Glutathione transferase genes in Zea mays (L.) seedlings infested with cereal Aphids revealed excessive superoxide anion radicals in response to insect treatments. In recent years, the progress has been made in our understanding of plant-aphid interaction, especially the molecular bases of plant resistance and defence against aphid feeding. The development of high-throughput technologies allows us a global view of gene expression changes during plant interactions with aphids (Dong et al. 2015; Xia et al. 2014). Advances in field of omics related disciplines, have led to an exponential growth of plant genomic,

transcriptomic, proteomic and metabolomic studies (Agarwal & Narayan 2015; Mochida & Shinozaki 2011; Katam 2015; Barh et al. 2015). Liang et al. (2015) used Deep-sequencing to analyse gene expression profiles in the whole genome to identify differentially expressed genes related to aphid resistance in cucumber (*Cucumis sativus* L.). RNA sequencing technology was used by Xia et al. (2014) to have a comprehensive view of gene expression changes induced by aphid feeding in *chrysthemum morifolium.* The use of multidimensional approaches for studying plant defences against insects is highly advocated by (Barah & Bones 2015) Gase & Baldwin (2012) demonstrated NGS to be a valuable tool to identify genetic loci for ecologically relevant traits to study transcriptomic changes in the non-model plant Nicotiana attenuata elicited by abiotic factors, as well as 34 different herbivore taxa and innumerable pathogens. High-throughput quantitative proteomics studies have gained substantive importance in plant research during the last few years to characterize proteomes and their differential modulation during plant development, and biotic and abiotic stresses (Barah & Bones 2015). A proteomic study was conducted to investigate physiological factors affecting feeding behaviour by larvae of the insect Plutella xylostella on herbivore-susceptible and herbivore-resistant A. thaliana recombinant inbred lines (RILs) (Collins et al. 2010). Proteomic analysis was used in rice (Oryza sativa L.) mutants to identify differentially induced proteins during infestation by brown plant hopper by Sangha et al. (2013). In spite of all these advances, still microarray technology is a highly preferred tool to study the gene expression changes. One big advantage of microarrays is that they have been around for decades, which means researchers have developed a lot of really useful data analysis tools. Also new technology like RNA-seq is cost prohibitive specially when dealing with large number

of samples. Not only this, microarrays have been used in cross-species gene expression studies as well (Broekgaarden et al. 2008; Lu et al. 2009; Broekgaarden et al. 2007). Microarrays designed for one species can be used to measure gene expression in another species because orthologous genes are likely to share high sequence similarity, especially between closely related species. As a result, probes designed for a gene in one species are able to hybridize with its ortholog. Broekgaarden et al.(2008) successfully used Arabidopsis 70-mer oligonucleotide microarray to study transcriptional responses of cabbage aphid feeding in two Brassica cultivars. The nuclear gene expression profiles of flowers of a Brassica napus CMS line were analysed using Arabidopsis thaliana flower-specific cDNA microarrays (Carlsson et al. 2007). The use of Arabidopsis array to study gene expression in Brassica species is highly used and accepted (Gaeta et al. 2009; Nishizawa et al. 2012; Lee et al. 2004).

In the current study, a microarray analysis approach is adopted to investigate transcriptional changes in non model plants of Brassica species in response to cabbage aphid feeding using Arabidopsis gene chip.

4.2.1 Chapter aim

The main aim of study was to compare the transcriptional response of resistant (non-preferred) and susceptible (preferred) genotypes in response to cabbage aphid feeding. This study adapted a unique approach of categorising the genotypes into two distinct groups (resistant or susceptible) based on aphid performance on genotypes in the field and feeding behaviour experiments as discussed in previous chapters (Chapter 2&3). This approach not only helps in targeting the significant differentially expressed genes between two groups, but also provides opportunity to investigate different set of

transcripts identified by using different categorisations. It is expected that this approach will provide insight about the location of resistance factors in Brassicas. In addition, identification of potential candidate genes for cabbage aphid resistance which can be further investigated and used for crop improvement is proposed.

4.3 Material and methods

4.3.1 Brassica plants and cabbage aphid

The transcriptional response of aphid feeding was assessed in fifteen Brassica genotypes shown in table 4.1 selected from the field and feeding behaviour assessments. The plants were grouped as resistant and susceptible based on analysis of previous chapter 2 & 3 (Table 4.1). The preparation of plant material and insect culture was done in the same way as described previously in sections 3.3.1 to 3.3.4. Briefly, seeds from each selected accessions were planted individually in 6cm diameter pots containing 6 parts peat based compost (Humax multipurpose) to 1 part Silvaperl under 16:8 hour light: dark regime at 24±3 °C with 60-70% humidity in a plant growth room for 12 weeks. The only exception to the above regime was B.fruticulosa accessions which were grown only for 3 weeks as they began flowering much earlier compared to other accessions used in the study. The cabbage aphid was maintained on the host plant B.nigra in the insect room.

For gene expression analysis under stress conditions, four plants (replicates) from each accession received 4 clip cages containing 15 aphids each, on lower surface of two fully expanded leaves while empty clip cages were put on the control plants. After 24 h of aphid infestation, the leaf area under the clip cages was collected and four leaf discs of one plant were pooled and immediately flash frozen in liquid nitrogen and stored at −80

°C until further use. Each accession had 4 replicates each aphid exposed and control treatment.

Table 4.1: List of Brassica genotypes. The grouping into resistant (non-preferred) or susceptible (preferred) in response to aphid feeding is based on the previous field experiment (chapter 2) and feeding behaviour assessment using EPG study (Chapter 3). NP (non-penetration), Pathway, and E2 (phloem feeding) are Electrical penetration Graph parameters used in grouping. The accession 502202452 was not analyzed for feeding behaviour analysis due to insufficient number of replicates.

S.no	Seed no.	Accession no.	Species	Resistant/Susceptible			
				Field	NP	Pathway	E2
1	24	BRA 2856	*B.incana*	Resistant	Resistant	Susceptible	Resistant
2	26	K 10373	*B.incana*	Resistant	Susceptible	Susceptible	Resistant
3	37	K 9404	*B.montana*	Susceptible	Resistant	Susceptible	Susceptible
4	38	BRA 1644	*B.montana*	Resistant	Resistant	Susceptible	Resistant
5	199	BRA 2401	*B.oleracea*	Susceptible	Susceptible	Resistant	Susceptible
6	321	57071	*B.oleracea*	Resistant	Susceptible	Resistant	Susceptible
7	398	K 6926	*B.villosa*	Susceptible	Resistant	Resistant	Resistant
8	401	K 10259	*B.villosa*	Resistant	Susceptible	Resistant	Resistant
9	453	Bol2009-0080	*B.fruticulosa*	Resistant	Susceptible	Resistant	Susceptible
10	454	Bol2009-0081	*B.fruticulosa*	Susceptible	Resistant	Susceptible	Susceptible
11	54	BRS-0103	*B.oleracea*	Resistant	Susceptible	Resistant	Susceptible
12	116	CGN18468	*B.oleracea*	Resistant	Resistant	Resistant	Resistant
13	229	HRIGRU 6568	*B.oleracea*	Resistant	Resistant	Resistant	Resistant
14	260	BRA 915	*B.oleracea*	Susceptible	Resistant	Susceptible	Susceptible
15	430	502202452	*B.oleracea*	Resistant	x	x	x

4.3.2 Sample RNA Preparation

The frozen leaf samples were then processed to extract RNA. The RNA extraction was carried out using TRIzol® Reagent (Invitrogen; 15596-026). Total RNA was isolated from leaves using a modified TRIZOL extraction method as follows. Approximately 0.1 g of plant leaf tissue was ground to a powder in liquid nitrogen using mortar and pestle, resuspended without thawing in 2 ml TRIZOL reagent (Invitrogen) by vortexing, and

incubated at 65◦C for 5 min with regular mixing. Cell debris was pelleted by centrifugation (30 min, 12,000 g, 4◦C) and the supernatant was extracted twice with 3ml chloroform with the aqueous phase recovered each time after centrifugation (20 min, 12,000 × g, 4◦C). RNA was precipitated from this phase at room temperature for 5 min with 0.5 volumes isopropanol. The RNA pellet obtained after centrifugation (30 min, 12,000 × g, 4◦C) was washed with 70% ethanol, recovered again by centrifugation, air dried for 5 min and resuspended in 200µl nuclease free water. QIAGEN's RNeasy Plant Mini kit was used for further RNA purification following the manufacturer's instructions. The quality of total RNA was then determined using Nanodrop ND-1000 UV-VIS Spectrophotometer (Bio-Rad). The RNA samples having 260/280 and 260 /230 ratios above 1.8 were selected for further analysis. The 260/280 ratio corresponds to any protein, phenol or alcohol contamination whereas 260/230 ratio indicates presence of genomic DNA. The RNA samples with a minimum of 100 ng/µl with 260/280 ratio of 1.8 or above ratio were then analysed with Agilent 2100 Bioanalyser (Agilent Technologies, Santa Clara CA,USA) to check the integrity of RNA before hybridization. The integrity inspection of the RNA sample is important to determine that RNA is not degraded during the extraction process. The samples showing RIN ≥7.0 were selected for gene expression microarray experiments.

4.3.3 Minimum Information about a Microarray Experiment (MIAME)

MIAME refers to the 'Minimum Information about a Microarray Experiment' that is required to enable the interpretation of the results of the microarray experiment unambiguously and potentially to reproduce the experiment (Brazma et al. 2001). This standardized method of performing microarray experiments was initiated by the Microarray Gene Expression Database group (MGED; http://www.mged.org). The

guidelines ensure that the data generated are| correctly established for others to use for reference. The MIAME document includes all information regarding the experiment like array design, control elements, procedures used, environmental conditions, plant material etc. The MIAME detail for this study is given in Appendix II.

4.3.4 Arabidopsis vs. Brassica array

A pilot experiment was conducted to evaluate the best available Affymetrix platforms suitable for study. The two chips tested were Affymetrix Brassica Exon 1.0ST array containing 2.3 million 25-base oligonucleotides probes, as 338195 probe sets representing 135201 gene models with 15 probes per gene based on three different Brassica species (B.rapa, B.napus and B.oleracea) with additional variant contributions. The other chip used was the Affymetrix Arabidopsis Gene 1.0 ST array designed with 600941 probes as 28500 gene level probe sets with a median of 22 probes gene. It is derived from the inbred Columbia *Arabidopsis* ecotype TAIR 10 genome annotation, containing 33602 genes and 14671 gene models [GEO Platform reference:GPL17416 (Barrett et al. 2013)]. The test RNA sample and the array signals were analysed using Partek Genomics Suite version 6.6 software (Partek.Inc 2014).

The results from pilot experiment showed that Affymetrix Arabidopsis Gene 1.0 ST was more suitable for this study. Both array chips have many probe sets that have not yet been functionally annotated, but the number was much greater for the *Brassica* array as compared to *Arabidopsis* array. Most of the annotations on the Brassica array refer back to *Arabidopsis* annotations rather than independent *Brassica* annotations. The Arabidopsis chip is derived from the most updated *Arabidopsis* genome database (TAIR 10) while the Brassica array is derived from TAIR9 annotations. In addition to the lack

of annotations and difficulty in downstream analysis, the Brassica arrays are 1.5 times more costly than the *Arabidopsis* array. Considering the large number of slides needed for the experiment, annotation richness, ease of downstream data analysis and more straightforward possibility to link to other research, the Affymetrix *Arabidopsis* Gene 1.0 ST array was selected for this study.

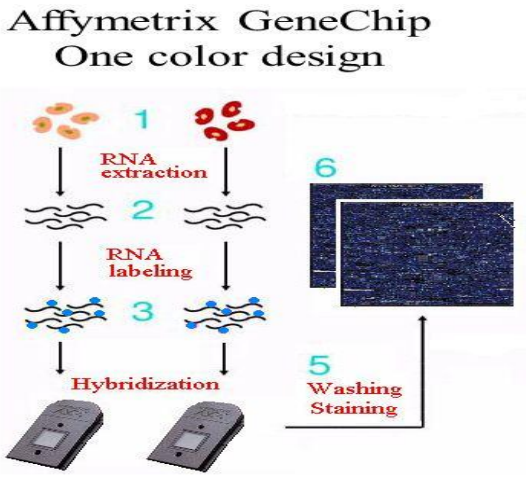


Figure 4.1: Schematic workflow for Affymetrix GeneChip one colour Microarray based gene expression experiment.

4.3.5 Microarray Hybridization

Sample preparation

The Ambion® WT Expression Kit (Ambion Inc, Austin, TX, USA) was used to prepare single stranded cDNA for labelling and hybridisation as per instructions provided by manufacturer. All starting total RNA samples were quality assessed and RNA sample having minimum 250ng/µl concentration and have RIN value >7 or above was used to cDNA.

End Labelling and hybridisation

The single stranded cDNA was end labelled and hybridized according to user manual using Affymetrix® GeneChip® WT Terminal Labelling Kit (Affymetrix Inc, Santa Clara, CA, USA. Briefly, each 5.5µg of single stranded cDNA were labelled and hybridised to probe arrays. The hybridisation was performed at 45°C at 60rpm for 17 h. Total RNA samples were processed as recommended by the manufacturer (Affymetrix, Santa Clara, CA, USA). In brief, 250 ng/µg of total RNA was reverse transcribed using SuperScript II RT (Invitrogen, Carlsbad, CA, USA) and T7-$(dT)_{24}$ primer. All the first strand cDNA was used for double-strand cDNA synthesis. Double-strand cDNA was purified by phenol–chloroform extraction and ethanol precipitation. One-half of the purified double-strand cDNA was used to generate biotin-labelled cRNA from an *in vitro* transcription reaction (IVT) using the Bio-Array High-Yield RNA Transcript Labeling Kit (Enzo Diagnostics, Farmingdale, NY, USA). The reaction product of IVT was purified using an RNeasy Mini Kit (Qiagen, Valencia, CA, USA) and quantified with a Biophotometer (Brinkmann, Westbury, NY, USA). Fifteen micrograms of fragmented cRNA was used to make 300 µL of hybridization cocktail, and 225 µL of the cocktail was used for target hybridization. The biotin-labelled targets were hybridized to GeneChip *Arabidopsis* Gene 1.0 ST Array (Affymetrix, Inc, Santa Clara, CA, USA) for 17 h at 45 °C with rotation at 60 r.p.m. in an Affymetrix GeneChip Hybridization Oven 640.

Detection

Staining, washing and scanning were performed according to GeneChip® Expression Wash, Stain and Scan user manual using the GeneChip® Hybridisation Wash and Stain Kit and Affymetrix 3000 7G scanner respectively. Scanning, signal intensity (.cel) files were then generated using the Affymetrix® Command console software.

4.3.6 Data Analysis

The Partek® Genomics Suite™ version 6.6 (Partek Inc., St. Louis, MO, USA), software was used for analysing array signal. The raw data (.CEL) files were uploaded into Partek and normalised by Robust Multichip Analysis (RMA). Normalised signal values for each sample were used to perform one way Analysis of Variance Analysis of Variance (ANOVA) or independent sample t-tests for comparisons between groups.

To analyse genome wide gene expression response(s) to aphid feeding in *Brassica* genotypes, 8 different comparisons were conducted. In each comparison, resistant vs. susceptible genotypes were compared based on either field assessment (Chapter 2) or feeding behaviour assessment (Chapter 3) in response to aphid feeding and absence of aphid feeding conditions. The main purpose of analysing the same data set in different ways is to have a clear insight of genes which get differentially regulated at different levels and provide information regarding location of insect resistance factors. This analysis resulted in 8 lists of candidate genes based on the different comparisons (Table 4.2). All candidate gene probes were selected based on a *P*-value ≤ 0.05 and a fold change in expression level of ≥ 1.5 (up-regulated genes) to ≤ -1.5 (down regulated genes). The significantly expressed gene probes from each list were then used for GO enrichment and pathway enrichment analysis using Partek® Genomics Suite™ software. The GO Enrichment Analysis was completed by setting 'Arabidopsis TAIR 10 gene model' as the reference and 'Fisher exact test' as the statistical test method, 'Hochberg False Discovery Rate (FDR)' as the multi-test adjustment method, 0.05 as the p-value cut-off and restricted analysis to functional groups with more than 2 genes as the minimum number of mapping entries. The gene probes were then annotated and defined according to the GO terms directly under the three main categories: biological

process, molecular function and cellular component. The focus of analysis was on the set of genes related to biological processes like biotic stress, response to wounding which could be linked to insect resistance.

Table 4.2: Number of genes showing significant differential gene expression patterns at p-value cut off 0.05 and fold change of ≥ 1.5 (up regulated) or ≤ -1.5 (down regulated) for comparison of accessions assessed as resistant (non-preferred) and susceptible (preferred) in the field trial and feeding behaviour experiments. The gene expression data was recorded for both presence and absence of aphid feeding for 24h. Each accession had 4 replicates.

Gene expression	Method of assessing Resistant (R) vs. Susceptible(S)	No. of samples (Accessions X reps)		No. Of genes	Up regulated	Down regulated	No. of annotated genes
		R	S				
Non Induced (Absence of aphid feeding)	Field trial	40	20	54	11	43	18
	EPG(NP based)	32	24	12	12	0	1
	EPG (Pathway Based)	32	24	153	12	141	31
	EPG(E2 Based)	28	28	94	20	74	46
Induced (Presence of aphid feeding)	Field trial	28	20	7	4	3	2
	EPG (NP based)	24	20	160	58	102	34
	EPG (Pathway Based)	28	16	105	101	4	23
	EPG (E2 Based)	20	24	143	109	34	40

4.4 Results

4.4.1 Comparison of gene expression between resistant and susceptible Brassica accessions (as determined in field trial) of brassica in the absence and presence of Cabbage Aphid feeding

Using Affymetrix Arabidopsis Gene 1.0 ST array, the transcriptional response of Brassica accessions in response to absence (non-induced) and presence (induced) of feeding by cabbage aphid was studied. A total of 54 gene probes (18 annotated shown

in table 4.3) were found to be differentially regulated in non-induced (absence of aphid feeding) plants when resistant accessions were compared to susceptible. Out of these 54 gene probes, 11 were found to be up-regulated whereas 43 were down-regulated. The complete list of gene probes with either known and unknown function or annotation is shown in supplementary table S4.1. In contrast to this, only 7 gene probes (4 up-regulated and 3 down-regulated) were found to be differentially expressed in response to cabbage aphid feeding for 24h (table 4.4) , when resistant accessions were compared to susceptible ones. Only 2 genes were found commonly expressed between both induced and non-induced conditions.

In non-induced plants, when 54 significant gene probes were further subjected to GO enrichment, it resulted in 417 GO enriched terms which were grouped under biological processes, molecular function and cellular component (shown in supplementary table S4.2). Only gene probes under biological process like response to stimulus, response to stress, developmental processes, protein metabolism, signal transduction and transport were further investigated as they could be linked with aphid resistance. The overlap between the categories was observed (see supplementary table S4.3). For example 3 out of 4 genes in response to stress category overlapped with genes in response to abiotic and biotic stimulus.

The transcript for gene AT3G142101 was found to be a part of all significantly enriched groups. AT3G14210, also known as ESM1 or epithiospecifier modifier 1, was found to be repressed in the resistant accessions when compared to susceptible ones in non induced conditions. Another gene AT1G04680 found in majority of categories was also down regulated in resistant plants. It was noted that the majority of genes present in response to biotic or abiotic stress, response to stimulus were down regulated in

resistant plants under non-induced conditions. Also AT5G03760 a glucomannan 4-beta-mannosyltransferase 9 was found to be down regulated. Among the Up-regulated genes under non induced conditions were AT5g45380 (DUR3), At5G24470 (PRR5), AT5G26570 (ATGWD3), AT2G3380 and AT3G46970 (PHS2).

The seven genes were found significant in their differential gene expression, out of them, only 2 had known annotations. These are NPC6 (AT3G48610) and MIOX2 (AT2G19800), and both these genes are involved in metabolic and cellular process. The results from GO enrichment for induced and non-induced comparison between resistant and susceptible accessions indicated the categories of genes important in response to insect attack. The comparison based on the field assessment showed the large number of significant constitutively expressed genes and only 7 significant genes in response to aphid feeding, where majority of genes were down regulated.

Table 4.3: List of differentially regulated significant probes in **non-induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **field assessment** were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation. Each genotype had 4 replicates.

S.no	Column ID	RefSeq	Locus ID (TAIR ID)	Gene Symbol	p-value	FC	gene_assignment
1	13416526	NM_179889	AT2G33830	AT2G33830	0.0000	1.78	NM_179889 // AT2G33830 // dormancy/auxin associated protein // --- // 817950 /// NM_128
2	13510863	NM_123906	AT5G45380	DUR3	0.0000	1.75	NM_123906 // DUR3 // urea-proton symporter DUR3 // --- // 834574 /// AT5G45380.1 // DUR
3	13505974	NM_122538	AT5G26570	ATGWD3	0.0000	1.61	NM_122538 // ATGWD3 // phosphoglucan, water dikinase // --- // 832706 /// NM_001125808
4	13456940	NM_114564	AT3G4697	PHS2	0.0001	1.56	NM_114564 // PHS2 // alpha-glucan phosphorylase isozyme H // --- // 823850 /// AT3G4697
5	13422411	NM_180142	AT2G47390	AT2G47390	0.0000	1.55	NM_180142 // AT2G47390 // probable glutamyl endopeptidase // --- // 819352 /// AT2G4739
6	13530385	NM_122355	AT5G24470	PRR5	0.0000	1.53	NM_122355 // PRR5 // pseudo-response regulator 5 // --- // 832518 /// AT5G24470.1 // PR
7	13521663	NM_120457	AT5G03760	ATCSLA09	0.0001	-1.52	NM_120457 // ATCSLA09 // glucomannan 4-beta-mannosyltransferase 9 // --- // 831734 ///
8	13444498	NM_111300	AT3G04290	LTL1	0.0000	-1.53	NM_111300 // LTL1 // Li-tolerant lipase 1 // --- // 819584 /// AT3G04290.1 // LTL1 // L
9	13474903	NM_119022	AT4G28780	AT4G28780	0.0001	-1.55	NM_119022 // AT4G28780 // GDSL esterase/lipase // --- // 828999 /// AT4G28780.1 // AT4G
10	13472434	NM_118481	AT4G23500	AT4G23500	0.0003	-1.57	NM_118481 // AT4G23500 // putative polygalacturonase // --- // 828450 /// AT4G23500.1 /
11	13363312	NM_100348	AT1G04680	AT1G04680	0.0001	-1.58	NM_100348 // AT1G04680 // putative pectate lyase 1 // --- // 839452 /// AT1G04680.1 //
12	13457531	NM_114720	AT3G48610	NPC6	0.0000	-1.58	NM_114720 // NPC6 // non-specific phospholipase C6 // --- // 824021 /// AT3G48610.1 //
13	13496908	NM_120408	AT5G03300	ADK2	0.0000	-1.58	NM_120408 // ADK2 // adenosine kinase 2 // --- // 831882 ///

							AT5G03300.1 // ADK2 // ade
14	13475502	AT4G30210.1	AT4G30210	ATR2	0.0001	-1.62	AT4G30210.1 // ATR2 // NADPH--cytochrome P450 reductase 2 // --- // 829144 /// AT4G3021
15	13545581	NM_112278	AT3G14210	ESM1	0.0001	-1.62	NM_112278 // ESM1 // epithiospecifier modifier 1 // --- // 820639 /// AT3G14210.1 // ES
16	13480348	NM_116322	AT4G00950	MEE47	0.0000	-1.65	NM_116322 // MEE47 // hypothetical protein // --- // 827948 /// AT4G00950.1 // MEE47 //
17	13448456	BX825917	AT3G12587	AT3G12587	0.0000	-1.92	BX825917 // AT3G12587 // Oligosaccaryltransferase // --- // 6240849
18	13366768	AT1G61880.1	AT1G61880	AT3G03847	0.0001	-1.99	AT1G61880.1 // AT3G03847 // SAUR-like auxin-responsive protein // --- // 821118 /// AT1

Table 4.4: List of differentially regulated significant probes in **Induced plants (presence of aphid feeding for 24h)** when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **field assessment** were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation. Each genotype had 4 replicates.

S.No	Column ID	RefSeq	Locus ID (TAIR ID)	Gene Symbol	p-value	Fold-Change	gene_assignment
1	13455618				0.0000	2.65	
2	13343527				0.0000	2.52	
3	13535928				0.0000	1.58	
4	13457531	NM_114720	AT3G48610.1	NPC6	0.0000	-1.58	NM_114720 // NPC6 // non-specific phospholipase C6 // --- // 824021 /// AT3G48610.1 //
5	13369773				0.0000	-1.62	
6	13447532				0.0000	-1.77	
7	13410499	NM_127538	AT2G19800.1	MIOX2	0.0000	-1.99	NM_127538 // MIOX2 // inositol oxygenase // --- // 816499 /// AT2G19800.1 // MIOX2 // i

4.4.2 **Differences in Transcriptional Response to Feeding by Cabbage Aphid (*Brevicoryne Brassicae*) comparing resistant accessions with susceptible ones based on different parameters of EPG assessment.**

The Brassica accessions were grouped into resistant or susceptible based on 3 parameters of EPG assessment; Non-Penetration, Pathway and E2 (Table 4.1).

4.4.2.1 **Gene expression analysis between resistant and susceptible accessions based on the non- probing parameter of EPG.**

The comparison between resistant and susceptible accessions based on non-penetration phase of EPG in response to presence or absence of cabbage aphid feeding was analysed. The results reveal that more genes were induced in response to aphid feeding as compared to non induced plants (Table 4.2). A total of 12 significant genes probes (all up-regulated) were found to be differentially regulated in non induced plants (Table 4.5) in comparison to 160 genes for plants exposed to aphid feeding for 24h (34 annotated gene probes shown in Table 4.6, complete list in Table S4.4). Out of 160 genes, only 58 gene probes were induced while 102 were repressed in response to cabbage aphid feeding. Only AT3G48610 (NPC6) gene probe had a defined gene annotation, so this gene list was not further subjected to GO enrichment. NPC6 was found to be significantly up regulated in non induced plants while in induced plants its expression levels were not significant.

In contrast to non induced gene expression, a higher number of genes were found expressing differentially due to aphid feeding. The 34 gene probes were subjected to GO enrichment and gene probes under biological process were further investigated (Table S4.5 & Table S4.6). The GO category response to stimulus, response to stress,

cell organization and biogenesis, developmental processes, protein metabolism, signal transduction were among highly enriched categories. The genes KAS, Delta-tip, MYB2, CPN20 were included in category of response to stress while in the case of biotic and abiotic stresses PAS2, FAD3, FAD6 were noted. Another gene probe enriched under both molecular and cellular function was PDCB3. The PDCB3 gene probe was found 1.6 fold up regulated in resistant plants in response to aphid feeding. The deposition of callose in response to wounding and abiotic stress is reported in the GO category.

Table 4.5: Complete list of differentially regulated significant probes in **non-induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Non-penetration parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

S.no	Column ID	RefSeq	Locus ID (TAIR ID)	Gene Symbol	p-value	Fold-Change	gene_assignment
1	13323795	---	---	---	0.0000	1.62	---
2	13324507	---	---	---	0.0000	1.62	---
3	13325533	---	---	---	0.0000	1.62	---
4	13330813	---	---	---	0.0000	1.62	---
5	13332045	---	---	---	0.0000	1.62	---
6	13332055	---	---	---	0.0000	1.62	---
7	13447311	---	---	---	0.0000	3.69	---
8	13457531	NM_114720	AT3G48610	NPC6	0.0000	1.51	NM_114720 // NPC6 // non-specific phospholipase C6 // --- // 824021 /// AT3G48610.1 //
9	13526012	---	---	---	0.0000	1.74	---
10	13414338	---	---	---	0.0000	3.01	---
11	13511981	---	---	---	0.0000	1.56	---
12	13490219	---	---	---	0.0000	1.59	---

Table 4.6 List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Non-penetration parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

S.No	Column ID	RefSeq	Locus ID (TAIR ID)	Gene Symbol	p-value	Fold-Change	gene_assignment
1	13460341	NM_115382	AT3G55240	AT3G55240	0.0000	2.51	NM_115382 // AT3G55240 // hypothetical protein // --- // 824690 /// AT3G55240.1 // AT3G
2	13429022	NM_112495	AT3G16240	DELTA-TIP	0.0000	2.25	NM_112495 // DELTA-TIP // aquaporin TIP2-1 // --- // 820870 /// AT3G16240.1 // DELTA-TI
3	13414696	NM_179808	AT2G29980	FAD3	0.0000	2	NM_179808 // FAD3 // omega-3 fatty acid desaturase // --- // 817548 /// AT2G29980.2 //
4	13538982	NM_124700	AT5G53210	SPCH	0.0000	1.98	NM_124700 // SPCH // transcription factor SPEECHLESS // --- // 835402 /// AT5G53210.1 /
5	13510281	NM_123769	AT5G44020	AT5G44020	0.0000	1.96	NM_123769 // AT5G44020 // HAD superfamily, subfamily IIIB acid phosphatase // --- // 83
6	13459298	NM_115154	AT3G52940	FK	0.0001	1.8	NM_115154 // FK // delta(14)-sterol reductase // --- // 824460 /// NM_202691 // FK // d
7	13458820	NM_115041	AT3G51820	G4	0.0000	1.76	NM_115041 // G4 // chlorophyll synthase // --- // 824345 /// AT3G51820.1 // G4 // chlor

8	13334904	NM_100137	AT1G02560	CLPP5	0.0003	1.72	NM_100137 // CLPP5 // ATP-dependent Clp protease proteolytic subunit 5 // --- // 839433
9	13369708	NM_101695	AT1G18370	HIK	0.0000	1.7	NM_101695 // HIK // kinesin HINKEL // --- // 838418 /// AT1G18370.1 // HIK // kinesin H
10	13369782	NM_202129	AT1G18640	PSP	0.0000	1.7	NM_202129 // PSP // phosphoserine phosphatase // --- // 838445 /// AT1G18640.2 // PSP /
11	13369792	NM_101723	AT1G18650	PDCB3	0.0000	1.65	NM_101723 // PDCB3 // plasmodesmata callose-binding protein 3 // --- // 838446 /// AT1G
12	13495802	NM_120132	AT4G39710	FKBP16-2	0.0007	1.62	NM_120132 // FKBP16-2 // FK506-binding protein 16-2 // --- // 830126 /// AT4G39710.1 //
13	13444762	NM_111372	AT3G05000	AT3G05000	0.0000	1.59	NM_111372 // AT3G05000 // transport protein particle (TRAPP) component // --- // 819661
14	13337416	NM_100725	AT1G08520	ALB1	0.0000	1.57	NM_100725 // ALB1 // magnesium-chelatase subunit chlD // --- // 837374 /// AT1G08520.1
15	13461438	NM_115674	AT3G58120	BZIP61	0.0000	1.56	NM_115674 // BZIP61 // basic-leucine zipper transcription factor family protein // ---
16	13504088	NM_001203421	AT5G20720	CPN20	0.0000	1.55	NM_001203421 // CPN20 // chaperonin 20 // --- // 832195 /// AT5G20720.3 // CPN20 // cha
17	13524399	NM_001203348	AT5G10480	PAS2	0.0000	1.54	NM_001203348 // PAS2 // very-long-chain (3R)-3-hydroxyacyl-[acyl-carrier protein] dehyd
18	13463479	NM_116156	AT3G62910	APG3	0.0000	1.52	NM_116156 // APG3 // peptide chain release factor 1 // --- // 825466 ///

							AT3G62910.1 //
19	13532040	NM_0010851 62	AT5G33370	AT5G33370	0.0004	1.52	NM_001085162 // AT5G33370 // GDSL esterase/lipase // --- // 833315 /// NM_122861 // AT5
20	13528015	NM_121875	AT5G18700	RUK	0.0000	1.52	NM_121875 // RUK // protein kinase family protein with ARM repeat domain // --- // 8319
21	13491800	NM_119243	AT4G30950	FAD6	0.0000	1.51	NM_119243 // FAD6 // omega-6 fatty acid desaturase // --- // 829220 /// AT4G30950.1 //
22	13536190	NM_123998	AT5G46290	KAS	0.0000	1.51	NM_123998 // KAS I // beta-ketoacyl-[acyl carrier protein] synthase I // --- // 834671
23	13500671	NM_121256	AT5G1218	CPK17	0.0000	-1.52	NM_121256 // CPK17 // calcium-dependent protein kinase 17 // --- // 831091 /// AT5G1218
24	13342632	NM_101911	AT1G20600	AT1G20600	0.0002	-1.53	NM_101911 // AT1G20600 // AP2/B3-like transcriptional factor family protein // --- // 8
25	13406076	NM_130287	AT2G47190	MYB2	0.0000	-1.54	NM_130287 // MYB2 // R2R3 MYB DNA binding domain transcription factor // --- // 819332
26	13545317	ATMG00770. 1	ATMG00770.1	ArthMp069	0.0002	-1.55	ATMG00770.1 // ArthMp069 // hypothetical protein // --- // 4024989
27	13545283	ATMG00470. 1	ATMG00470.1	ArthMp039	0.0000	-1.56	ATMG00470.1 // ArthMp039 // hypothetical protein // --- // 3371332
28	13392218	NM_126731	AT2G07674	AT2G07674	0.0001	-1.58	NM_126731 // AT2G07674 // hypothetical protein // --- // 815346 /// AT2G07674.1 // AT2G
29	13408306	NM_147272	AT2G07772	AT2G07772	0.0002	-1.6	NM_147272 // AT2G07772 // hypothetical protein // --- // 815359 /// AT2G07772.1 // AT2G

30	13424117	NM_111315	AT3G04440	AT3G04440	0.0001	-1.79	NM_111315 // AT3G04440 // plasma-membrane choline transporter family protein // --- //
31	13545363	ATMG01110.1	ATMG01110.1	ArthMp097	0.0002	-1.88	ATMG01110.1 // ArthMp097 // hypothetical protein // --- // 4024973
32	13545411	ATMG01410.1	ATMG01410.1	ArthMp112	0.0000	-2.45	ATMG01410.1 // ArthMp112 // hypothetical protein // --- // 4024966
33	13545128	NM_147273	AT2G07774	AT2G07774	0.0000	-2.57	NM_147273 // AT2G07774 // hypothetical protein // --- // 815361 /// AT2G07774.1 // AT2G
34	13545190	NM_147273	AT2G07774	AT2G07774	0.0000	-2.57	NM_147273 // AT2G07774 // hypothetical protein // --- // 815361 /// AT2G07774.1 // AT2G

4.4.2.2 **Gene expression analysis between resistant and susceptible accessions based on the Pathway parameter of EPG.**.

The results from the gene expression comparison between resistant and susceptible accessions under induced and non-induced conditions reveal that constitutively, the gene expression is negatively regulated (i.e. more genes are down regulated) . Out of a total of 153 significant genes differentially expressed in the absence of aphid feeding, 141 were down-regulated and only 12 were up-regulated (Table 4.7, only 13 annotated gene probes shown and complete list in table S4.7). In contrast to this, although the number of significant gene probes differentially expressed in response to aphid feeding was 105, 96.1% (101) were up-regulated and only 3.9% (4) down-regulated (Table 4.8 ,14 only annotated gene probes and complete list in table S4.10).

The GO enrichment analysis of non induced plants identified 14 significantly overrepresented groups under biological processes, cellular component and molecular function. These groups were further categorised according to functional characterisation (see table S4.8 and Table S4.9). The groups enriched under cellular component and molecular function were organelles part, membrane; cell part, binding, catalytic activity, transport activity, and electron carrier activity. In addition response to stress, response to stimulus, developmental process and reproductive process were among significantly overrepresented biological processes. On other hand the GO enrichment for induced plants showed 10 significantly overrepresented groups. The main categories enriched for biological process were response to stimulus, biological regulation and cellular process while membrane component, extracellular region, cell part, cell junction, and

organelle were overrepresented under cellular component along with catalytic activity and binding under molecular function (Table S4.11 and Table S4.12). When the groups were further analysed for non induced expression the gene probe AT5G25190 (ESE3) a member of ethylene induced response factor family, involved in ethylene activated signalling pathways was identified. Among the up-regulated genes probes AT2G07695, AT2G30600, ATMG00990 and AT2G43150 are reported to be located in the membrane or cytoplasm.

When the gene probes in induced plants were investigated most of them were annotated as hypothetical proteins. However, AT1G68760 (NUDX1) first defined nudix hydrolase in *Arabidopsis* had more than 2 fold up-regulated change in its expression in induced plants when resistant accessions were compared to susceptible ones.

Table 4.7: List of differentially regulated significant probes in **Non-Induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Pathway parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

S. no	Column ID	RefSeq	Gene Symbol	Locus ID (TAIR ID)	p-value	Fold-Change	gene_assignment
1	13392220	NM_147270	AT2G07751	AT2G07751	0.0000	2.17	NM_147270 // AT2G07751 // NADH-ubiquinone oxidoreductase chain 3 // --- // 815347 /// A
2	13545349	ATMG00990.1	ArthMp086	ATMG00990.1	0.0000	1.87	ATMG00990.1 // ArthMp086 // NADH dehydrogenase subunit 3 // --- // 3890475 /// ATMG0099
3	13392253	NM_126745	AT2G07695	AT2G07695	0.0001	1.67	NM_126745 // AT2G07695 // cytochrome C oxidase subunit II-like, transmembrane domain //
4	13398991	NM_001202713	AT2G30600	AT2G30600	0.0000	1.56	NM_001202713 // AT2G30600 // BTB/POZ domain-containing protein // --- // 817610 /// AT2
5	13503687	NM_121978	AT5G19730	AT5G19730	0.0000	1.56	NM_121978 // AT5G19730 // probable pectinesterase 53 // --- // 832093 /// AT5G19730.1 /
6	13404388	NM_129877	AT2G43150	AT2G43150	0.0000	1.51	NM_129877 // AT2G43150 // Proline-rich extensin-like family protein // --- // 818917 //
7	13530775	NM_122428	ESE3	AT5G25190	0.0000	-1.54	NM_122428 // AT5G25190 // ethylene-responsive transcription factor ERF003 // --- // 832
8	13425273	NM_111570	AT3G06895	AT3G06895	0.0003	-1.64	NM_111570 // AT3G06895 // hypothetical protein // --- // 819875 /// AT3G06895.1 // AT3G
9	13335368	NM_202776	AT4G03410	AT4G03410	0.0000	-2.24	NM_202776 // AT4G03410 // peroxisomal membrane (Mpv17/PMP22) family protein // --- // 8
10	13356880	NM_118771	AT4G26380	AT4G26380	0.0000	-2.58	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /

11	13402182	NM_114974	AT3G51140	AT3G51140	0.0000	-2.81	NM_114974 // AT3G51140 // hypothetical protein // --- // 824278 /// AT3G51140.1 // AT3G
12	13544460	NM_001203547	AT5G45720	AT5G45720	0.0000	-4.73	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
13	13445024	NM_001161121	AT3G05520	AT3G05520	0.0000	-4.76	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///

Table 4.8: List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Pathway parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

S.no	Column ID	Locus-ID (TAIR ID)	Gene Symbol	RefSeq	p-value	Fold-Change	gene_assignment
1	13545411	ATMG01410.1	ArthMp112	ATMG01410.1	0.000	2.46041	ATMG01410.1 // ArthMp112 // hypothetical protein // --- // 4024966
2	13545363	ATMG01110.1	ArthMp097	ATMG01110.1	0.000	2.36064	ATMG01110.1 // ArthMp097 // hypothetical protein // --- // 4024973
3	13428711	AT3G15534	AT3G15534	NM_001125164	0.000	2.13931	NM_001125164 // AT3G15534 // hypothetical protein // --- // 6240914 /// AT3G15534.1 //
4	13408306	AT2G07772	ATT2G07772	NM_147272	0.000	2.11543	NM_147272 // AT2G07772 // hypothetical protein // --- // 815359 /// AT2G07772.1 // AT2G
5	13452267	AT1G68760.1	NUDX1	AT1G68760.1	0.000	2.02776	AT1G68760.1 // NUDX1 // nudix hydrolase 1 // --- // 843207
6	13475949	AT2G24390	AT2G24390	NM_001202662	0.000	2.01753	NM_001202662 // AT2G24390 // AIG2-like family protein // --- // 816974 /// AT2G24390.3
7	13545317	ATMG00770.1	ArthMp069	ATMG00770.1	0.000	1.82219	ATMG00770.1 // ArthMp069 // hypothetical protein // --- // 4024989
8	13545283	ATMG00470.1	ArthMp039	ATMG00470.1	0.000	1.71183	ATMG00470.1 // ArthMp039 // hypothetical protein // --- // 3371332

9	13455565	AT3G42475	AT3G42475	NM_001125284	0.000	1.69745	NM_001125284 // AT3G42475 // hypothetical protein // --- // 6241006 /// AT3G42475.1 //
10	13392342	AT2G07787	AT2G07787	NM_201712	0.000	1.60657	NM_201712 // AT2G07787 // hypothetical protein // --- // 2745469 /// AT2G07787.1 // AT2
11	13392334	AT2G07827	AT2G07827	NM_001161035	0.000	1.55211	NM_001161035 // AT2G07827 // hypothetical protein // --- // 7922309 /// AT2G07827.2 //
12	13383470	AT1G65346	AT1G65346	NM_001160976	0.000	1.53487	NM_001160976 // AT1G65346 // hypothetical protein // --- // 7922305 /// AT1G65346.1 //
13	13427325	AT3G12760	AT3G12760	AT3G12760.1	0.000	-1.58079	AT3G12760.1 // AT3G12760 // hypothetical protein // --- // 820458 /// NM_112112 // AT3G
14	13490383	AT4G27450	AT4G27450	NM_118880	0.000	-1.61607	NM_118880 // AT4G27450 // aluminum induced protein with YGL and LRDR motifs // --- // 8

4.4.2.3 **Gene expression analysis between resistant and susceptible accessions based on the E2 parameter of EPG.**

The gene expression comparison of resistant and susceptible accessions in induced and non-induced plants showed that number of significant gene probes differentially expressed upon insect attack (143) is more in comparison with non-induced plants (94). In absence of aphid feeding, 74 genes were down-regulated in comparison with 20 up-regulated genes (30 gene probes shown in table 4.9, while supplementary table S4.13 show complete list of all significant gene probes). In contrast, more genes were found to be up-regulated (109) after exposure to aphids in resistant vs. susceptible plants (Table 4.10 and Table S4.16).

The GO enrichment analysis of the non induced group showed 548 GO terms which were characterised under biological processes, cellular component and molecular function (Supplementary Table S4.14 and S4.15). The important GO categories overrepresented under biological process category in non induced comparison were noted to be response to stimulus, single cell organism process, establishment of localisation, cellular and metabolic process, biological regulation. The response to stimulus was further investigated as this showed the highest enrichment. The gene probe AT5G09650 (PPa6) encoding a protein with pyrophosphatase activity was enriched for detection and response to biotic stimulus, glucosinolate biosynthetic process, and jasmonic acid mediated signalling pathway. The gene probe AT1G10760 (SEX1) also known as starch excess gene was found up regulated in resistant plants. In the non-induced plants, gene probe AT1G10155 (PP2-A10) was also found to be two-fold down-regulated differentially expressed gene. The Gene enrichment analysis details for the induced group are shown in supplementary table S4.17 and 4.18. The gene probe in

response to stress in induced plants AT4G13480 (IIL1) is involved in glucosinolate biosynthesis. This was found up regulated in resistant plants in response to aphid feeding in our result. The other gene probe AT4G26870 is class11 aminoacyl-tRNA and biotinsynthetase super family protein involved in regulation of meristem growth, RNA interference and virus induced gene silencing was 4.8 fold up-regulated in resistant plants in response to aphid feeding. Also AT5G23540, AT1G11840, were found up-regulated while AT3G21860 (SK10) is SKP1-like protein 10 involved in response to stress and response to misfolded protein and AT3G13750 (BGAL1) involved in carbohydrate metabolic process and regulation of protein localisation, at cell wall and plasmodesma was found down regulated. All these gene probes were enriched under response to stress in Gene Ontology terms.

Table 4.9: List of differentially regulated significant probes in **Non-Induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **E2 parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

S. no	Column ID	RefSeq	Locus ID (TAIR ID)	Gene Symbol	p-value	Fold-Change	gene_assignment
1	13366330	NM_100952	AT1G10760	SEX1	0.0004	1.89	NM_100952 // SEX1 // alpha-glucan water dikinase 1 // --- // 837619 /// AT1G10760.1 //
2	13416526	NM_179889	AT2G33830	AT2G33830	0.0000	1.88	NM_179889 // AT2G33830 // dormancy/auxin associated protein // --- // 817950 /// NM_128
3	13381189	NM_104696	AT1G60040	AGL49	0.0000	1.71	NM_104696 // AGL49 // protein agamous-like 49 // --- // 842298 /// AT1G60040.1 // AGL49
4	13444738	NM_001125102	AT3G04855	AT3G04855	0.0000	1.70	NM_001125102 // AT3G04855 // hypothetical protein // --- // 6241394 /// AT3G04855.1 //
5	13456940	NM_114564	AT3G46970	PHS2	0.0000	1.68	NM_114564 // PHS2 // alpha-glucan phosphorylase isozyme H // --- // 823850 /// AT3G4697
6	13530385	NM_122355	AT5G24470	PRR5	0.0000	1.66	NM_122355 // PRR5 // pseudo-response regulator 5 // --- // 832518 /// AT5G24470.1 // PR
7	13350898	NM_104039	AT1G51610	AT1G51610	0.0000	1.58	NM_104039 // AT1G51610 // metal tolerance protein C4 // --- // 841586 /// AT1G51610.1 /
8	13519470	NM_180940	AT5G64940	ATH13	0.0000	-1.50	NM_180940 // ATH13 // putative ABC transporter // --- // 836618 /// AT5G64940.1 // ATH1
9	13432672	NM_113324	AT3G24190	AT3G24190	0.0002	-1.51	NM_113324 // AT3G24190 // ABC1 protein kinase 6 // --- // 822005 /// AT3G24190.1 // AT3
10	13509784	NM_123638	AT5G42740	AT5G42740	0.0000	-1.53	NM_123638 // AT5G42740 // glucose-6-phosphate isomerase // --- // 834283 /// AT5G42740.
11	13545151	NM_126752	AT2G07708	AT2G07708	0.0000	-1.54	NM_126752 // AT2G07708 // hypothetical protein // --- // 815383 /// AT2G07708.1 // Arth
12	13496908	NM_120408	AT5G03300	ADK2	0.0001	-1.55	NM_120408 // ADK2 // adenosine kinase 2 // --- // 831882 /// AT5G03300.1 // ADK2 // ade
13	13532469	NM_122988	AT5G35970	AT5G35970	0.0004	-1.55	NM_122988 // AT5G35970 // P-loop containing nucleoside

							triphosphate hydrolases superfam
14	13460707	NM_115482	AT3G56240	CCH	0.0000	-1.56	NM_115482 // CCH // copper chaperone // --- // 824790 /// AT3G56240.1 // CCH // copper
15	13490658	NM_118942	AT4G28030	AT4G28030	0.0005	-1.56	NM_118942 // AT4G28030 // GCN5-related N-acetyltransferase (GNAT) family protein // ---
16	13523904	NM_121002	AT5G09650	PPa6	0.0000	-1.56	NM_121002 // PPa6 // soluble inorganic pyrophosphatase 1 // --- // 830824 /// AT5G09650
17	13361877	NM_099996	AT1G01140	CIPK9	0.0000	-1.56	NM_099996 // CIPK9 // CBL-interacting serine/threonine-protein kinase 9 // --- // 83934
18	13341008	NM_101533	AT1G16720	HCF173	0.0003	-1.58	NM_101533 // HCF173 // high chlorophyll fluorescence phenotype 173 protein // --- // 83
19	13426735	NM_111953	AT3G11170	FAD7	0.0001	-1.59	NM_111953 // FAD7 // fatty acid desaturase 7 // --- // 820288 /// AT3G11170.1 // FAD7 /
20	13390224	NM_106724	AT1G80760	NIP6;1	0.0000	-1.59	NM_106724 // NIP6;1 // aquaporin NIP6-1 // --- // 844415 /// AT1G80760.1 // NIP6;1 // a
21	13450603	NM_112631	AT3G17510	CIPK1	0.0000	-1.61	NM_112631 // CIPK1 // CBL-interacting serine/threonine-protein kinase 1 // --- // 82101
22	13529974	NM_122271	AT5G23660	MTN3	0.0000	-1.62	NM_122271 // MTN3 // bidirectional sugar transporter SWEET12 // --- // 832431 /// AT5G2
23	13429022	NM_112495	AT3G16240	DELTA-TIP	0.0003	-1.66	NM_112495 // DELTA-TIP // aquaporin TIP2-1 // --- // 820870 /// AT3G16240.1 // DELTA-TI
24	13475502	AT4G30210.1	AT4G30210	ATR2	0.0000	-1.69	AT4G30210.1 // ATR2 // NADPH--cytochrome P450 reductase 2 // --- // 829144 /// AT4G3021
25	13366768	AT1G61880.1	AT3G03847	AT3G03847	0.0001	-1.96	AT1G61880.1 // AT3G03847 // SAUR-like auxin-responsive protein // --- // 821118 /// AT1
26	13366039	NM_148455	AT1G10155	PP2-A10	0.0000	-2.01	NM_148455 // PP2-A10 // phloem protein 2-A10 // --- // 837553 /// AT1G10155.1 // PP2-A1
27	13387345	AK221799	AT5G60280	AT5G60280	0.0000	-2.08	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
28	13540082	NM_125031	AT5G56480	AT5G56480	0.0000	-2.12	NM_125031 // AT5G56480 // bifunctional inhibitor/lipid-transfer protein/seed storage 2S
29	13337838	NM_00120354	AT5G45720	AT5G45720	0.0002	-3.58	NM_001203547 // AT5G45720 // AAA-type ATPase family

		7					protein // --- // 834611 /// AT5G45
30	13445024	NM_001161121	AT3G05520	AT3G05520	0.0002	-3.58	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///

Table 4.10: List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **E2 parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

S.no	Column ID	Locus ID (TAIR ID)	Gene Symbol	RefSeq	p-value	Fold-Change	gene_assignment
1	13408386	AT2G07825	AT2G07825	NM_001124816	0.0000	6.17	NM_001124816 // AT2G07825 // hypothetical protein // --- // 6240956 /// AT2G07825.1 //
2	13474015	AT4G26870	AT4G26870	NM_118821	0.0000	4.82	NM_118821 // AT4G26870 // Asx tRNA synthetase (AspRS/AsnRS) class II core domain-contat
3	13428711	AT3G15534	AT3G15534	NM_001125164	0.0000	2.21	NM_001125164 // AT3G15534 // hypothetical protein // --- // 6240914 /// AT3G15534.1 //
4	13375015	AT1G32900.	AT1G32900	NM_103023	0.0000	1.91	NM_103023 // AT1G32900 // granule-bound starch synthase // --- // 840184 /// AT1G32900.
5	13370518	AT1G20575	AT1G20575	NM_101908	0.0009	1.91	NM_101908 // AT1G20575 // dolichol-phosphate mannosyltransferase // --- // 838646 /// A
6	13503687	AT5G19730	AT5G19730	NM_121978	0.0000	1.86	NM_121978 // AT5G19730 // probable pectinesterase 53 // --- // 832093 /// AT5G19730.1 /
7	13428254	AT3G14720	MPK19	NM_112333	0.0000	1.78	NM_112333 // MPK19 // mitogen-activated protein kinase 19 // --- // 820700 /// AT3G1472
8	13342559	AT1G20260	AT1G20260	NM_101877	0.0005	1.71	NM_101877 // AT1G20260 // V-type proton ATPase subunit B3 // --- // 838614 /// AT1G2026
9	13484009	AT4G13430	IIL1	NM_117417	0.0000	1.69	NM_117417 // IIL1 // 3-isopropylmalate dehydratase // --- // 826975 /// AT4G13430.1 //
10	13392224	AT2G07676	AT2G07676	NM_126733	0.0004	1.68	NM_126733 // AT2G07676 // hypothetical protein // ---

							// 815350 /// AT2G07676.1 // AT2G
11	13355217	AT1G64680	AT1G64680	NM_105143	0.0000	1.67	NM_105143 // AT1G64680 // hypothetical protein // --- // 842776 /// AT1G64680.1 // AT1G
12	13445595	AT3G06550	AT3G06550	NM_001084647	0.0000	1.67	NM_001084647 // AT3G06550 // O-acetyltransferase family protein // --- // 819834 /// AT
13	13524637	AT5G10946	AT5G10946	BX832040	0.0000	1.67	BX832040 // AT5G10946 // hypothetical protein // --- // 6241503
14	13355407	AT1G65260	PTAC4	NM_105199	0.0000	1.62	NM_105199 // PTAC4 // plastid transcriptionally active 4 // --- // 842833 /// AT1G65260
15	13524498	AT5G10730	AT5G10730	NM_121111	0.0001	1.61	NM_121111 // AT5G10730 // Rossmann-fold NAD(P)-binding domain-containing protein // ---
16	13465459	AT4G03950	AT4G03950	NM_116633	0.0000	1.59	NM_116633 // AT4G03950 // Nucleotide/sugar transporter family protein // --- // 825705
17	13416771	AT2G34470	UREG	NM_128999	0.0000	1.57	NM_128999 // UREG // urease accessory protein G // --- // 818010 /// NM_001036404 // UR
18	13455565	AT3G42475	AT3G42475	NM_001125284	0.0006	1.57	NM_001125284 // AT3G42475 // hypothetical protein // --- // 6241006 /// AT3G42475.1 //
19	13377872	AT1G50430	DWF5	NM_103926	0.0000	1.56	NM_103926 // DWF5 // 7-dehydrocholesterol reductase // --- // 841465 /// NM_001084224 /
20	13461010	AT3G57050	CBL	NM_180382	0.0003	1.56	NM_180382 // CBL // cystathionine beta-lyase // --- // 824872 /// AT3G57050.3 // CBL //
21	13462884	AT3G61620	RRP41	NM_001125406	0.0001	1.56	NM_001125406 // RRP41 // exonuclease RRP41 // --- // 825335 /// AT3G61620.2 // RRP41 //
22	13417589	AT2G36360	AT2G36360	NM_001036416	0.0002	1.54	NM_001036416 // AT2G36360 // galactose oxidase/kelch repeat-containing protein // --- /
23	13339000	AT1G11840	GLX1	NM_001198039	0.0007	1.53	NM_001198039 // GLX1 // glyoxalase I homolog // --- // 837731 /// AT1G11840.6 // GLX1 /
24	13402202	AT2G38080	IRX12	NM_129364	0.0006	1.53	NM_129364 // IRX12 // laccase-4 // --- // 818386 /// AT2G38080.1 // IRX12 // laccase-4
25	13479460	AT4G39210	APL3	NM_120081	0.0000	1.52	NM_120081 // APL3 // glucose-1-phosphate adenylyltransferase large subunit 3 // --- //
26	13504946	AT5G23540	AT5G23540	NM_122261	0.0000	1.52	NM_122261 // AT5G23540 // 26S proteasome non-

							ATPase regulatory subunit 14 // --- // 832
27	13446447	AT3G08530	AT3G08530	NM_111688	0.0000	1.52	NM_111688 // AT3G08530 // Clathrin, heavy chain // --- // 820001 /// AT3G08530.1 // AT3
28	13465878	AT4G05010	AT4G05010	NM_116740	0.0000	1.51	NM_116740 // AT4G05010 // F-box protein // --- // 825843 /// AT4G05010.1 // AT4G05010 /
29	13344895	AT1G26940	KDR	NM_179381	0.0001	-1.52	NM_179381 // KDR // basic helix-loop-helix protein KIDARI // --- // 839585 /// AT1G2694
30	13501752	AT5G15230	GASA4	NM_121527	0.0001	-1.52	NM_121527 // GASA4 // gibberellin-regulated protein 4 // --- // 831375 /// AT5G15230.1
31	13335137	AT1G03090	MCCA	NM_179252	0.0002	-1.58	NM_179252 // MCCA // methylcrotonoyl-CoA carboxylase subunit alpha // --- // 838362 ///
32	13473882	AT4G26530	AT4G26530	NM_001036644	0.0000	-1.58	NM_001036644 // AT4G26530 // fructose-bisphosphate aldolase 5 // --- // 828759 /// NM_1
33	13385632	AT1G70260	AT1G70260	NM_105694	0.0003	-1.59	NM_105694 // AT1G70260 // nodulin MtN21-like transporter UMAMIT36 // --- // 843362 ///
34	13494491	AT4G36780	BEH2	NM_119842	0.0000	-1.59	NM_119842 // BEH2 // BES1/BZR1-like protein 2 // --- // 829831 /// AT4G36780.1 // BEH2
35	13537386	AT5G49360	BXL1	NM_124313	0.0000	-1.62	NM_124313 // BXL1 // bifunctional {beta}-D-xylosidase/{alpha}-L-arabinofuranosidase //
36	13432048	AT3G23050	IAA7	NM_113205	0.0000	-1.66	NM_113205 // IAA7 // auxin-responsive protein IAA7 // --- // 821879 /// AT3G23050.1 //
37	13545206	AT2G07671	AT2G07671	ATMG01080.1	0.0008	-1.96	ATMG01080.1 // AT2G07671 // ATP synthase subunit 9 // --- // 815343 /// ATMG01080.1 //
38	13427803	AT3G13750	BGAL1	NM_112225	0.0000	-2.18	NM_112225 // BGAL1 // beta galactosidase 1 // --- // 820584 /// AT3G13750.1 // BGAL1 //
39	13452464	AT3G21860	SK10	NM_113081	0.0008	-2.20	NM_113081 // SK10 // SKP1-like protein 10 // --- // 821740 /// AT3G21860.1 // SK10 // S
40	13502105	AT5G16030	AT5G16030	NM_001125761	0.0001	-2.96	NM_001125761 // AT5G16030 // hypothetical protein // --- // 831460 /// NM_001036812 //
41	13392313	AT2G07722	AT2G07722	NM_126759	0.0001	-5.96	NM_126759 // AT2G07722 // hypothetical protein // --- // 815395 /// AT2G07722.1 // AT2G

4.4.3 **Comparison of transcriptional profiles in response to presence or absence of aphid feeding in distinct groups as determined by field assessment and EPG parameters**

The gene expression responses were analysed in 15 accessions of Brassica grouped as resistant (non-preferred) or susceptible (preferred) based on field assessment (Chapter 2) and EPG parameters (Chapters 3) revealed the distinct transcriptional responses based on grouping. A large variation in number of significant gene probes was noticed that were significantly different in their gene expression between all eight comparisons (Figure 4.2). This shows the overall comparison between the induced and non-induced condition when resistant accessions were compared with susceptible ones.

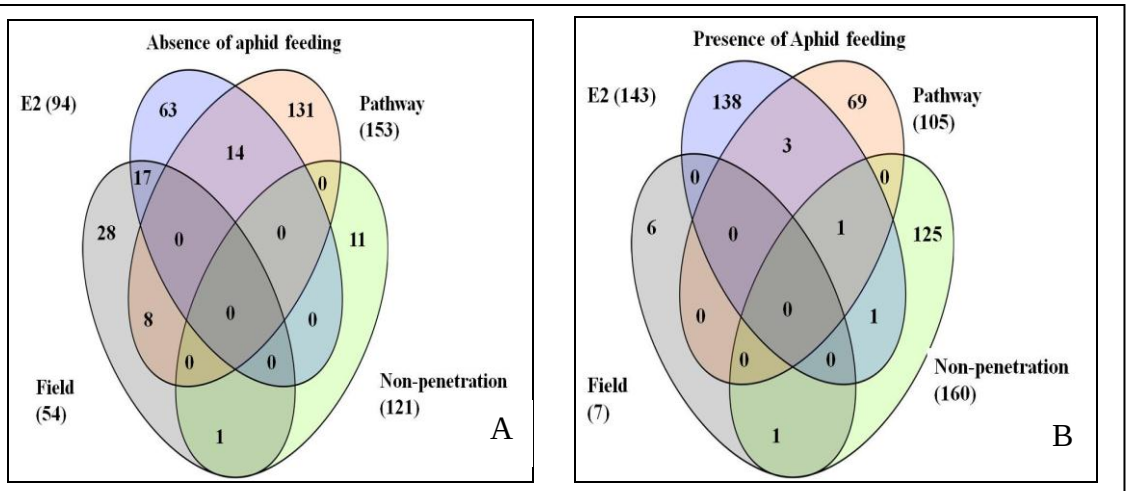


Figure 4.2: Comparison of number of genes, in response to aphid feeding in different groups when resistant (non-preferred) Brassica genotypes were compared with susceptible (preferred) genotypes. (A) Venn diagram representing the distribution of genes in absence of aphid feeding. (B) Venn diagram representing the distribution of genes in presence of aphid feeding for 24 h. The numbers in the overlapping area indicate the shared number of genes in the comparisons and include genes with an average expression ratio ≥ 1.5 -fold and a P value < 0.05 in both experiments. Numbers outside the overlapping area represent genes only in that group.

The number of significant genes expressing between resistant and susceptible genotypes in case of field assessment in absence of aphid feeding were more (54) as compared to only 7 in response to aphid feeding. The number of genes in all comparison based on EPG parameters showed that more genes were induced in response to aphid feeding when resistant genotypes were compared with susceptible ones. Only 1 gene was found common in comparison between non-penetration parameter and field assessment in absence of aphid feeding while more overlap of genes was noticed with E2 parameter (17 genes) and pathway parameter (8 genes). In contrast, none of the gene was found common in E2 parameter and field assessment in response to aphid feeding. When E2 parameter was compared to pathway parameter in absence of aphid feeding it showed 14 genes in common as compared to only 4 after aphid feeding. The comparison between pathway parameter and non-penetration parameters showed absence of any common gene expressing between these two groups in absence of aphid feeding as compared to 33 genes in presence of aphid feeding. The details are shown in figure 4.2.

4.5 Discussion

The plant defences against insect herbivores involve various morphological, biochemical and molecular mechanisms in order to overcome insect attack (War et al. 2012). In this study we evaluated the changes in transcriptional response of Brassica accessions to aphid feeding. The Affymetrix Arabidopsis Gene 1.0 ST arrays were used to compare the gene expression changes between 12 week old resistant (non-preferred) and susceptible (preferred) plants in response to presence or absence of aphid feeding. The novel approach of comparison was adapted, and as anticipated it resulted in providing insight into the transcriptional response variation depending upon the

classification group based on EPG parameters and also indicated towards the location specific gene activation.

4.5.1 *Arabidopsis* Gene 1.0 ST array microarrays are applicable to Brassica studies

This study was aimed at getting insight into the transcriptional responses of Brassica genotypes grouped as resistant (non-preferred) and susceptible (preferred) in response to aphid feeding based on the field assessment (Chapter 2) and feeding assessment using EGP parameters (Chapter 3) using full genome microarray analyses. Although, Brassica Exon 1.0 ST array were available, but they posed challenge during data analysis in the pilot experiment as described earlier in section 4.3.4 (Chapter 4). There were no enhanced annotations directly available and other published annotations referred back to Arabidopsis genome. On balance because of cost and current annotation consideration, *Arabidopsis* Gene 1.0 ST array was used for this study. The successful use of Cross-species arrays have been reported in literature (Carlsson et al. 2007; Gaeta et al. 2009; Nishizawa et al. 2012). Arabidopsis oligonucleotide microarray have been used to study gene expression changes in Brassica plants (Broekgaarden et al. 2007; Lee et al. 2004; Broekgaarden, Snoeren, et al. 2011). Overall, the gene expression was successfully studied for all 15 Brassica genotypes and further downstream analysis was done. In accordance with results mentioned above, it is expected that all species within the Brassicaceae can be analyzed with A. thaliana based microarrays. Of course, genes specific for Brassica will not be detected using these microarrays.

4.5.2 **Transcriptional profile differ in Brassica genotypes in response to aphid feeding depending on grouping criteria**

This study suggests that 24h of *B.brassicae* infestation of *Brassica* accessions led to induction of multiple transcriptional responses. The microarray results indicated that transcriptional responses to aphid feeding mainly overlapped between gene ontology groups like response to biotic stress, wounding, oxidative stress response, ethylene response pathway, cell wall modifications, cellular processes which were significantly overrepresented in the GO enrichment analysis. The difference in the gene probes for each comparison is in line with our hypothesis that different genes are activated at different locations in order to counter the aphid attack.

For example, in comparison based on field assessment, gene transcript AT3G142101 was significantly enriched in all significant gene ontology groups. This gene is already known to be involved in response to insects (https://www.arabidopsis.org/servlets/TairObject?name=AT3G14210&type=locus). AT3G14210, also known as ESMI is known as a semi-dominant QTL, https://www.arabidopsis.org/servlets/TairObject?id=39264&type=gene) which has an epistatic effect on the epithiospecifier gene. This gene represses nitrile formation and favours isothiocyanate production during glucosinolate hydrolysis, and its functional allele deters insect herbivory (Zhang et al. 2006). AT1G04680 is reported as a pectin lyase-like superfamily protein, present in extracellular region/membrane (http://www.arabidopsis.org/servlets/TairObject?type=locus&name=At1g04680). It is involved in biological processes such as plant cell wall organisation, polysaccharide biosynthesis processes, regulation of cell size etc. From the above information, we speculate that it may well be directly or indirectly affecting aphid behaviour because of

its cell size and cell wall organisation functions. The DUR3 gene is a known major transporter for high affinity urea across plasma membranes of nitrogen deficient *Arabidopsis* roots (Kojima et al. 2007). The GO annotation implicated this gene in both response to stress and response to abiotic and biotic stimulus. The comparison based on field assessment showed more constitutively expressed genes (54) as compared to only 7 in response to aphid feeding. It is possible that this variation in gene expression is due to the taxonomic differences between accessions and has no relation to aphid feeding.

The comparison between resistant and susceptible genotypes based on non-penetration parameter of EPG showed many gene transcripts which can be speculated to be linked with surface resistance. The up-regulated AT3G48610 (NPC6) is involved in hydrolase activity and it is known to be expressed during the developmental stages of a plant. Its involvement in hydrolase activity acting on ester bonds, is already established (http://www.arabidopsis.org/servlets/TairObject?type=locus&name=At3g48610). This indicates a possible role of NPC6 in response to aphid feeding in development of plants and may hinder growth of plants. The gene MYB2 encodes a transcription factor R2R3 MYB DNA binding domain which regulated expression of salt and dehydration genes. It is also reported to regulate the response of wounding in plants and has been shown to bind to calmodium (http://www.ncbi.nlm.nih.gov/gene/819332). In addition to this it is also involved in hormone mediated signalling pathways and response to abscisic acid and ethylene, both plant hormones involved in response to stress. The probe AT1G18650, known as the PDCB3 gene is assigned as plasmodesmata callose-binding protein 3, and encodes a member of the X8-GPI family of proteins known to localize to the plasmodesmata and to regulate cell -to- cell trafficking (http://www.arabidopsis.org/servlets/TairObject?type=locus&name=At1g18650).

Simpson et al. (2009) reported the role of PDCB family members PDCB1, PDCB2 and PDCB3 in accumulation of callose and its role in cell to cell communication. PDCB3 was found up-regulated in resistant plants in our study. Callose blocking of sieve tubes is a known defence of sieve tubes to physical damage. To conclude, from the assessment of transcriptional changes of resistant and susceptible plants in response to presence or absence of aphid feeding based on the non penetration parameter of EPG indicates some gene probes (NPC6 and PDCB3) may be associated with plant defence associated with surface cues of initial probing.

The time period from first stylet insertion by aphids into leaf tissue to reaching phloem sap is defined as the pathway phase. It is the time when the aphid is exploring the leaf tissue, moving stylet from one cell to another in order to locate the phloem sap. As proposed in an earlier chapter (chapter 3), presence of resistance factors to aphid feeding are present in the pathway. Aphid stylets secrete a variety of proteins including lipases, peroxidases, pectinases, and glucosidases, as they make their way through and around other cells in the leaf to reach the phloem and activate plant defence responses (Walling 2008; Miles 1999; Elzinga & Jander 2013). A massive transcriptional response was evoked in Brassica plants in response to aphid feeding when resistant genotypes were compared to susceptible ones based on pathway parameter of EPG. Down regulation of AT5G25190 (ESE3) was noticed in our study. It is located in intracellular/nucleus and aids in DNA binding and sequence specific DNA binding transcription factor (https://www.arabidopsis.org/servlets/TairObject?type=locus&name=At5g25190). It was found to be down regulated under non induced conditions. This gene is reported to play a role in response to salt induced stress in *Arabidopsis* (L. Zhang et al. 2011). Most

of the gene probes in response to aphid feeding under pathway parameter were annotated as hypothetical proteins so their values in regard to resistance are unknown. NUDX1, first defined nudix hydrolase in Arabidopsis was found to be 2 fold upregulated in resistant plants. Members of NUDX family are shown to be regulated in cellular response to biotic and abiotic stress (Ishikawa et al. 2010).

The distinct set of genes was found when comparison was done based on the E2 parameter of EPG. The changes in gene expression of resistant and susceptible accessions in response to presence and absence of aphid feeding can provide vital information regarding phloem based resistance. The gene probe AT5G09650 (PPa6) is known to have a negative regulation of defence response which supports our result as this probe was down regulated in our result also (https://www.arabidopsis.org/servlets/TairObject?id=130686&type=locus). The gene probe AT1G10760 (SEX1) gene encodes for α-glucan, a water dikinase required in starch degradation. It is reported that in SEX1 mutant, there has been 3-7 fold increase in level of starch content. This gene is up regulated in resistant plants, indicating lower starch content in resistant plants as compared to susceptible ones. A phloem tissue transport sucrose which is converted into starch in the sink cell. It is speculated that SEX1 gene up-regulation at phloem level in resistant plants may result in changes in sucrose content which lowers starch production in sink cells hence making plant less favourable to aphid feeding. Gene probe AT1G10155 is annotated as Phloem Protein 2-A10 (PP2-A10), and was enriched as part of cellular component. Phloem protein 2 (PP2) , a known conserved phloem lectin is found in abundance in phloem sap (Dinant et al. 2003). It is often believed to have a important role in the establishing phloem based defence (PBD) due to insect attacks (Kehr 2006; Dinant et al. 2010; C. Zhang et

al. 2011) and other stresses, such as wounding (Alvarez et al. 2006) and oxidative conditions. In our study PP2-A10 was found to be significantly down-regulated in absence of aphid feeding in resistant plants. Gene probe AT4G13480 (IIL1) was found up-regulated in our study. It is involved in glucosinolate biosynthesis. Glucosinolates have a well established role in response to insect herbivory (Cole 1997b; Barah et al. 2013; Mithöfer & Boland 2012; Jones et al. 2002). In resistant plants a short E2 phase (chapter 3). Glucosinolates are negatively associated to aphid performance (Kim et al. 2008) The up-regulation of IIL1 can be linked to shorted E2 phase in resistant plants.

It was noted that many growth and development related gene probes were down regulated in this study. This may be due to the fact that plants reallocate resources for defence at the cost of growth and development (Broekgaarden, Voorrips, et al. 2011). The large differences in the induced and non induced plants were found in this study. This is also reported in other studies where plant response to sap feeding insects is investigated (De Vos et al. 2005; Broekgaarden et al. 2008; Pelgrom et al. 2014). The expression of gene probes typical to response to stress (abiotic and biotic) is a common finding in response to insect feeding and they are expected to reconfigure the primary metabolism (Schwachtje & Baldwin 2008). The SNF1/AMPK/SnRK1/CPK protein kinases family are thought to mediate defence response for stress in plants (Hong & Carlson 2007; Polge et al. 2008; Crozet et al. 2014). The gene probe AT3G24190, AT1g01141and AT5G1218 belonging to CPK protein kinases family were all down regulated in our study. The SnRk1 is reported to be down regulated in response to M.persicae feeding in *Arabidopsis* (Appel et al. 2014), and also regulates relocation of photoassimilates in response to chewing herbivores in *Nicotiana attenuata* (Schwachtje et al. 2006). The gene probe AT1G10155 (PP2-A10), which is a phloem protein gene

was down regulated in non induced plants in this study. However Zhang et al. (2011) have reported reduced phloem ingestion by M.persicae as a result of PPa2-A1 in presence of Harpin-induced expression. The different in gene expression responses of phloem feeding insects M.persicae and B.brassicae on Brassica species are reported (Broekgaarden et al. 2007; Broekgaarden, Voorrips, et al. 2011). Although photosynthesis was not measured in our study but down regulation of gene probe ATR2 (AT4G30210), a NADPH-cytochrome 450 reductase2, which play role in controlling circadian rhythm in plants, was noticed. Down-regulation of photosynthesis is a common response to biotic stress and has been previously reported to also occur in response to herbivory (Giri et al. 2006; Tang et al. 2006; Bilgin et al. 2010). Some of the plant transcriptional responses to insects involved sugars (Bolouri Moghaddam & Van den Ende 2012; Morkunas & Ratajczak 2014). Previous work also has shown that herbivores can stimulate localized sink strength in the tissues they attack (Appel et al. 2014; Mewis et al. 2005). While we did not directly studied this in our study but gene expression changes in many of the gene probes linked to hexoses family support the view that hexose signalling are part of plant insect interaction. Up-regulation of genes encoding proteins involved in biosynthesis of glucosinolate in response to aphid feeding was noticed which is also reported previously (Kim et al. 2008; Barah et al. 2013).

The results from this study provided clear evidence that cabbage aphid faces resistance at different levels while feeding on the Brassica plants. The genes like NPC6 ,PDCB3 PP2-A10, IIL1, At3g56240; At5g09650, are strong candidates to be tested for their specific role against cabbage aphid resistance in plants by studying them further using more advanced “omics” technology and testing the hypothesis that they are source of surface level resistance and for phloem based resistance in plants against aphid feeding.

Recently many new studies have utilised the next generation sequencing and proteomics techniques to further evaluate the results obtained from microarray experiments. The candidate genes reported in our study need to be further studied, and doing the quantitative real time (qRT) PCR may prove useful in validating the microarray results. In addition, doing the knock-in/knock-out studies using these candidates in model plant Arabidopsis can help in confirming the role of these candidate genes.

The complexity of the situation that is described goes some way towards explaining why it has been so difficult to identify strong enduring resistance to aphids by Brassicas in breeding programmes. Any utilizable resistance in the plant breeding sense will be highly polygenic; the genes underpinning any resistance will therefore only be likely to explain relatively small percentages of the total variation in resistance and susceptibility. A number of strong candidate genes as discussed in the results section for resistance have been revealed in both induced and non-induced situations, and some of these clearly contribute in different physiological ways related to modifying feeding behaviour of the aphids. This suggests that for future crop improvement, several or many different resistance genes will need to be 'stacked' as part of the plant breeding process in order to enhance effective resistance. Shan et al. (2013) reported another very efficient approach CRISPR for genomic engineering in various organisms including rice. CRISPR provides the possibility of producing knock-outs for multiple genes and for inducing mutants for those genes, and is being developed as multiplex genome editing tool for crop plants (Sander & Joung 2014; Shan et al. 2013; Bortesi & Fischer 2014; Belhaj et al. 2015) . Seemingly it could be used for modifying the expression of a number of genes at a time to confer resistance, and could therefore be a very useful tool to demonstrate the combined effects of our candidate genes for resistance in a single

genotype of Brassica. Genomics including sequencing of the model plant and plant pathogen genome has progressed rapidly and opened several opportunities for genetic improvement of crop plants (Agarwal & Narayan 2015). Comparative genomics could help in achieving improvement of yields in rice, maize, and other related grass crops such as barley, rye, sugarcane and wheat (Mochida & Shinozaki 2011). Bioinformatics is now playing a significant role in the development of agriculture sector (Agarwal & Narayan 2015). It has opened up many avenues of utilising the vast data already available in databases to be used again to answer different biological questions and efforts are being carried out internationally to link existing related databases around the whole world.

In this study we used a variety of Brassica genotypes which included both CWR and LR. The data generated in this chapter was further evaluated to study the gene expression changes in CWR and LR in response to aphid feeding which are represented in next chapter (Chapter 5). In addition to this, it was also attempted to utilise the microarray data set generated in this study to develop a prediction model as a tool to test insect resistance in Brassica germplasm (Chapter 6). Although, the dataset used in next two chapters will be same but it will used to look at two different aspects and prove the hypothesis that same dataset can provide different results depending on the way of analysis (Tarca et al. 2006).

4.6 Conclusion

In summary, our results provide a comprehensive overview of transcriptional response of *Brevicoryne brassicae* feeding in resistant and susceptible accessions. It was noted that expression of the gene probes was overrepresented in GO terms for response to stress, response to stimulus, transport, and cell wall modification in response to aphid

feeding as indicated by GO enrichment analysis. The result from the study strongly support our hypothesis that insect feeding evokes defence responses at different levels in plants making plant defence mechanism a multilevel and complex system. The candidate gene probes specific to surface, pathway and phloem were successfully identified. The genes like NPC6, PDCB3 PP2-A10, IIL1, At3g56240; At5g09650, are strong candidates to be tested for their specific role against cabbage aphid resistance One of our very strong candidate genes for resistance (At1g10155) has already been shown to be effective in aphid resistance in *Arabidopsis* by way of a gene knock-out, and therefore it is extremely likely to play a role in phloem-based resistance in Brassica plants' defence against aphids as well. This justifies our location specific resistance approach and adds credence to our other probes indentified also being good candidates. Following on from this it could be recommended that further knock-outs in *Arabidopsis* be explored for our other strong candidates so that gene stacking could be developed. Somewhat as expected, there are no genes of major effect that have been revealed as conferring resistance to aphids in Brassica germplasm. The resistance against aphids is likely multigenic and acts by way of contrasting physiological routes. Our results would certainly support this conclusion, and therefore any advances in breeding for resistance to aphids in Brassicas should take this into account. Among the identified genes in resistant and susceptible genotypes, the genes that were not characterized before constitute the most novel target genes and would be of great interest in future to functionally validate the role of these genes which can then be used for biotechnological manipulation to improve the insect resistance in Brassica crops.

List of Supplementary tables

Table S4.1: List of differentially regulated significant probes in **non-induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **field assessment** were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation. Each genotype had 4 replicates

Table S4.2: Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **field assessment** in **non-induced** (absence of aphid feeding) plants.

Table S4.3 List of GO terms under biological process for **non-induced**, **field assessment** for comparison between resistant and susceptible accessions using TAIR's Gene Ontology annotation search and functional categorization web tool analysis.

Table S4.4: List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Non-penetration parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

Table S4.5 Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **Non-penetration parameter** of EPG assessment in **induced** (presence of aphid feeding).

Table S4.6 List of GO terms under biological process for **induced**, **non-penetration parameter of EPG** for comparison between resistant and susceptible accessions using TAIR's Gene Ontology annotation search and functional categorization web tool analysis.

Table S4.7: List of differentially regulated significant probes in **Non-Induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in

response to aphid feeding in the **Pathway parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

Table S4.8 Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **Pathway parameter** of EPG assessment in **non-induced** (absence of aphid feeding) plants.

Table S4.9: GO Enrichment analysis of significant differentially expressed genes between resistant and susceptible accession based on **pathway parameter** of EPG assessment in **non-induced** (absence of aphid feeding) plants

Table S4.10: List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **Pathway parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation

Table S4.11: Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **Pathway parameter** of EPG assessment in **induced** (presence of aphid feeding)

Table S4.12: GO Enrichment analysis of significant differentially expressed genes between resistant and susceptible accession based on **pathway parameter** of EPG assessment in induced (presence of aphid feeding) plants.

Table S4.13: List of differentially regulated significant probes in **Non-Induced (absence of aphid feeding)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **E2 parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. P < 0.05;

Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

Table S4.14 Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **E2 parameter** of EPG assessment in **non-induced** (absence of aphid feeding) plants.

Table S4.15: GO Enrichment analysis of significant differentially expressed genes between resistant and susceptible accession based on **E2 parameter** of EPG assessment in **non-induced** (absence of aphid feeding) plants.

Table S4.16: List of differentially regulated significant probes in **Induced (presence of aphid feeding for 24h)** plants when resistant (non-preferred) and susceptible (preferred) genotypes in response to aphid feeding in the **E2 parameter** of EPG experiment were compared. Table shows Affymetrix column Id, Refseq ID, Locus ID, Gene Symbol, p-value, fold-change (FC) and gene assignment according to available annotations or BLAST search ;. $P < 0.05$; Benjamini-Hochberg false discovery correction applied. A positive FC indicates that infestation causes up regulation; a negative FC indicates that infestation causes down regulation.

Table S4.17 Functional characterisation of significant differentially expressed genes between resistant and susceptible accession based on **E2 parameter** of EPG assessment in **induced** (presence of aphid feeding) plants.

Table S4.18: GO Enrichment analysis of significant differentially expressed genes between resistant and susceptible accession based on **E2 parameter** of EPG assessment in **induced** (presence of aphid feeding) plants.

CHAPTER 5

Transcriptional response of Brassica wild relatives and landraces to infestation by cabbage aphid (B.brassicae)

5.1 Abstract

Crop domestication is one of the most important processes in human history. In order to meet the feeding needs of human populations, farmers started selecting from wild species which resulted in landrace varieties. Both crop wild relatives (CWR) and landraces (LR) are very important resources for crop improvement and can provide resistance against abiotic and biotic stress. In this chapter the gene expression differences and changes that may have resulted from domestication between Brassica CWR and LR were studied in response to presence and absence of aphid (*B.brassicae*) feeding.

The results show that the comparison of induced and non-induced gene expression of CWR and LR is highly affected by aphid feeding. A larger number of genes (261) were found significantly differentially expressed in response to aphid feeding in contrast to only 75 genes expressed differentially in the absence of aphid feeding. It was also noticed that there was a greater number of up-regulated genes involved in response to aphid feeding in CWR as compared to LR, which is in line with the domestication hypothesis that selection for higher yield, and desired characters results in decreased plant defence.

In addition, up-regulated genes like PTR3, NAC083, TIFY7, HCHIB, RAP2.10 and ERD6 are indicated as strong candidates for resistance against aphid feeding in *Brassica* and are found in CWR. It is strongly suggested that identified candidate genes should be further investigated to confirm their association with aphid feeding in future.

5.2 Introduction

In today's world our crops are affected by many abiotic and biotic stresses which threaten the world's food supplies. The wild relatives of our crops (CWR) and the local varieties, grown traditionally and adapted to local environments (landraces) are considered as useful resources for crop improvement. Their value for food security is widely recognized (Hyten et al. 2006; Chaudhary 2013).

The origin of landraces is the result of the domestication process of wild crop species over several hundred years for desired traits which were selected and grown every season. This resulted in area specific varieties with traits which separate them from their wild relatives. Mostly these crops were grown and maintained by farmers themselves. As a result of this highly selective cultivation, the majority of plant species showed changes in a variety of traits like enhanced yield, fruit size, reduced seed dormancy, and relative susceptibility to pathogens and insect pests (Burger et al. 2008; Hyten et al. 2006; Tanksley & McCouch 1997). This is thought to be maintained by predicted resource allocation tradeoffs between the benefits of reduced herbivory and the costs of resistance and differential selection pressures. The landrace crops have now become a very important source of germplasm which can be used in crop improvement by crop breeders. However, the selection process has narrowed the genetic diversity in crop species which is also assumed to be an important reason for the reduced resistance of crop plants to important traits like resistance to biotic and abiotic stresses (Chaudhary 2013).

Crop wild relatives are an important and beneficial germplasm resource for modern agriculture. They provide plant breeders with a potentially useful gene pool. That CWR

genes can be used to improve crop performance is an established fact supported by many examples dating back more than 60 years (Hajjar & Hodgkin 2007; Ford-Lloyd et al. 2011; Nevo 2014; Tew & Bekeko 2014). CWR can have beneficial traits relating to both biotic and abiotic stress. Studies have been conducted to search for extended gene pools for pest and disease resistance that can be transferred to agricultural crops (Brar & Khush 1997; Brar & Khush 2002). Disease resistance genes have been reported in tomato wild relatives and used in commercial cultivars (Rick & Chetelat 1995). In millet *Pyricularia grisea* resistance was introgressed from its wild relative *Pennisetum glaucum subsp. monodii* (Burton & Wilson 1995). Similarly in bananas, resistance against Black Sigatoka caused by the fungus *Mycosphaerella fijiensis* has been developed from its wild relative *Musa acuminata* most commonly known as "Calcutta 4" ((Hajjar & Hodgkin 2007) The wild relatives of crop Brassicas (*Brassica fruticulosa Cyr., B. spinescens Pomel., B. insularis Moris. and B. villosa Biv*) have been reported to possess resistance against cabbage aphid, *Brevicoryne brassicae* (Ellis et al. 2000).

There is much evidence that transcriptional studies are helpful in finding the genes which can prove useful in generating more resistant crops to both biotic and abiotic stresses (Redden 2013; Yumurtaci 2015). With this evidence of success, the current study was conducted to look at differences in gene expression changes in response to aphid feeding between the CWR and LR of Brassicas. To study these changes the Affymetrix array was used and gene expression levels of wild relatives of *Brassica* i.e. *B.incana, B.montana, B.villosa* were compared with landrace *B.oleracea* accessions in response to presence and absence of aphid feeding. The data set obtained in chapter 4 is evaluated again in this chapter in order to explore the gene expression variation present between CWR and LR in response to aphid feeding. Although the data set is the same,

in chapter 4 the whole data set was evaluated in relation to feeding behaviour parameters from chapter 3 as criteria for classification of genotypes. In this chapter (chapter 5) a direct comparison of cabbage aphid feeding response between CWR and LR grouped as resistant based on feeding preference in the field (chapter 2) and EPG analysis (chapter 3) is explored. The susceptible accessions were not included due to very small numbers of replicates available.

5.2.1 Chapter aim

This study aims at investigating the differences in the gene expression of CWR and LR of Brassica germplasm under control conditions (absence of aphid) and following exposure to aphids for 24 h. At the end of the study we will have identified a list of significantly differentially expressed genes between resistant CWR and resistant LR under control and stress conditions and would have identified candidate resistance genes to cabbage aphid using the microarray approach.

5.3 Material and methods

These are the same as previously described in chapter 4 (section 4.3).

5.4 Results

5.4.1 Determination of significantly differentially expressed genes between CWR and LR using ANOVA.

The gene expression data for 4 CWR and 4 LR Brassica accessions were analyzed using Partek® Genomics Suite™ software. The Analysis of Variance was used to obtain a list of genes that were significantly differentially expressed between CWR and LR after 24 h of aphid exposure and absence of aphid exposure. All samples were subjected to the

same statistical parameters to identify the constitutive and aphid-induced gene expression patterns. The results from the one-way ANOVA produced a spreadsheet containing all 38,408 genes. This consists of the overall statistical results from the experiment which were further reduced to a list of genes that pass statistical criteria of significant difference in expression at FDR corrected p- value < 0.05 and fold change 1.5 to ≤ -1.5 selected to cover the maximum number of genes showing change in expression.

5.4.2 Comparison of gene expression between CWR and LR accessions of Brassica in the absence of Cabbage Aphid

All Brassica accessions used in this experiment were untreated i.e. no aphid infestation in order to identify the differences in gene expression between wild and cultivated germplasm in absence of aphid feeding.

From the available 38,408 gene probes on the Affymetrix Arabidopsis Gene 1.0 ST array, at FDR corrected p- value < 0.05 and fold change ≥1.5 to ≤ -1.5 only 75 genes were found that were constitutively differentially expressed between CWR and LR. The complete list of up and down regulated gene probes along with gene symbols, fold change and gene assignment is provided in table 5.1and S5.1. Out of 75 significant gene probes 41 were up regulated in CWR while 34 were down regulated in CWR as compared to LR. The fold changes varied between 22.82 fold up regulated for ATMG01220 to -3.39 fold down regulated probe ID (13451950) with unknown gene ID.

Out of 75 significantly expressed gene probes, 32 with available locus ID were then used for GO enrichment analysis using Partek® Genomics Suite™ software as

mentioned previously in section 4.36. The gene probes were then annotated and defined according to the GO terms directly under the three main categories: biological process, molecular function and cellular component. Further analysis was focused on the set of genes related to biological processes like biotic stress, response to wounding which could be linked to insect resistance. The detailed information of functional characterisation and Go enrichment is shown in Table S5.2 and S5.3

These differentially expressed significant genes were functionally categorised as majorly involved in cellular and metabolic processes, transport, response to stress, response to biotic or abiotic stimulus, electron transport or energy pathways, developmental processes, protein metabolism, DNA dependent transcription etc. Additionally genes involved in cell organisation and biogenesis were down regulated in CWR when compared to LR. The list of genes under biological processes with detailed gene assignment is shown in supplementary table S5.2

5.4.2.1 Genes involved in Transport

Four genes **AT1G08930** (ERD6), **AT1G15210** (PDR7), **AT5G24030** (SLAH3), and **ATG46050** (PTR3) were found differentially expressing between CWR and LR under GO category of gene involved in transport. Out of these 3 genes **AT1G08930**, AT1G1520, ATG46050 were up regulated in CWR while **AT5G24030** was found to be down regulated.

5.4.2.2 Response to Stress and response to stimulus

Five genes (**AT1G08930, AT3G03780, AT5G13180, AT5G4650 and AT5G46050**) were categorised as response to stress. These 5 genes were enriched in both response to stress and response to stimulus category in GO enrichment analysis. All these genes

were found to be up regulated in CWR when compared with the LR. The gene **AT1G08930** (ERD6) is known to be involved in sugar transport in plants. The genes **AT3G03780** (MS2), encodes cytosolic methionine synthase and is involved in methionine regeneration via activated methyl cycle**. AT5G13180** (NACO83) is known to negatively regulate xylem vessel formation, and **AT5G46050** (PTR3) up regulated in CWR, and encodes for a di- and tri-peptide transporter involved in response to wounding (Karim et al. 2005). Karim et al. (2007) reported that AtPTR3 gene mutant had increased susceptibility to virulent pathogenic bacteria *Erwinia carotovora* subsp. *corotovora* and *Pseudomonas syringae* and protects plants against biotic and abiotic stresses (Karim et al. 2007). The 1.9 fold up regulation of PTR3 gene in CWR in the current study may be speculated as one of the reasons for higher levelsof protection to wounding present in wild plant species as compared to LR.

5.4.2.3 Genes involved in Signal Transduction and transcription

The gene **AT5G25190** was enriched in both signal transduction and transcriptional response categories. It was found to be 1.57 fold up-regulated in CWR when compared to LR in the absence of aphid feeding. In addition to this **AT5G13180** (NAC083) was also found 2.4 fold up-regulated. The details of all GO enrichments is provided in the supplementary table S5.2

Table 5.1 Differentially Expressed genes: Number of genes with significant different in expression at FDR corrected p- Value < 0.05 and fold change ≥1.5 to ≤ -1.5 for comparison of resistant Brassica Crop Wild relative (CWR) and Landraces (LR) in absence of aphid feeding.

S.No	Transcript ID	Gene Symbol	RefSeq	Locus ID	Fold-Change	Fold-Change (Description)	gene_assignment
1	13524842	ASP3	NM_121190	AT5G11520	1.70	Up-regulated	NM_121190 // ASP3 // aspartate aminotransferase 3 // --- // 831024 /// AT5G11520.1 // A
2	13343847	AT1G23465	NM_202168	AT1G23465	2.22	Up-regulated	NM_202168 // AT1G23465 // peptidase-S24/S26 domain-containing protein // --- // 2745760
3	13383271	AT1G64850	NM_105159	AT1G64850	1.64	Up-regulated	NM_105159 // AT1G64850 // calcium-binding EF-hand-containing protein // --- // 842793 /
4	13545217	AT2G07696	NM_126746	AT2G07696	1.62	Up-regulated	NM_126746 // AT2G07696 // ribosomal protein S7 // --- // 815372 /// AT2G07696.1 // rps7
5	13408365	AT2G07721	NM_126758	AT2G07721	4.12	Up-regulated	NM_126758 // AT2G07721 // hypothetical protein // --- // 815394 /// AT2G07721.1 // AT2G
6	13437647	AT3G50123	NM_001125335	AT3G50123	1.55	Up-regulated	NM_001125335 // AT3G50123 // hypothetical protein // --- // 6240873 /// AT3G50123.1 //
7	13442656	AT3G62510	NM_116116	AT3G62510	1.60	Up-regulated	NM_116116 // AT3G62510 // protein disulfide isomerase-like protein // --- // 825425 ///
8	13530775	AT5G25190	NM_122428	AT5G25190	1.57	Up-regulated	NM_122428 // AT5G25190 // ethylene-responsive transcription factor ERF003 // --- // 832

9	13374992	CID11	NM_001036054	AT1G32790	1.74	Up-regulated	NM_001036054 // CID11 // CTC-interacting domain 11 protein // --- // 840173 /// AT1G327
10	13536092	DELTA-OAT	NM_123987	AT5G46180	1.65	Up-regulated	NM_123987 // DELTA-OAT // ornithine-delta-aminotransferase // --- // 834660 /// AT5G461
11	13337731	ERD6	NM_001035929	AT1G08930	1.64	Up-regulated	NM_001035929 // ERD6 // sugar transporter ERD6 // --- // 837414 /// NM_100765 // ERD6 /
12	13423873	MS2	NM_180176	AT3G03780	1.61	Up-regulated	NM_180176 // MS2 // methionine synthase 2 // --- // 821147 /// NM_001125092 // MS2 // m
13	13500896	NAC083	NM_121321	AT5G13180	2.05	Up-regulated	NM_121321 // NAC083 // NAC domain containing protein 83 // --- // 831157 /// AT5G13180.
14	13545387	nad1	ATMG01220.1	ATMG01220	22.82	Up-regulated	ATMG01220.1 // nad1 // NADH dehydrogenase subunit 1 // --- // 3890477
15	13368314	PDR7	AT1G15210.1	AT1G15210	1.50	Up-regulated	AT1G15210.1 // PDR7 // ABC transporter G family member 35 // --- // 838087 /// NM_10138
16	13536042	PTR3	NM_123973	AT5G46050	1.95	Up-regulated	NM_123973 // PTR3 // peptide transporter 3 // --- // 834646 /// AT5G46050.1 // PTR3 //
17	13374451	RUB1	NM_102873	AT1G31340	1.52	Up-regulated	NM_102873 // RUB1 // ubiquitin-NEDD8-like protein RUB1 // --- // 840023 /// AT1G31340.1
18	13345523	AT1G28307	NM_001123899	AT1G28307	-1.56	Down-regulated	NM_001123899 // AT1G28307 // hypothetical protein // --- // 6241164 /// AT1G28307.1 //

19	13347647	AT1G34041	NM_001123935	AT1G34041	-1.51	Down-regulated	NM_001123935 // AT1G34041 // hypothetical protein // --- // 6241199 /// AT1G34041.1 //
20	13385899	AT1G70949	AT1G70949.1	AT1G70949	-1.92	Down-regulated	AT1G70949.1 // AT1G70949 // hypothetical protein // --- // 6240811 /// NM_001124109 //
21	13392218	AT2G07674	NM_126731	AT2G07674	-1.98	Down-regulated	NM_126731 // AT2G07674 // hypothetical protein // --- // 815346 /// AT2G07674.1 // AT2G
22	13421794	AT2G45860	NM_130149	AT2G45860	-1.78	Down-regulated	NM_130149 // AT2G45860 // hypothetical protein // --- // 819194 /// AT2G45860.1 // AT2G
23	13431841	AT3G22415	NM_202624	AT3G22415	-1.67	Down-regulated	NM_202624 // AT3G22415 // hypothetical protein // --- // 2745889 /// AT3G22415.1 // AT3
24	13380596	CYP96A15	NM_104570	AT1G57750	-1.55	Down-regulated	NM_104570 // CYP96A15 // alkane hydroxylase CYP96A15 // --- // 842150 /// AT1G57750.1 /
25	13353669	E1	NM_104682	AT1G59900	-1.54	Down-regulated	NM_104682 // E1 ALPHA // pyruvate dehydrogenase complex E1 alpha subunit // --- // 8422
26	13407392	EXPA15	NM_126361	AT2G03090	-1.54	Down-regulated	NM_126361 // EXPA15 // expansin A15 // --- // 814838 /// AT2G03090.1 // EXPA15 // expan
27	13355059	GH9C2	NM_105114	AT1G64390	-1.94	Down-regulated	NM_105114 // GH9C2 // glycosyl hydrolase 9C2 // --- // 842747 /// AT1G64390.1 // GH9C2
28	13339000	GLX1	NM_001198039	AT1G11840	-1.74	Down-regulated	NM_001198039 // GLX1 // glyoxalase I homolog // --- // 837731 /// AT1G11840.6 // GLX1 /

29	13369782	PSP	NM_202129	AT1G18640	-1.50	Down-regulated	NM_202129 // PSP // phosphoserine phosphatase // --- // 838445 /// AT1G18640.2 // PSP /
30	13401542	SBE2.1	NM_129196	AT2G36390	-1.52	Down-regulated	NM_129196 // SBE2.1 // 1,4-alpha-glucan branching enzyme 2-1 // --- // 818212 /// AT2G3
31	13530101	SLAH3	NM_122308	AT5G24030	-1.56	Down-regulated	NM_122308 // SLAH3 // SLAC1 homologue 3 // --- // 832468 /// AT5G24030.1 // SLAH3 // SL
32	13356178	URE	NM_105422	AT1G67550	-1.50	Down-regulated	NM_105422 // URE // urease // --- // 843076 /// AT1G67550.1 // URE // urease // --- //

5.4.3 **Comparison of gene expression between resistant CWR and resistant LR accessions of Brassica after 24 hours of Cabbage Aphid feeding.**

A total of 261 genes were differentially expressed when resistant CWR were compared to resistant LR after 24 hours of aphid feeding. Out of these, 102 genes were found to be induced (up-regulated) while 159 genes were repressed (down regulated) in CWR. All significantly differentially expressed genes are listed in table 5.2and S5.4. These genes were further categorised according to their functional categorisations of biological process, cellular component and molecular function. Still a large set of genes were not studied as there were no annotations available for these genes. The fold changes varied between 43.92 fold up regulated in gene ATMG01220 (nad1) to -10.65 fold down regulated in an annotated gene. Out of 261 genes, 142 annotated genes were then used for GO enrichment analysis using Partek® Genomics Suite™ software as described earlier. The focus was mainly on genes enriched under the biological categories like response to stress, response to biotic stress, response to wounding etc. which can be linked to aphid infestation. On functional categorisation of loci under biological process, 70 genes encoding for cellular and 70 genes for metabolic process were found, 8 genes encoding for transport, 12 for developmental processes, 25 for response to stress, 27 for response to abiotic or biotic stress , 14 for response to protein metabolism, 10 for transcription and DNA-dependant process, 7 for signal transduction and 3 for electron transport pathway along with 3 genes involved in DNA or RNA metabolism (Table S5.5). The details of functional categorizations and GO enrichment is provided in supplementary data table S5.5 and Table S5.6

Table 5.2 Differentially Expressed genes: Number of genes with significant different in expression at FDR corrected p- Value < 0.05 and fold change ≥1.5 to ≤ -1.5 for comparison of resistant Brassica Crop Wild relative (CWR) and Landraces (LR) in presence of aphid feeding.

S.no	Transcript ID	Gene Symbol	RefSeq	Locus ID	Fold-Change	Fold-Change(Description)	Gene_assignment
1	13457204	ABCA2	NM_114641	AT3G47730	1.52	Up-regulated	NM_114641 // ABCA2 // ABC transporter A family member 2 // --- // 823927 /// AT3G47730.
2	13528308	ACL5	NM_001085136	AT5G19530	1.84	Up-regulated	NM_001085136 // ACL5 // Thermospermine synthase ACAULIS5 // --- // 832073 /// AT5G19530
3	13545283	ArthMp039	ATMG00470.1	ATMG00470	1.62	Up-regulated	ATMG00470.1 // ArthMp039 // hypothetical protein // --- // 3371332
4	13545359	ArthMp091	ATMG01040.1	ATMG01040	1.84	Up-regulated	ATMG01040.1 // ArthMp091 // hypothetical protein // --- // 4024969
5	13545411	ArthMp112	ATMG01410.1	ATMG01410	1.51	Up-regulated	ATMG01410.1 // ArthMp112 // hypothetical protein // --- // 4024966
6	13524842	ASP3	NM_121190	AT5G115200	1.64	Up-regulated	NM_121190 // ASP3 // aspartate aminotransferase 3 // --- // 831024 /// AT5G11520.1 // A
7	13342911	AT1G21245	NM_148476	AT1G21245	1.67	Up-regulated	NM_148476 // AT1G21245 // Protein kinase superfamily protein // --- // 838720 /// AT1G2
8	13343847	AT1G23465	NM_202168	AT1G23465	3.21	Up-regulated	NM_202168 // AT1G23465 // peptidase-S24/S26 domain-containing protein // --- // 2745760
9	13354650	AT1G63220	NM_105001	AT1G63220	1.51	Up-regulated	NM_105001 // AT1G63220 // calcium-dependent lipid-binding domain-containing protein //
10	13356629	AT1G68570	NM_105528	AT1G68570	1.67	Up-regulated	NM_105528 // AT1G68570 // putative nitrite transporter // --- // 843186 /// AT1G68570.1
11	13384976	AT1G68570	NM_105528	AT1G68570	1.57	Up-regulated	NM_105528 // AT1G68570 // putative nitrite transporter // --- // 843186 /// AT1G68570.1
12	13358274	AT1G72290	NM_105888	AT1G72290	1.83	Up-regulated	NM_105888 // AT1G72290 // trypsin inhibitor (Kunitz) domain-containing protein // --- /

13	13386441	AT1G72540	NM_105913	AT1G72540	1.94	Up-regulated	NM_105913 // AT1G72540 // putative receptor-like protein kinase // --- // 843586 /// AT
14	13407824	AT2G04650	NM_126494	AT2G04650	1.67	Up-regulated	NM_126494 // AT2G04650 // ADP-glucose pyrophosphorylase-like protein // --- // 815007 /
15	13408365	AT2G07721	NM_126758	AT2G07721	11.11	Up-regulated	NM_126758 // AT2G07721 // hypothetical protein // --- // 815394 /// AT2G07721.1 // AT2G
16	13392342	AT2G07787	NM_201712	AT2G07787	1.53	Up-regulated	NM_201712 // AT2G07787 // hypothetical protein // --- // 2745469 /// AT2G07787.1 // AT2
17	13411645	AT2G22370	NM_127802	AT2G22370	1.61	Up-regulated	NM_127802 // AT2G22370 // hypothetical protein // --- // 816769 /// AT2G22370.1 // AT2G
18	13401971	AT2G37440	NM_129299	AT2G37440	1.71	Up-regulated	NM_129299 // AT2G37440 // DNAse I-like superfamily protein // --- // 818321 /// AT2G374
19	13404388	AT2G43150	NM_129877	AT2G43150	1.53	Up-regulated	NM_129877 // AT2G43150 // Proline-rich extensin-like family protein // --- // 818917 //
20	13428400	AT3G14990	NM_001035621	AT3G14990	1.55	Up-regulated	NM_001035621 // AT3G14990 // protein DJ-1-like A // --- // 820728 /// NM_112361 // AT3G
21	13452781	AT3G23090	NM_113210	AT3G23090	1.54	Up-regulated	NM_113210 // AT3G23090 // targeting protein for Xklp2-like protein // --- // 821884 ///
22	13432310	AT3G23510	NM_113254	AT3G23510	1.54	Up-regulated	NM_113254 // AT3G23510 // cyclopropane-fatty-acyl-phospholipid synthase // --- // 82193
23	13453160	AT3G24180	NM_001035680	AT3G24180	1.64	Up-regulated	NM_001035680 // AT3G24180 // Beta-glucosidase, GBA2 type family protein // --- // 82200
24	13457337	AT3G48030	NM_114672	AT3G48030	1.56	Up-regulated	NM_114672 // AT3G48030 // RING-H2 finger protein ATL48 // --- // 823958 /// AT3G48030.1
25	13437332	AT3G49130	NM_114772	AT3G49130	1.52	Up-regulated	NM_114772 // AT3G49130 // SWAP (Suppressor-of-White-APricot)/surp RNA-binding domain-co
26	13460997	AT3G57020	NM_115561	AT3G57020	1.73	Up-regulated	NM_115561 // AT3G57020 // strictosidine synthase family protein // --- // 824869 /// AT

27	13480634	AT4G01700	NM_116400	AT4G01700	1.73	Up-regulated	NM_116400 // AT4G01700 // Chitinase family protein // --- // 828131 /// AT4G01700.1 //
28	13465352	AT4G03420	NM_116580	AT4G03420	1.52	Up-regulated	NM_116580 // AT4G03420 // hypothetical protein // --- // 827928 /// AT4G03420.1 // AT4G
29	13465878	AT4G05010	NM_116740	AT4G05010	1.57	Up-regulated	NM_116740 // AT4G05010 // F-box protein // --- // 825843 /// AT4G05010.1 // AT4G05010 /
30	13485148	AT4G15885	NM_148338	AT4G15885	1.56	Up-regulated	NM_148338 // AT4G15885 // kinesin motor protein-like protein // --- // 827270 /// AT4G1
31	13485638	AT4G16765	NM_179213	AT4G16765	1.55	Up-regulated	NM_179213 // AT4G16765 // oxidoreductase, 2OG-Fe(II) oxygenase family protein // --- //
32	13477038	AT4G33440	NM_119498	AT4G33440	1.66	Up-regulated	NM_119498 // AT4G33440 // pectin lyase-like superfamily protein // --- // 829481 /// AT
33	13529672	AT5G22860	NM_180728	AT5G22860	1.58	Up-regulated	NM_180728 // AT5G22860 // Serine carboxypeptidase S28 family protein // --- // 832349 /
34	13542294	AT5G61820	AT5G61820.1	AT5G61820	1.54	Up-regulated	AT5G61820.1 // AT5G61820 // hypothetical protein // --- // 836304 /// NM_125576 // AT5G
35	13543310	AT5G64250	NM_125821	AT5G64250	1.62	Up-regulated	NM_125821 // AT5G64250 // Aldolase-type TIM barrel family protein // --- // 836546 ///
36	13537205	ATSDI1	NM_124262	AT5G48850	2.02	Up-regulated	NM_124262 // ATSDI1 // protein SULPHUR DEFICIENCY-INDUCED 1 // --- // 834943 /// AT5G48
37	13484924	BAM5	NM_117609	AT4G15210	1.78	Up-regulated	NM_117609 // BAM5 // beta-amylase 5 // --- // 827185 /// AT4G15210.1 // BAM5 // beta-am
38	13363863	CALS1	NM_100436	AT1G05570	1.58	Up-regulated	NM_100436 // CALS1 // callose synthase 1 // --- // 837059 /// AT1G05570.1 // CALS1 // c
39	13374992	CID11	NM_001036054	AT1G32790	1.65	Up-regulated	NM_001036054 // CID11 // CTC-interacting domain 11 protein // --- // 840173 /// AT1G327

40	13420535	CSY3	NM_129840	AT2G42790	1.71	Up-regulated	NM_129840 // CSY3 // citrate synthase 3 // --- // 818879 /// AT2G42790.1 // CSY3 // cit
41	13536092	DELTA-OAT	NM_123987	AT5G46180	1.77	Up-regulated	NM_123987 // DELTA-OAT // ornithine-delta-aminotransferase // --- // 834660 /// AT5G461
42	13505359	DMR6	NM_122361	AT5G24530	2.36	Up-regulated	NM_122361 // DMR6 // downy mildew resistance 6 protein / oxidoreductase // --- // 83252
43	13337731	ERD6	NM_001035929	AT1G08930	1.63	Up-regulated	NM_001035929 // ERD6 // sugar transporter ERD6 // --- // 837414 /// NM_100765 // ERD6 /
44	13545581	ESM1	NM_112278	AT3G14210	4.03	Up-regulated	NM_112278 // ESM1 // epithiospecifier modifier 1 // --- // 820639 /// AT3G14210.1 // ES
45	13418497	FAC1	NM_129384	AT2G38280	1.59	Up-regulated	NM_129384 // FAC1 // AMP deaminase // --- // 818408 /// NM_179963 // FAC1 // AMP deamin
46	13543347	FKBP12	NM_125831	AT5G64350	1.53	Up-regulated	NM_125831 // FKBP12 // peptidyl-prolyl isomerase FKBP12 // --- // 836556 /// AT5G64350.
47	13448399	HCHIB	NM_112085	AT3G12500	1.77	Up-regulated	NM_112085 // HCHIB // basic chitinase B // --- // 820429 /// AT3G12500.1 // HCHIB // ba
48	13402202	IRX12	NM_129364	AT2G38080	1.64	Up-regulated	NM_129364 // IRX12 // laccase-4 // --- // 818386 /// AT2G38080.1 // IRX12 // laccase-4
49	13456324	IVD	NM_114399	AT3G45300	1.66	Up-regulated	NM_114399 // IVD // isovaleryl-CoA-dehydrogenase // --- // 823668 /// AT3G45300.1 // IV
50	13459262	MDAR1	NM_001161202	AT3G52880	1.53	Up-regulated	NM_001161202 // MDAR1 // monodehydroascorbate reductase (NADH) // --- // 824454 /// AT3
51	13454169	MUB4	NM_113612	AT3G26980	1.64	Up-regulated	NM_113612 // MUB4 // membrane-anchored ubiquitin-fold protein 4 // --- // 822315 /// AT
52	13344465	MYB116	NM_102344	AT1G25340	1.67	Up-regulated	NM_102344 // MYB116 // putative transcription factor (MYB116) // --- // 839118 /// AT1G

53	13518334	NAC101	NM_125632	AT5G62380	1.56	Up-regulated	NM_125632 // NAC101 // NAC-domain transcription factor // --- // 836359 /// AT5G62380.1
54	13545387	nad1	ATMG01220.1	ATMG01220	43.92	Up-regulated	ATMG01220.1 // nad1 // NADH dehydrogenase subunit 1 // --- // 3890477
55	13528833	NPH4	NM_180715	AT5G20730	1.62	Up-regulated	NM_180715 // NPH4 // auxin-regulated transcriptional activator NPH4 // --- // 832196 //
56	13448873	PA200	NM_112178	AT3G13330	1.56	Up-regulated	NM_112178 // PA200 // proteasome activating protein 200 // --- // 820533 /// AT3G13330.
57	13494916	PCK1	NM_119948	AT4G37870	1.64	Up-regulated	NM_119948 // PCK1 // phosphoenolpyruvate carboxykinase [ATP] // --- // 829943 /// AT4G3
58	13368314	PDR7	AT1G15210.1	AT1G15210	1.59	Up-regulated	AT1G15210.1 // PDR7 // ABC transporter G family member 35 // --- // 838087 /// NM_10138
59	13429722	PYD1	NM_112662	AT3G17810	1.56	Up-regulated	NM_112662 // PYD1 // putative dihydropyrimidine dehydrogenase // --- // 821049 /// AT3G
60	13478537	RAP2.10	NM_119854	AT4G36900	1.53	Up-regulated	NM_119854 // RAP2.10 // ethylene-responsive transcription factor RAP2-10 // --- // 8298
61	13357574	TIFY7	NM_105738	AT1G70700	1.54	Up-regulated	NM_105738 // TIFY7 // protein TIFY 7 // --- // 843407 /// AT1G70700.1 // TIFY7 // prote
62	13361730	WRKY40	NM_106732	AT1G80840	1.77	Up-regulated	NM_106732 // WRKY40 // putative WRKY transcription factor 40 // --- // 844423 /// AT1G8
63	13519410	WRKY51	NM_125877	AT5G64810	1.66	Up-regulated	NM_125877 // WRKY51 // putative WRKY transcription factor 51 // --- // 836602 /// AT5G6
64	13386963	ALPHA	NM_106027	AT1G73680	-1.56	Down-regulated	NM_106027 // ALPHA DOX2 // alpha dioxygenase // --- // 843703 /// AT1G73680.1 // ALPHA
65	13367035	AT1G12423	NM_001123802	AT1G12423	-1.80	Down-regulated	NM_001123802 // AT1G12423 // hypothetical protein // --- // 6240731 /// AT1G12423.1 //

66	13374726	AT1G32080	NM_102942	AT1G32080	-1.53	Down-regulated	NM_102942 // AT1G32080 // plastidal glycolate/glycerate translocator 1 // --- // 840100
67	13378599	AT1G52220	AT1G52220.1	AT1G52220	-1.91	Down-regulated	AT1G52220.1 // AT1G52220 // hypothetical protein // --- // 841652 /// AT1G52220.2 // AT
68	13385899	AT1G70949	AT1G70949.1	AT1G70949	-2.07	Down-regulated	AT1G70949.1 // AT1G70949 // hypothetical protein // --- // 6240811 /// NM_001124109 //
69	13386385	AT1G72430	NM_105902	AT1G72430	-1.68	Down-regulated	NM_105902 // AT1G72430 // SAUR-like auxin-responsive protein family // --- // 843575 //
70	13415609	AT2G31830	NM_128741	AT2G31830	-1.55	Down-regulated	NM_128741 // AT2G31830 // Type II inositol-1,4,5-trisphosphate 5-phosphatase 14 // ---
71	13437797	AT3G05520	NM_001161121	AT3G05520	-5.76	Down-regulated	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///
72	13445024	AT3G05520	NM_001161121	AT3G05520	-5.77	Down-regulated	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///
73	13387889	AT3G05520	NM_001161121	AT3G05520	-5.83	Down-regulated	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///
74	13458322	AT3G05520	NM_001161121	AT3G05520	-5.83	Down-regulated	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///
75	13425033	AT3G06470	NM_111522	AT3G06470	-1.73	Down-regulated	NM_111522 // AT3G06470 // GNS1/SUR4 membrane protein family // --- // 819824 /// AT3G06
76	13544974	AT3G20362	AT3G20362.1	AT3G20362	-1.62	Down-regulated	AT3G20362.1 // AT3G20362 // hypothetical protein // --- // 821581 /// NM_148737 // AT3G
77	13453771	AT3G25880	NM_113491	AT3G25880	-1.99	Down-regulated	NM_113491 // AT3G25880 // hypothetical protein // --- // 822183 /// AT3G25880.1 // AT3G
78	13435143	AT3G43570	NM_114225	AT3G43570	-1.90	Down-regulated	NM_114225 // AT3G43570 // GDSL esterase/lipase // --- // 823453 /// AT3G43570.1 // AT3G

79	13460341	AT3G55240	NM_115382	AT3G55240	-3.60	Down-regulated	NM_115382 // AT3G55240 // hypothetical protein // --- // 824690 /// AT3G55240.1 // AT3G
80	13471375	AT4G20940	NM_118212	AT4G20940	-1.54	Down-regulated	NM_118212 // AT4G20940 // leucine-rich receptor-like protein kinase // --- // 827842 //
81	13487592	AT4G21520	NM_118272	AT4G21520	-1.60	Down-regulated	NM_118272 // AT4G21520 // WD40 repeat family protein // --- // 828237 /// AT4G21520.1 /
82	13346598	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
83	13356880	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
84	13395867	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
85	13411603	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
86	13453704	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
87	13489860	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
88	13489862	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
89	13501968	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
90	13476724	AT4G32710	NM_119424	AT4G32710	-1.63	Down-regulated	NM_119424 // AT4G32710 // proline-rich receptor-like protein kinase PERK14 // --- // 37
91	13476949	AT4G33080	NM_119462	AT4G33080	-1.52	Down-regulated	NM_119462 // AT4G33080 // AGC (cAMP-dependent, cGMP-dependent and protein kinase C) kin
92	13522706	AT5G06530	NM_001036766	AT5G06530	-1.67	Down-regulated	NM_001036766 // AT5G06530 // ABC transporter G family member 22 // --- // 830541 /// AT

93	13525208	AT5G12323	NM_001125743	AT5G12323	-1.50	Down-regulated	NM_001125743 // AT5G12323 // hypothetical protein // --- // 6240742 /// AT5G12323.1 //
94	13526694	AT5G15780	NM_121583	AT5G15780	-1.64	Down-regulated	NM_121583 // AT5G15780 // pollen Ole e 1 allergen and extensin family protein // --- //
95	13532237	AT5G35480	NM_122939	AT5G35480	-2.65	Down-regulated	NM_122939 // AT5G35480 // hypothetical protein // --- // 833512 /// AT5G35480.1 // AT5G
96	13510281	AT5G44020	NM_123769	AT5G44020	-1.95	Down-regulated	NM_123769 // AT5G44020 // HAD superfamily, subfamily IIIB acid phosphatase // --- // 83
97	13337838	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
98	13370341	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
99	13494665	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
100	13531589	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
101	13544460	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
102	13517400	AT5G60280	AK221799	AT5G60280	-4.38	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
103	13387345	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
104	13411454	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
105	13413099	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
106	13427522	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615

107	13442068	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
108	13468093	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
109	13499770	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
110	13508977	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
111	13519435	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
112	13521637	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
113	13338203	BCAT-1	NM_001123785	AT1G10060	-2.28	Down-regulated	NM_001123785 // BCAT-1 // branched-chain-amino-acid aminotransferase 1 // --- // 837542
114	13461438	BZIP61	NM_115674	AT3G58120	-1.73	Down-regulated	NM_115674 // BZIP61 // basic-leucine zipper transcription factor family protein // ---
115	13516311	CER3	NM_125164	AT5G57800	-1.78	Down-regulated	NM_125164 // CER3 // protein ECERIFERUM 3 // --- // 835889 /// AT5G57800.1 // CER3 // p
116	13477383	CYCD3;1	NM_119579	AT4G34160	-1.57	Down-regulated	NM_119579 // CYCD3;1 // cyclin-D3-1 // --- // 829564 /// AT4G34160.1 // CYCD3;1 // cycl
117	13380596	CYP96A15	NM_104570	AT1G57750	-1.52	Down-regulated	NM_104570 // CYP96A15 // alkane hydroxylase CYP96A15 // --- // 842150 /// AT1G57750.1 /
118	13452868	EDA6	NM_113247	AT3G23440	-1.51	Down-regulated	NM_113247 // EDA6 // embryo sac development arrest 6 protein // --- // 821925 /// AT3G2
119	13407392	EXPA15	NM_126361	AT2G03090	-1.80	Down-regulated	NM_126361 // EXPA15 // expansin A15 // --- // 814838 /// AT2G03090.1 // EXPA15 // expan
120	13538937	FTSH11	NM_124696	AT5G53170	-1.53	Down-regulated	NM_124696 // FTSH11 // ATP-dependent zinc metalloprotease FTSH 11 // --- // 835398 ///

121	13355545	HEXO3	NM_105233	AT1G65590	-1.69	Down-regulated	NM_105233 // HEXO3 // beta-hexosaminidase 3 // --- // 842871 /// AT1G65590.1 // HEXO3 /
122	13426364	KCS14	NM_111863	AT3G10280	-2.49	Down-regulated	NM_111863 // KCS14 // 3-ketoacyl-CoA synthase 14 // --- // 820190 /// AT3G10280.1 // KC
123	13545262	nad6	ATMG00270.1	ATMG00270.1	-1.54	Down-regulated	ATMG00270.1 // nad6 // NADH dehydrogenase subunit 6 // --- // 814580
124	13524399	PAS2	NM_001203348	AT5G10480	-1.74	Down-regulated	NM_001203348 // PAS2 // very-long-chain (3R)-3-hydroxyacyl-[acyl-carrier protein] dehyd
125	13369792	PDCB3	NM_101723	AT1G18650	-1.50	Down-regulated	NM_101723 // PDCB3 // plasmodesmata callose-binding protein 3 // --- // 838446 /// AT1G
126	13367598	PGL1	NM_101239	AT1G13700	-1.67	Down-regulated	NM_101239 // PGL1 // 6-phosphogluconolactonase 1 // --- // 837931 /// AT1G13700.1 // PG
127	13374442	PSAF	NM_102871	AT1G31330	-1.59	Down-regulated	NM_102871 // PSAF // photosystem I subunit F // --- // 840021 /// AT1G31330.1 // PSAF /
128	13346202	PSAK	NM_102775	AT1G30380	-1.56	Down-regulated	NM_102775 // PSAK // photosystem I reaction center subunit psaK // --- // 839918 /// AT
129	13458456	PSBO2	NM_114942	AT3G50820	-1.58	Down-regulated	NM_114942 // PSBO2 // oxygen-evolving enhancer protein 1-2 // --- // 824246 /// AT3G508
130	13369782	PSP	NM_202129	AT1G18640	-1.80	Down-regulated	NM_202129 // PSP // phosphoserine phosphatase // --- // 838445 /// AT1G18640.2 // PSP /
131	13444474	PTAC3	NM_111297	AT3G04260	-1.51	Down-regulated	NM_111297 // PTAC3 // plastid transcriptionally active 3 // --- // 819581 /// AT3G04260
132	13489109	RBP31	AT4G24770.1	AT4G24770	-1.67	Down-regulated	AT4G24770.1 // RBP31 // ribonucleoprotein // --- // 828579 /// NM_118610 // RBP31 // ri
133	13445203	RCI2A	NM_111462	AT3G05880	-1.93	Down-regulated	NM_111462 // RCI2A // Hydrophobic protein RCI2A // --- // 819757 /// AT3G05880.1 // RCI
134	13382582	RFC2	NM_104994	AT1G63160	-1.55	Down-regulated	NM_104994 // RFC2 // replication factor C 2 // --- // 842620 /// AT1G63160.1 // RFC2 //

135	13544913	rpl32	ATCG01020.1	ATCG01020	-2.47	Down-regulated	ATCG01020.1 // rpl32 // ribosomal protein L32 // --- // 844704
136	13439976	SBPASE	NM_115438	AT3G55800	-1.57	Down-regulated	NM_115438 // SBPASE // Sedoheptulose-1,7-bisphosphatase // --- // 824746 /// AT3G55800.
137	13530101	SLAH3	NM_122308	AT5G24030	-1.75	Down-regulated	NM_122308 // SLAH3 // SLAC1 homologue 3 // --- // 832468 /// AT5G24030.1 // SLAH3 // SL
138	13525738	SMT1	NM_121374	AT5G13710	-1.64	Down-regulated	NM_121374 // SMT1 // cycloartenol-c-24-methyltransferase // --- // 831216 /// AT5G13710
139	13469677	TPS03	NM_117775	AT4G16740	-1.58	Down-regulated	NM_117775 // TPS03 // tricyclene synthase // --- // 827377 /// AT4G16740.1 // TPS03 //
140	13474801	TYRDC	NM_001203932	AT4G28680	-1.65	Down-regulated	NM_001203932 // TYRDC // tyrosine decarboxylase // --- // 828986 /// AT4G28680.3 // TYR
141	13519853	XTH6	NM_125970	AT5G65730	-1.51	Down-regulated	NM_125970 // XTH6 // probable xyloglucan endotransglucosylase/hydrolase protein 6 // --
142	13352563	ZKT	NM_104423	AT1G55480	-1.61	Down-regulated	NM_104423 // ZKT // ZKT protein containing PDZ, K-box and a TPR region // --- // 841995

5.4.3.1 Genes involved in Signal Transduction

In contrast to non-induced (aphid feeding absent) plants, 7 genes were found significantly differentially expressed in the signal transduction category when resistant CWR were compared to resistant LR after aphid feeding. Out of these 7 genes 5 were up-regulated. These genes included **AT1G70700** (TIFY7), a protein presumed to be involved in jasmonate signalling (https://www.arabidopsis.org/servlets/Tairobject?id=30459&type=Locus); **AT3G12500** (HCHIB) which encodes a basic chitinase involved in ethylene/jasmonic acid mediated signalling pathway during systemic acquired resistance based on expression analyses; **AT4G36900** (RAP2.10) ethylene-responsive transcription factor which encodes a member of the DREB subfamily A-5 of ERF/AP2 transcription factor family (RAP2.10) - it is involved in ethylene-activated signalling pathway and regulation of transcription; **AT5G20730** (NPH4), encodes an auxin-regulated transcriptional activator; **AT5G64810** (WRKY51) is a putative transcription factor. The two genes found down regulated were **AT4G20940** and **AT4G26380**. Both these genes encode for putative leucine-rich receptor-like protein kinase. Transcription factors (TFs) and signal transduction genes have been reported to be significant for crop domestication (Doebley et al. 2006). Moreover these genes are considered to play a central role in regulating gene expression and control innumerable biological processes by cabbage aphid feeding. According to one theory of domestication, trait change is usually a result of loss of function in order to adapt to new environments (Lester 1989).

5.4.3.2 Genes in response to stress

Changes in transcription of genes associated with stress increased from 5 in non-stressed plants to 25 after aphid feeding. The majority of genes responding to stress were up-regulated in response to aphid feeding in CWR when compared to LR. In total 25 genes were found in this category, 15 were found to be induced (up-regulated by aphid feeding) while 10 were down- regulated in CWR.

The up-regulated genes in response to aphid feeding were **AT1G08930** (ERD6), **AT1G70700** (TIFY7), **AT1G80840** (WRKY40), **AT2G22370**, **AT3G12500** (HCHIB), **AT3G13330** (PA200), **AT3G14210** (ESM1), **AT3G14990, AT3G48030, AT4G01700, AT4G37870** (PCK1), **AT5G24530** (DMR6), **AT5G46180** (DELTA-OAT), AT5G48850 (ATSDI1), **AT5G64810** (WRKY51). The down regulated genes were **AT1G55480** (ZKT), **AT1G73680** (ALPHA), **AT3G05880** (RCI2A), **AT3G5580**0 (SBPASE), **AT4G16740** (TPS03)**, AT4G24770**, (RBP31)**, AT4G28680** (TYRDC), **AT5G06530**, **AT5G53170** (FTSH11), **AT5G65730** (XTH6). The details of all genes are provided in supplementary table S5.6 and will further be discussed later in the chapter.

5.4.3.3 Genes in response to stimulus

The majority of genes that were found in this category overlapped with 'response to stress' category as expected because GO term 'Response to stress' is a subdivision of this category as generally stress occurs as a result of changes due to some stimulus in plants. Twenty seven were found to be significantly enriched under response to stimulus category. Out of these, 16 were found to be induced in response to 24 hours of aphid feeding while 11 were found to be repressed when CWR were compared to LR.

Twenty one genes were common to response to stress category while 6 genes were unique to response to stimulus. Genes falling only under this category after aphid feeding were **AT1G72290, AT5G64250, AT4G15210, AT5G62380, AT5G20730, AT3G50820.** The higher number of genes in response to stress and response to stimulus category indicates the variation in transcription response to CWR and LR towards aphid feeding and these genes may be good candidates for aphid resistance in Brassica plants.

5.4.4 Comparison of gene expression of CWR and LR in response to presence or absence of aphid feeding

The comparison of induced and non-induced gene expression showed that a large number of genes are induced by aphid feeding. A total of 75 genes had significant difference in expression when resistant CWR were compared with resistant LR in the absence of aphid feeding (Table 5.1). In contrast, a total of 261 genes were significantly differentially expressed in response to aphid feeding (Table 5.2).

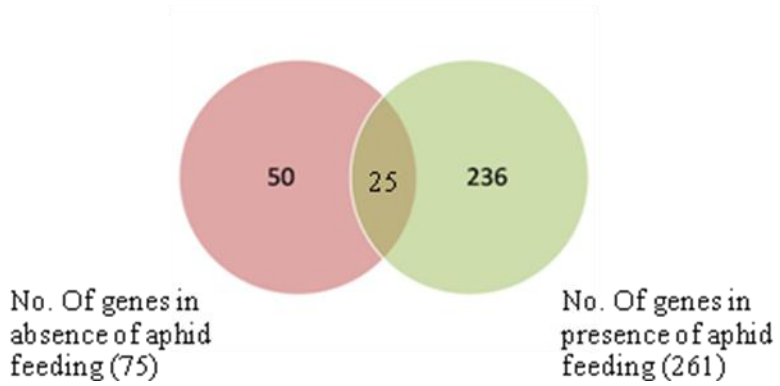


Figure 5.1: Venn diagram: Venn diagram shows the number of significantly differentially expressed among CWR and LR in presence and absence of aphid feeding.

Twenty five common genes were significantly differentially expressed between non induced and induced plants.

Out of these 25 common genes, 15 genes were up regulated in CWR while 10 were down regulated. The details of these genes are provided in table 5.3. The common genes are involved in functions like Cellular and metabolic processes, transport, protein metabolism, response to stress, and response to stimulus. Similarly a list of uniquely expressed genes in presence of aphid feeding (table 5.4) and in absence of aphid feeding (table 5.5) is provided. The uniquely up regulated genes can be important indicators of response to aphid feeding and need further investigation in future. The number of candidate genes for aphid resistance has been revealed in response to aphid feeding. These candidate genes should be further explored in future to confirm their role in resistance in Brassica plants due to aphid feeding.

Table 5.3: List of 25 genes common between induced (presence of aphid feeding) and non -induced (absence of aphid feeding) treatment: List of common genes between two comparisons with significant differences in expression at FDR corrected p-value <0.05 and fold change ≥1.5 to ≤-1.5. A positive FC indicates up regulation ;negative FC indicates down regulation

S.no	Column ID	RefSeq	Gene Symbol	Fold-Change	Fold-Change (Description)	gene_assignment
1	13524842	NM_121190	ASP3	1.64	Up-regulated	NM_121190 // ASP3 // aspartate aminotransferase 3 // --- // 831024 /// AT5G11520.1 // A
2	13343847	NM_202168	AT1G23465	3.21	Up-regulated	NM_202168 // AT1G23465 // peptidase-S24/S26 domain-containing protein // --- // 2745760
3	13385899	AT1G70949.1	AT1G70949	-2.07	Up-regulated	AT1G70949.1 // AT1G70949 // hypothetical protein // --- // 6240811 /// NM_001124109 //
4	13408365	NM_126758	AT2G07721	11.11	Up-regulated	NM_126758 // AT2G07721 // hypothetical protein // --- // 815394 /// AT2G07721.1 // AT2G
5	13374992	NM_001036054	CID11	1.65	Down-regulated	NM_001036054 // CID11 // CTC-interacting domain 11 protein // --- // 840173 /// AT1G327
6	13380596	NM_104570	CYP96A15	-1.52	Up-regulated	NM_104570 // CYP96A15 // alkane hydroxylase CYP96A15 // --- // 842150 /// AT1G57750.1 /
7	13536092	NM_123987	DELTA-OAT	1.77	Down-regulated	NM_123987 // DELTA-OAT // ornithine-delta-aminotransferase // --- // 834660 /// AT5G461
8	13337731	NM_001035929	ERD6	1.63	Up-regulated	NM_001035929 // ERD6 // sugar transporter ERD6 // --- // 837414 /// NM_100765 // ERD6 /
9	13407392	NM_126361	EXPA15	-1.80	Down-regulated	NM_126361 // EXPA15 // expansin A15 // --- // 814838 /// AT2G03090.1 // EXPA15 // expan
10	13545387	ATMG01220.1	nad1	43.92	Down-regulated	ATMG01220.1 // nad1 // NADH dehydrogenase subunit 1 // --- // 3890477

11	13368314	AT1G15210.1	PDR7	1.59	Up-regulated	AT1G15210.1 // PDR7 // ABC transporter G family member 35 // --- // 838087 /// NM_10138
12	13369782	NM_202129	PSP	-1.80	Down-regulated	NM_202129 // PSP // phosphoserine phosphatase // --- // 838445 /// AT1G18640.2 // PSP /
13	13530101	NM_122308	SLAH3	-1.75	Down-regulated	NM_122308 // SLAH3 // SLAC1 homologue 3 // --- // 832468 /// AT5G24030.1 // SLAH3 // SL
14	13348616			7.99	Up-regulated	---
15	13361670			4.00	Down-regulated	---
16	13367956			-2.51	Down-regulated	---
17	13406143			2.33	Up-regulated	---
18	13406711			-1.76	Down-regulated	---
19	13442074			-2.18	Up-regulated	---
20	13451950			-3.84	Up-regulated	---
21	13466328			2.41	Up-regulated	---
22	13475068			-1.79	Down-regulated	---
23	13482213			5.81	Up-regulated	---
24	13511965			2.26	Up-regulated	---
25	13531979			2.53	Up-regulated	---

Table 5.4: List of 236 uniquely expressed induced (presence of aphid feeding) treatment: List of unique genes in presence of aphid feeding with significant differences in expression at FDR corrected p-value <0.05 and fold change ≥1.5 to ≤-1.5. A positive FC indicates up regulation ;negative FC indicates down regulation

S.no	Transcript ID	Gene Symbol	RefSeq	Locus ID	Fold-Change	Fold-Change(Description)	Gene_assignment
1	13355545	HEXO3	NM_105233	AT1G65590	-1.69	Down-regulated	NM_105233 // HEXO3 // beta-hexosaminidase 3 // --- // 842871 /// AT1G65590.1 // HEXO3 /
2	13457204	ABCA2	NM_114641	AT3G47730	1.52	Up-regulated	NM_114641 // ABCA2 // ABC transporter A family member 2 // --- // 823927 /// AT3G47730.
3	13439976	SBPASE	NM_115438	AT3G55800	-1.57	Down-regulated	NM_115438 // SBPASE // Sedoheptulose-1,7-bisphosphatase // --- // 824746 /// AT3G55800.
4	13477383	CYCD3;1	NM_119579	AT4G34160	-1.57	Down-regulated	NM_119579 // CYCD3;1 // cyclin-D3-1 // --- // 829564 /// AT4G34160.1 // CYCD3;1 // cycl
5	13528308	ACL5	NM_001085136	AT5G19530	1.84	Up-regulated	NM_001085136 // ACL5 // Thermospermine synthase ACAULIS5 // --- // 832073 /// AT5G19530
6	13363863	CALS1	NM_100436	AT1G05570	1.58	Up-regulated	NM_100436 // CALS1 // callose synthase 1 // --- // 837059 /// AT1G05570.1 // CALS1 // c
7	13338203	BCAT-1	NM_001123785	AT1G10060	-2.28	Down-regulated	NM_001123785 // BCAT-1 // branched-chain-amino-acid aminotransferase 1 // --- // 837542
8	13367035	AT1G12423	NM_001123802	AT1G12423	-1.80	Down-regulated	NM_001123802 // AT1G12423 // hypothetical protein // --- // 6240731 /// AT1G12423.1 //
9	13367598	PGL1	NM_101239	AT1G13700	-1.67	Down-regulated	NM_101239 // PGL1 // 6-phosphogluconolactonase 1 // --- //

							837931 /// AT1G13700.1 // PG
10	13369792	PDCB3	NM_101723	AT1G18650	-1.50	Down-regulated	NM_101723 // PDCB3 // plasmodesmata callose-binding protein 3 // --- // 838446 /// AT1G
11	13342911	AT1G21245	NM_148476	AT1G21245	1.67	Up-regulated	NM_148476 // AT1G21245 // Protein kinase superfamily protein // --- // 838720 /// AT1G2
12	13344465	MYB116	NM_102344	AT1G25340	1.67	Up-regulated	NM_102344 // MYB116 // putative transcription factor (MYB116) // --- // 839118 /// AT1G
13	13346202	PSAK	NM_102775	AT1G30380	-1.56	Down-regulated	NM_102775 // PSAK // photosystem I reaction center subunit psaK // --- // 839918 /// AT
14	13374442	PSAF	NM_102871	AT1G31330	-1.59	Down-regulated	NM_102871 // PSAF // photosystem I subunit F // --- // 840021 /// AT1G31330.1 // PSAF /
15	13374726	AT1G32080	NM_102942	AT1G32080	-1.53	Down-regulated	NM_102942 // AT1G32080 // plastidal glycolate/glycerate translocator 1 // --- // 840100
16	13378599	AT1G52220	AT1G52220.1	AT1G52220	-1.91	Down-regulated	AT1G52220.1 // AT1G52220 // hypothetical protein // --- // 841652 /// AT1G52220.2 // AT
17	13352563	ZKT	NM_104423	AT1G55480	-1.61	Down-regulated	NM_104423 // ZKT // ZKT protein containing PDZ, K-box and a TPR region // --- // 841995
18	13382582	RFC2	NM_104994	AT1G63160	-1.55	Down-regulated	NM_104994 // RFC2 // replication factor C 2 // --- // 842620 /// AT1G63160.1 // RFC2 //
19	13354650	AT1G63220	NM_105001	AT1G63220	1.51	Up-regulated	NM_105001 // AT1G63220 // calcium-dependent lipid-binding domain-containing protein //

20	13356629	AT1G68570	NM_105528	AT1G68570	1.67	Up-regulated	NM_105528 // AT1G68570 // putative nitrite transporter // --- // 843186 /// AT1G68570.1
21	13384976	AT1G68570	NM_105528	AT1G68570	1.57	Up-regulated	NM_105528 // AT1G68570 // putative nitrite transporter // --- // 843186 /// AT1G68570.1
22	13357574	TIFY7	NM_105738	AT1G70700	1.54	Up-regulated	NM_105738 // TIFY7 // protein TIFY 7 // --- // 843407 /// AT1G70700.1 // TIFY7 // prote
23	13358274	AT1G72290	NM_105888	AT1G72290	1.83	Up-regulated	NM_105888 // AT1G72290 // trypsin inhibitor (Kunitz) domain-containing protein // --- /
24	13386385	AT1G72430	NM_105902	AT1G72430	-1.68	Down-regulated	NM_105902 // AT1G72430 // SAUR-like auxin-responsive protein family // --- // 843575 //
25	13386441	AT1G72540	NM_105913	AT1G72540	1.94	Up-regulated	NM_105913 // AT1G72540 // putative receptor-like protein kinase // --- // 843586 /// AT
26	13386963	ALPHA	NM_106027	AT1G73680	-1.56	Down-regulated	NM_106027 // ALPHA DOX2 // alpha dioxygenase // --- // 843703 /// AT1G73680.1 // ALPHA
27	13361730	WRKY40	NM_106732	AT1G80840	1.77	Up-regulated	NM_106732 // WRKY40 // putative WRKY transcription factor 40 // --- // 844423 /// AT1G8
28	13407824	AT2G04650	NM_126494	AT2G04650	1.67	Up-regulated	NM_126494 // AT2G04650 // ADP-glucose pyrophosphorylase-like protein // --- // 815007 /
29	13392342	AT2G07787	NM_201712	AT2G07787	1.53	Up-regulated	NM_201712 // AT2G07787 // hypothetical protein // --- // 2745469 /// AT2G07787.1 // AT2
30	13411645	AT2G22370	NM_127802	AT2G22370	1.61	Up-regulated	NM_127802 // AT2G22370 // hypothetical protein // --- // 816769 /// AT2G22370.1 // AT2G

31	13415609	AT2G31830	NM_128741	AT2G31830	-1.55	Down-regulated	NM_128741 // AT2G31830 // Type II inositol-1,4,5-trisphosphate 5-phosphatase 14 // ---
32	13401971	AT2G37440	NM_129299	AT2G37440	1.71	Up-regulated	NM_129299 // AT2G37440 // DNAse I-like superfamily protein // --- // 818321 /// AT2G374
33	13402202	IRX12	NM_129364	AT2G38080	1.64	Up-regulated	NM_129364 // IRX12 // laccase-4 // --- // 818386 /// AT2G38080.1 // IRX12 // laccase-4
34	13418497	FAC1	NM_129384	AT2G38280	1.59	Up-regulated	NM_129384 // FAC1 // AMP deaminase // --- // 818408 /// NM_179963 // FAC1 // AMP deamin
35	13420535	CSY3	NM_129840	AT2G42790	1.71	Up-regulated	NM_129840 // CSY3 // citrate synthase 3 // --- // 818879 /// AT2G42790.1 // CSY3 // cit
36	13404388	AT2G43150	NM_129877	AT2G43150	1.53	Up-regulated	NM_129877 // AT2G43150 // Proline-rich extensin-like family protein // --- // 818917 //
37	13444474	PTAC3	NM_111297	AT3G04260	-1.51	Down-regulated	NM_111297 // PTAC3 // plastid transcriptionally active 3 // --- // 819581 /// AT3G04260
38	13387889	AT3G05520	NM_001161121	AT3G05520	-5.83	Down-regulated	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///
39	13437797	AT3G05520	NM_001161121	AT3G05520	-5.76	Down-regulated	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///
40	13445024	AT3G05520	NM_001161121	AT3G05520	-5.77	Down-regulated	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///
41	13458322	AT3G05520	NM_001161121	AT3G05520	-5.83	Down-regulated	NM_001161121 // AT3G05520 // F-actin-capping protein subunit alpha // --- // 819717 ///

42	13445203	RCI2A	NM_111462	AT3G05880	-1.93	Down-regulated	NM_111462 // RCI2A // Hydrophobic protein RCI2A // --- // 819757 /// AT3G05880.1 // RCI
43	13425033	AT3G06470	NM_111522	AT3G06470	-1.73	Down-regulated	NM_111522 // AT3G06470 // GNS1/SUR4 membrane protein family // --- // 819824 /// AT3G06
44	13426364	KCS14	NM_111863	AT3G10280	-2.49	Down-regulated	NM_111863 // KCS14 // 3-ketoacyl-CoA synthase 14 // --- // 820190 /// AT3G10280.1 // KC
45	13448399	HCHIB	NM_112085	AT3G12500	1.77	Up-regulated	NM_112085 // HCHIB // basic chitinase B // --- // 820429 /// AT3G12500.1 // HCHIB // ba
46	13448873	PA200	NM_112178	AT3G13330	1.56	Up-regulated	NM_112178 // PA200 // proteasome activating protein 200 // --- // 820533 /// AT3G13330.
47	13545581	ESM1	NM_112278	AT3G14210	4.03	Up-regulated	NM_112278 // ESM1 // epithiospecifier modifier 1 // --- // 820639 /// AT3G14210.1 // ES
48	13428400	AT3G14990	NM_001035621	AT3G14990	1.55	Up-regulated	NM_001035621 // AT3G14990 // protein DJ-1-like A // --- // 820728 /// NM_112361 // AT3G
49	13429722	PYD1	NM_112662	AT3G17810	1.56	Up-regulated	NM_112662 // PYD1 // putative dihydropyrimidine dehydrogenase // --- // 821049 /// AT3G
50	13544974	AT3G20362	AT3G20362.1	AT3G20362	-1.62	Down-regulated	AT3G20362.1 // AT3G20362 // hypothetical protein // --- // 821581 /// NM_148737 // AT3G
51	13452781	AT3G23090	NM_113210	AT3G23090	1.54	Up-regulated	NM_113210 // AT3G23090 // targeting protein for Xklp2-like protein // --- // 821884 ///
52	13452868	EDA6	NM_113247	AT3G23440	-1.51	Down-regulated	NM_113247 // EDA6 // embryo sac development arrest 6 protein // --- // 821925 /// AT3G2

53	13432310	AT3G23510	NM_113254	AT3G23510	1.54	Up-regulated	NM_113254 // AT3G23510 // cyclopropane-fatty-acyl-phospholipid synthase // --- // 82193
54	13453160	AT3G24180	NM_001035680	AT3G24180	1.64	Up-regulated	NM_001035680 // AT3G24180 // Beta-glucosidase, GBA2 type family protein // --- // 82200
55	13453771	AT3G25880	NM_113491	AT3G25880	-1.99	Down-regulated	NM_113491 // AT3G25880 // hypothetical protein // --- // 822183 /// AT3G25880.1 // AT3G
56	13454169	MUB4	NM_113612	AT3G26980	1.64	Up-regulated	NM_113612 // MUB4 // membrane-anchored ubiquitin-fold protein 4 // --- // 822315 /// AT
57	13435143	AT3G43570	NM_114225	AT3G43570	-1.90	Down-regulated	NM_114225 // AT3G43570 // GDSL esterase/lipase // --- // 823453 /// AT3G43570.1 // AT3G
58	13456324	IVD	NM_114399	AT3G45300	1.66	Up-regulated	NM_114399 // IVD // isovaleryl-CoA-dehydrogenase // --- // 823668 /// AT3G45300.1 // IV
59	13457337	AT3G48030	NM_114672	AT3G48030	1.56	Up-regulated	NM_114672 // AT3G48030 // RING-H2 finger protein ATL48 // --- // 823958 /// AT3G48030.1
60	13437332	AT3G49130	NM_114772	AT3G49130	1.52	Up-regulated	NM_114772 // AT3G49130 // SWAP (Suppressor-of-White-APricot)/surp RNA-binding domain-co
61	13458456	PSBO2	NM_114942	AT3G50820	-1.58	Down-regulated	NM_114942 // PSBO2 // oxygen-evolving enhancer protein 1-2 // --- // 824246 /// AT3G508
62	13459262	MDAR1	NM_001161202	AT3G52880	1.53	Up-regulated	NM_001161202 // MDAR1 // monodehydroascorbate reductase (NADH) // --- // 824454 /// AT3
63	13460341	AT3G55240	NM_115382	AT3G55240	-3.60	Down-regulated	NM_115382 // AT3G55240 // hypothetical protein // --- // 824690 /// AT3G55240.1 // AT3G

64	13460997	AT3G57020	NM_115561	AT3G57020	1.73	Up-regulated	NM_115561 // AT3G57020 // strictosidine synthase family protein // --- // 824869 /// AT
65	13461438	BZIP61	NM_115674	AT3G58120	-1.73	Down-regulated	NM_115674 // BZIP61 // basic-leucine zipper transcription factor family protein // ---
66	13480634	AT4G01700	NM_116400	AT4G01700	1.73	Up-regulated	NM_116400 // AT4G01700 // Chitinase family protein // --- // 828131 /// AT4G01700.1 //
67	13465352	AT4G03420	NM_116580	AT4G03420	1.52	Up-regulated	NM_116580 // AT4G03420 // hypothetical protein // --- // 827928 /// AT4G03420.1 // AT4G
68	13465878	AT4G05010	NM_116740	AT4G05010	1.57	Up-regulated	NM_116740 // AT4G05010 // F-box protein // --- // 825843 /// AT4G05010.1 // AT4G05010 /
69	13484924	BAM5	NM_117609	AT4G15210	1.78	Up-regulated	NM_117609 // BAM5 // beta-amylase 5 // --- // 827185 /// AT4G15210.1 // BAM5 // beta-am
70	13485148	AT4G15885	NM_148338	AT4G15885	1.56	Up-regulated	NM_148338 // AT4G15885 // kinesin motor protein-like protein // --- // 827270 /// AT4G1
71	13469677	TPS03	NM_117775	AT4G16740	-1.58	Down-regulated	NM_117775 // TPS03 // tricyclene synthase // --- // 827377 /// AT4G16740.1 // TPS03 //
72	13485638	AT4G16765	NM_179213	AT4G16765	1.55	Up-regulated	NM_179213 // AT4G16765 // oxidoreductase, 2OG-Fe(II) oxygenase family protein // --- //
73	13471375	AT4G20940	NM_118212	AT4G20940	-1.54	Down-regulated	NM_118212 // AT4G20940 // leucine-rich receptor-like protein kinase // --- // 827842 //
74	13487592	AT4G21520	NM_118272	AT4G21520	-1.60	Down-regulated	NM_118272 // AT4G21520 // WD40 repeat family protein // --- // 828237 /// AT4G21520.1 /

75	13489109	RBP31	AT4G24770.1	AT4G24770	-1.67	Down-regulated	AT4G24770.1 // RBP31 // ribonucleoprotein // --- // 828579 /// NM_118610 // RBP31 // ri
76	13346598	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
77	13356880	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
78	13395867	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
79	13411603	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
80	13453704	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
81	13489860	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
82	13489862	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
83	13501968	AT4G26380	NM_118771	AT4G26380	-2.28	Down-regulated	NM_118771 // AT4G26380 // cysteine/histidine-rich C1 domain-containing protein // --- /
84	13474801	TYRDC	NM_001203932	AT4G28680	-1.65	Down-regulated	NM_001203932 // TYRDC // tyrosine decarboxylase // --- // 828986 /// AT4G28680.3 // TYR
85	13476724	AT4G32710	NM_119424	AT4G32710	-1.63	Down-regulated	NM_119424 // AT4G32710 // proline-rich receptor-like protein kinase PERK14 // --- // 37

86	13476949	AT4G33080	NM_119462	AT4G33080	-1.52	Down-regulated	NM_119462 // AT4G33080 // AGC (cAMP-dependent, cGMP-dependent and protein kinase C) kin
87	13477038	AT4G33440	NM_119498	AT4G33440	1.66	Up-regulated	NM_119498 // AT4G33440 // pectin lyase-like superfamily protein // --- // 829481 /// AT
88	13478537	RAP2.10	NM_119854	AT4G36900	1.53	Up-regulated	NM_119854 // RAP2.10 // ethylene-responsive transcription factor RAP2-10 // --- // 8298
89	13494916	PCK1	NM_119948	AT4G37870	1.64	Up-regulated	NM_119948 // PCK1 // phosphoenolpyruvate carboxykinase [ATP] // --- // 829943 /// AT4G3
90	13522706	AT5G06530	NM_001036766	AT5G06530	-1.67	Down-regulated	NM_001036766 // AT5G06530 // ABC transporter G family member 22 // --- // 830541 /// AT
91	13524399	PAS2	NM_001203348	AT5G10480	-1.74	Down-regulated	NM_001203348 // PAS2 // very-long-chain (3R)-3-hydroxyacyl-[acyl-carrier protein] dehyd
92	13525208	AT5G12323	NM_001125743	AT5G12323	-1.50	Down-regulated	NM_001125743 // AT5G12323 // hypothetical protein // --- // 6240742 /// AT5G12323.1 //
93	13525738	SMT1	NM_121374	AT5G13710	-1.64	Down-regulated	NM_121374 // SMT1 // cycloartenol-c-24-methyltransferase // --- // 831216 /// AT5G13710
94	13526694	AT5G15780	NM_121583	AT5G15780	-1.64	Down-regulated	NM_121583 // AT5G15780 // pollen Ole e 1 allergen and extensin family protein // --- //
95	13528833	NPH4	NM_180715	AT5G20730	1.62	Up-regulated	NM_180715 // NPH4 // auxin-regulated transcriptional activator NPH4 // --- // 832196 //
96	13529672	AT5G22860	NM_180728	AT5G22860	1.58	Up-regulated	NM_180728 // AT5G22860 // Serine carboxypeptidase S28 family protein // --- // 832349 /

97	13505359	DMR6	NM_122361	AT5G24530	2.36	Up-regulated	NM_122361 // DMR6 // downy mildew resistance 6 protein / oxidoreductase // --- // 83252
98	13532237	AT5G35480	NM_122939	AT5G35480	-2.65	Down-regulated	NM_122939 // AT5G35480 // hypothetical protein // --- // 833512 /// AT5G35480.1 // AT5G
99	13510281	AT5G44020	NM_123769	AT5G44020	-1.95	Down-regulated	NM_123769 // AT5G44020 // HAD superfamily, subfamily IIIB acid phosphatase // --- // 83
100	13337838	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
101	13370341	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
102	13494665	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
103	13531589	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
104	13544460	AT5G45720	NM_001203547	AT5G45720	-5.81	Down-regulated	NM_001203547 // AT5G45720 // AAA-type ATPase family protein // --- // 834611 /// AT5G45
105	13537205	ATSDI1	NM_124262	AT5G48850	2.02	Up-regulated	NM_124262 // ATSDI1 // protein SULPHUR DEFICIENCY-INDUCED 1 // --- // 834943 /// AT5G48
106	13538937	FTSH11	NM_124696	AT5G53170	-1.53	Down-regulated	NM_124696 // FTSH11 // ATP-dependent zinc metalloprotease FTSH 11 // --- // 835398 ///
107	13516311	CER3	NM_125164	AT5G57800	-1.78	Down-regulated	NM_125164 // CER3 // protein ECERIFERUM 3 // --- // 835889 /// AT5G57800.1 // CER3 // p

108	13387345	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
109	13411454	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
110	13413099	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
111	13427522	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
112	13442068	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
113	13468093	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
114	13499770	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
115	13508977	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
116	13517400	AT5G60280	AK221799	AT5G60280	-4.38	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
117	13519435	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615
118	13521637	AT5G60280	AK221799	AT5G60280	-3.05	Down-regulated	AK221799 // AT5G60280 // concanavalin A-like lectin kinase-like protein // --- // 83615

119	13542294	AT5G61820	AT5G61820.1	AT5G61820	1.54	Up-regulated	AT5G61820.1 // AT5G61820 // hypothetical protein // --- // 836304 /// NM_125576 // AT5G
120	13518334	NAC101	NM_125632	AT5G62380	1.56	Up-regulated	NM_125632 // NAC101 // NAC-domain transcription factor // --- // 836359 /// AT5G62380.1
121	13543310	AT5G64250	NM_125821	AT5G64250	1.62	Up-regulated	NM_125821 // AT5G64250 // Aldolase-type TIM barrel family protein // --- // 836546 ///
122	13543347	FKBP12	NM_125831	AT5G64350	1.53	Up-regulated	NM_125831 // FKBP12 // peptidyl-prolyl isomerase FKBP12 // --- // 836556 /// AT5G64350.
123	13519410	WRKY51	NM_125877	AT5G64810	1.66	Up-regulated	NM_125877 // WRKY51 // putative WRKY transcription factor 51 // --- // 836602 /// AT5G6
124	13519853	XTH6	NM_125970	AT5G65730	-1.51	Down-regulated	NM_125970 // XTH6 // probable xyloglucan endotransglucosylase/hydrolase protein 6 // --
125	13544913	rpl32	ATCG01020.1	ATCG01020	-2.47	Down-regulated	ATCG01020.1 // rpl32 // ribosomal protein L32 // --- // 844704
126	13545262	nad6	ATMG00270.1	ATMG00270.1	-1.54	Down-regulated	ATMG00270.1 // nad6 // NADH dehydrogenase subunit 6 // --- // 814580
127	13545283	ArthMp039	ATMG00470.1	ATMG00470	1.62	Up-regulated	ATMG00470.1 // ArthMp039 // hypothetical protein // --- // 3371332
128	13545359	ArthMp091	ATMG01040.1	ATMG01040	1.84	Up-regulated	ATMG01040.1 // ArthMp091 // hypothetical protein // --- // 4024969
129	13545411	ArthMp112	ATMG01410.1	ATMG01410	1.51	Up-regulated	ATMG01410.1 // ArthMp112 // hypothetical protein // --- // 4024966

Table 5.5: List of 50 uniquely expressed gene in non -induced (absence of aphid feeding) treatment: List of unique genes with significant differences in expression at FDR corrected p-value <0.05 and fold change ≥1.5 to ≤-1.5. A positive FC indicates up regulation ;negative FC indicates down regulation

S.No	Transcript ID	Gene Symbol	RefSeq	Locus ID	Fold-Change	Fold-Change (Description)	gene_assignment
1	13383271	AT1G64850	NM_105159	AT1G64850	1.64165	Up-regulated	NM_105159 // AT1G64850 // calcium-binding EF-hand-containing protein // --- // 842793 /
2	13545217	AT2G07696	NM_126746	AT2G07696	1.62455	Up-regulated	NM_126746 // AT2G07696 // ribosomal protein S7 // --- // 815372 /// AT2G07696.1 // rps7
3	13437647	AT3G50123	NM_001125335	AT3G50123	1.54883	Up-regulated	NM_001125335 // AT3G50123 // hypothetical protein // --- // 6240873 /// AT3G50123.1 //
4	13442656	AT3G62510	NM_116116	AT3G62510	1.59736	Up-regulated	NM_116116 // AT3G62510 // protein disulfide isomerase-like protein // --- // 825425 ///
5	13530775	AT5G25190	NM_122428	AT5G25190	1.57433	Up-regulated	NM_122428 // AT5G25190 // ethylene-responsive transcription factor ERF003 // --- // 832
6	13423873	MS2	NM_180176	AT3G03780	1.61016	Up-regulated	NM_180176 // MS2 // methionine synthase 2 // --- // 821147 /// NM_001125092 // MS2 // m
7	13500896	NAC083	NM_121321	AT5G13180	2.04555	Up-regulated	NM_121321 // NAC083 // NAC domain containing protein 83 // --- // 831157 /// AT5G13180.
8	13536042	PTR3	NM_123973	AT5G46050	1.95149	Up-regulated	NM_123973 // PTR3 // peptide transporter 3 // --- // 834646 /// AT5G46050.1 // PTR3 //
9	13374451	RUB1	NM_102873	AT1G31340	1.52173	Up-regulated	NM_102873 // RUB1 // ubiquitin-NEDD8-like protein RUB1 // --- // 840023 /// AT1G31340.1
10	13345523	AT1G28307	NM_001123899	AT1G28307	-1.55934	Down-regulated	NM_001123899 // AT1G28307 // hypothetical protein // --- // 6241164 /// AT1G28307.1 //
11	13347647	AT1G34041	NM_001123935	AT1G34041	-1.50672	Down-regulated	NM_001123935 // AT1G34041 // hypothetical protein // --- // 6241199 /// AT1G34041.1 //

12	13385899	AT1G70949	AT1G70949.1	AT1G70949	-1.92252	Down-regulated	AT1G70949.1 // AT1G70949 // hypothetical protein // --- // 6240811 /// NM_001124109 //
13	13392218	AT2G07674	NM_126731	AT2G07674	-1.9804	Down-regulated	NM_126731 // AT2G07674 // hypothetical protein // --- // 815346 /// AT2G07674.1 // AT2G
14	13421794	AT2G45860	NM_130149	AT2G45860	-1.78321	Down-regulated	NM_130149 // AT2G45860 // hypothetical protein // --- // 819194 /// AT2G45860.1 // AT2G
15	13431841	AT3G22415	NM_202624	AT3G22415	-1.67373	Down-regulated	NM_202624 // AT3G22415 // hypothetical protein // --- // 2745889 /// AT3G22415.1 // AT3
16	13353669	E1	NM_104682	AT1G59900	-1.54269	Down-regulated	NM_104682 // E1 ALPHA // pyruvate dehydrogenase complex E1 alpha subunit // --- // 8422
17	13355059	GH9C2	NM_105114	AT1G64390	-1.93732	Down-regulated	NM_105114 // GH9C2 // glycosyl hydrolase 9C2 // --- // 842747 /// AT1G64390.1 // GH9C2
18	13339000	GLX1	NM_001198039	AT1G11840	-1.73548	Down-regulated	NM_001198039 // GLX1 // glyoxalase I homolog // --- // 837731 /// AT1G11840.6 // GLX1 /
19	13401542	SBE2.1	NM_129196	AT2G36390	-1.52435	Down-regulated	NM_129196 // SBE2.1 // 1,4-alpha-glucan branching enzyme 2-1 // --- // 818212 /// AT2G3
20	13356178	URE	NM_105422	AT1G67550	-1.5013	Down-regulated	NM_105422 // URE // urease // --- // 843076 /// AT1G67550.1 // URE // urease // --- //

5.5 Discussion

5.5.1 Transcriptional responses of Brassica wild relatives and Landraces in absence and presence of aphid feeding.

Plants show phenotypic changes when challenged with herbivorous insects. The mechanisms underlying these changes include the activation of transcriptional responses. Comparing transcriptional responses of wild and cultivated members of the same plant family may contribute to the understanding of the evolution of plant-herbivore interactions. The results from this study show the variation in gene expression in resistant CWR and LR in response to presence or absence of aphid feeding. It was observed that aphid feeding plays a role in the change of gene expression when CWR were compared to LR which is evident by the fact that more genes were significantly differentially expressed as a response to aphid feeding in comparison to its absence. The significantly differentially expressed genes in categories like response to stress, response to stimulus, signalling, developmental processes can be good indicators of aphid resistance or tolerance in the CWR as compared to LR and may prove useful in the development of aphid resistant Brassica varieties in future. It was noted that in CWR, 24 h of *B. brassicae* feeding resulted in the induction of more defence-related genes in comparison to absence of aphid feeding, indicating the activation of certain defence mechanisms. Several photosynthesis- and/or development related genes were repressed upon aphid feeding (Table 5.2 and TableS5.6). Since defence activation has been shown to be costly (Wu & Baldwin 2010; Herms & Mattson 1992), it is likely that plants have reallocated resources for defence at the expense of growth and/or photosynthesis (Bilgin et al. 2010).

Some of the important transcripts are discussed which are thought to be playing a role in response to aphid feeding. AT5G46050 (PTR) was found to be 1.57 fold up-regulated in CWR suggesting that there is a higher level of protection to wounding present in wild plant species as compared to LR. In addition to this, another study showed that expression level of AtPTR3 was regulated by signalling compounds like salicylic acid (SA), methyl jasmonate (MeJA) and abscisic acid and PTR3 mutants have increased susceptibility to virulent bacterial pathogens (Karim et al. 2005). AT1G08930 (ERD6) encodes a putative sucrose transporter whose gene expression is induced by dehydration and cold. AT1G70700 (TIFY7or JAZ9) protein is presumed to be involved in jasmonate signalling. TIFY7 transcript levels rise in response to a jasmonate stimulus which is a defence response (http://www.ncbi.nlm.nih.gov/gene?cmd=DetailsSearch&term=AT1G70700). The plant hormone jasmonate plays a crucial role in regulating plant responses to herbivorous insects and microbial pathogens and is an important regulator of plant growth and development (F. Zhang et al. 2015). AT1G80840 (WRKY40) is a pathogen-induced transcription factor, and binds W-box sequences in vitro and forms protein complexes with itself and WRKY60. Coexpression with WRKY18 or WRKY60 made plants more susceptible to both *P. syringae* and *B. cinerea*. WRKY18, WRKY40, and WRKY60 have partially redundant roles in response to the hemibiotrophic bacterial pathogen *Pseudomonas syringae* and the necrotrophic fungal pathogen *Botrytis cinerea*, with WRKY18 playing a more important role than the other two (https://www.arabidopsis.org/servlets/TairObject?id=29182&type=locus).

AT3G14210, also known as ESM1 or epithiospecifier modifier 1, is known as a semi-dominant QTL, which has an epistatic effect on the epithiospecifier gene. This gene

represses nitrile formation and favours isothiocyanate production during glucosinolate hydrolysis, and its functional allele deters insect herbivory (Zhang et al. 2006).This gene was found up-regulated in CWR compared to LR in the presence of aphid feeding. AT5G46180 (DELTA-OAT) gene, encodes an ornithine delta-aminotransferase that is transcriptionally up-regulated in young seedlings and in response to salt stress (https://www.arabidopsis.org/servlets/TairObject?id=132945&type=locus). The gene AT4G33420, a peroxidase superfamily protein is also involved in oxidative stress, nitrate transport, and oxidation-reduction process. (https://www.arabidopsis.org/servlets/TairObject?name=AT4G33420&type=locus). AT3G05880 (RC12A) a RARE-COLD-INDUCIBLE 2A gene had repressed gene expression in CWR. AT1G55480 (ZKT), encodes a member of a novel plant protein family containing a PDZ, a K-box, and a TPR motif. mRNA but not protein levels decrease after wounding. ZKT is phosphorylated at Thr and Ser residues after wounding. Ishikawa et al 2005 reported a decrease level of ZKT mRNA in response to wounding in *Arabidopsis.* This gene was -1.6 fold down regulated in response to aphid feeding in CWR. All these genes and many others (table S5.3 and S5.6) showing significant changes in gene expression may yield valuable information and may be used as candidates for aphid resistance in Brassica.

5.5.2 Crop Wild Relatives and their potential for crop improvement.

Plant domestication and agronomic selection for increased yield may have had an associated effect of reducing plant defence against herbivorous insects, and evident in LR. Rosenthal & Dirzo(1997) proposed that for plants in which domestication and crop development constitute strong selection for increased growth and reproduction, reallocation of resources may result and indeed have resulted in, lower defence against

insects. This trade-off has been reported by (Wu & Baldwin 2010). It has frequently been observed that some domesticated plants have lower levels of defensive chemicals (and other structures interpretable as defensive) than do their wild relatives (Evans, 1996). Wild plant relatives or even less domesticated landraces are therefore a promising source of traits that could enable crop plants to withstand insect attack and other stressful conditions. Disease resistance genes have been reported in tomato wild relatives and used in commercial cultivars (Chetelat et al. 1995). In millet *Pyricularia grisea* resistance was introgressed from its wild relative *Pennisetum glaucum* subsp. *monodii* (Burton & Wilson 1995). Similarly in bananas, resistance against Black Sigatoka. caused by the fungus *Mycosphaerella fijiensis* has been developed from its wild relative *Musa acuminata* most commonly known as "Calcutta 4" (Hajjar & Hodgkin 2007). Finding new sources of resistance against insect pests by evaluating CWR and identifying candidate resistance genes may enable us to develop more resistant varieties in future.

In the current study an effort has been made to identify the candidate genes which when fully evaluated for function, may result in a source of new insect resistance varieties. New strategies are required to identify resistance traits and introgress or engineer them into domesticated crop germplasm. Introgression lines of wild species into crop germplasm provide a powerful resource for bringing in the traits required. However, it is often difficult to cross crops with wild relatives when they are distantly related and there is the problem of 'linkage drag' during which undesirable genes genetically closely linked to the desired ones are brought across (Zamir 2001). Use of Next generation sequencing (NGS) technologies to generate whole genome sequences for a wide range of crop species, when combined with precise phenotyping methods, can provide

powerful and rapid tools for identifying the genetic basis of agriculturally important traits and for predicting the breeding value of individuals in a plant breeding population (Varshney et al. 2014). These approaches will greatly facilitate the identification of useful traits. However the phenotyping is often more time consuming than the genotyping and it is not always possible to evaluate (Furbank & Tester 2011; Chen et al. 2012).

In order to overcome the hurdle of large scale phenotyping, efforts are being made to develop computational models based on gene expression data from known or established crop phenotypes like known resistant and susceptible varieties against biotic stresses, tolerant and susceptible varieties against abiotic stress in crop germplasm. These models can facilitate the prediction of germplasm of unknown phenotype based on gene expression from known ones. This approach is now used in the case of disease prediction in many human diseases (Trevino et al. 2011; Bucca et al. 2004; Reis-Filho & Pusztai 2011; Cooper-Knock et al. 2012). In the next chapter an attempt will be made to develop a model for prediction using the data set obtained for the gene expression study between crop wild relatives and LR in absence of aphid feeding.

In summary, this study has proved the usefulness of transcriptomics approaches for identifying significantly differentially expressing candidate genes in the absence and presence of aphid feeding in the plant. It should be borne in mind that the analysis indicated that all 75 genes in absence of aphid feeding and 261 genes in presence of aphid feeding in this experiment are potential candidates for resistance against aphid feeding. Although this discussion has managed to cover only a few of the significantly differentially expressed genes from the lists, the results presented here are able to facilitate the prioritization of strong candidate genes for further detailed investigation

such as in functional genomics study. Perhaps most importantly, the study has also highlighted the fact that resistant genes may well have been lost during the domestication process and hence can only be found in crop wild relatives, and much less so in landraces.

5.6 CONCLUSION

This chapter presents the study of differences in the pattern of gene expression between crop wild relatives (CWR) and landraces (LR) of *Brassica*. An attempt has been made to explore the changes in gene expression patterns in response to aphid feeding in these genotypes representing wide natural genetic variation in CWR and LR crop. In addition to this, an attempt was made to look at variation in gene expression in the context of domestication of crops from wild to cultivated forms. It is often reported that domestication is predicted to reduce the resistance of plants against herbivores (Turcotte et al. 2014). There are several studies where comparisons have been made between the crop and their wild relatives to check the effect of domestication on plant defence. Most of the studies report reduced resistance in cultivated crops as compared to wild crop relative (Wink 1988; Benrey et al. 1998; Rosenthal & Dirzo 1997). Our study confirms the need to survey wild relatives for resistance, where their presence may be greater than that in landraces.

Our results also indicate towards more changes in differential gene expression in CWR as compared to LR in response to aphid feeding. Aphid feeding response was compared in CWR and LR and was found to induce different transcriptional responses. The number of stress related genes induced in CWR in response to aphid feeding were more than in the absence of aphid feeding. The uniquely up-regulated genes in induced plants and commonly up-regulated among induced and non-induced plants are expected to be

the strong candidates for insect resistance in Brassica. In addition, genes like PTR3, NAC083, TIFY7, HCHIB, RAP210 and ERD6 which are up regulated provide the strongest candidates for resistance against aphids in Brassica and are found in CWR. It is strongly suggested that identified candidate genes should be further investigated to confirm their association with aphid feeding in future. The way forward in first instance is to map these against the known aphid markers in Brassica QTL and also to create knock-outs with these candidate genes to study and confirm their function in resistance against aphids. There are a large number of genes which are still not annotated but were found significant, leaving a future possibility to investigate their function but at this time no further investigations were undertaken.

List of supplementary tables

Table S5.1: Differentially Expressed genes: Number of genes with significant different in expression at FDR corrected p- Value < 0.05 and fold change ≥1.5 to ≤ -1.5 for comparison of resistant Brassica Crop Wild relative (CWR) and Landraces (LR) in absence of aphid feeding.

Table S5.2: Functional characterization of significant differentially expressed genes between Brassica CWR and LR in absence of aphid feeding.

Table S5.3: List of GO terms under biological process for constitutive differentially expressed significant genes for comparison between CWR and LR using TAIR's Gene Ontology annotation search and functional categorization web tool analysis.

Table S5.4: Differentially Expressed genes: Number of genes with significant different in expression at FDR corrected p- Value < 0.05 and fold change ≥1.5 to ≤ -1.5 for comparison of resistant Brassica Crop Wild relative (CWR) and Landraces (LR) in presence of aphid feeding.

Table S5.5: Functional characterization of significant differentially expressed genes between resistant CWR and resistant LR of Brassica after 24 hours of aphid feeding.

Table S5.6: List of GO terms under biological process for induced (after 24 hours of aphid feeding) differentially expressed significant genes for comparison between CWR and LR using TAIR's Gene Ontology annotation search and functional categorization web tool analysis.

CHAPTER 6

Gene expression arrays as a tool for prediction of insect resistance in Brassica germplasm

6.1 Abstract

The use of 'Omics' technologies like genomics, proteomics and metabolomics have reduced the time and expense of producing better quality food crops and enabled the consistency and predictability in plant breeding. The knowledge of resistance status of a plant is the first and most important requirement for any plant breeding programme and usually accomplished via morpho-physiological studies. These studies are time and highly resources consuming. So a better and effective tool to predict the resistance status is the need of today's plant breeding programme. The current study is an attempt to develop a model based on the microarray gene expression data of known resistant and susceptible genotypes and then test this model to predict the status of known genotypes. Prediction analysis was performed using the model selection tool of Partek Genomics Suite version 6.6 (Partek Inc., St. Louis, MO, USA), software. The model is based on discriminant analysis as classification method; ANOVA as variable selection method; nearest centroid and K-nearest neighbour followed by 2-level 5x5 cross-validation was performed to get the best model for each species prediction. The results from this study show that gene expression data can be successfully used in the prediction of resistance or susceptibility status in unknown genotypes present in the germplasm. Even though there is a strong separation of the resistant and susceptible group which is a proof of good model building, but still further validation is required before it can be used as a benchmark for the resistance and susceptibility for aphid feeding in Brassicas. The results from this study are promising, and can be further developed to have a single consistent model based on gene expression data in response to cabbage aphid feeding of established resistant and susceptible genotypes which can be applied to all datasets to test cabbage aphid resistance in Brassica plants.

6.2 Introduction

With ever growing human population, it is estimated that we need to increase the agricultural production by 70% in the next four decades to feed world population (FAO 2009). One way of doing this is by increasing the production of crops by developing new, more durable and resistant varieties of crops which can resist both biotic and abiotic stresses. The plant breeding programmes need to be re-oriented towards the need of these 'smart' crops which can protect themselves from environmental stresses (Mba et al. 2012). Knowing the status of tolerance or resistance of plants to various biotic and abiotic stresses is the essential prerequisite for any crop improvement programmes and usually accomplished via morpho-physiological studies which require extensive planning and design, laborious experiments and collection of data for a number of different growth, physiological, metabolic and stress parameters etc. These screening techniques, although very time and labour consuming have served the purpose to recognise resistant and tolerant genotypes in many crops species e.g. Rice LR Pokkali is tolerant to salt, CWR *B.fruticulosa* reported resistant to insect pests (Singh et al. 2014; Kumar et al. 2011). The knowledge of known resistant and susceptible genotypes for various biotic and abiotic stresses in crop species is exploited for crop improvement.

With the advances in the genomic and post-genomic technologies, it is now possible to have a better understanding of genomic, proteomic and metabolomic datasets to obtain a genome wider picture of molecular functions of various important genes. In spite of these advances, there is still a lack of knowledge regarding the available source of resistance to many biotic and abiotic factors present in plants. Researchers use several phenomic, metabolomic, and proteomic techniques to know the status of resistance of plants to biotic and abiotic stresses.

Gene expression analysis is widely used these days to understand plant-insect interactions (Hancock et al. 2015; Kerchev et al. 2012; Broekgaarden et al. 2007; Narusaka et al. 2006; Poelman et al. 2010; Poelman 2015). The whole genome microarray data is used to detect differentially expressed genes in organisms under different conditions (Torres-Avilés et al. 2014). Gene expression microarrays generate data for thousands of genes simultaneously (Duggan et al. 1999; Lockhart & Winzeler 2000). Microarray data from several plants have been used to identify genes responsive to insect feeding (Reymond et al. 2004; Reymond 2000; Elzinga et al. 2014). Jaouannet et al. (2015) characterised Arabidopsis transcriptional response to different aphid species in order to reveal genes that contribute to host susceptibility and non-host resistance. Kettles et al. (2013) reported resistance of Arabidopsis to green peach aphid involves camalexin. Similarly, there are many studies in other crop species also where gene expression data is used to improve crop varieties. Recently the submergence tolerance genes sub1A, sub1B and sub1C were identified in some wild species O. rufipogon and O. nivara and landraces e.g., Kurkaruppan and Goda Heenati (Bailey-Serres et al. 2010; Fukao et al. 2009) and were successfully utilized for developing submergence tolerant crop varieties such as FR13A and FR43B and breeding lines such as Thalavu, BKNFR76106-16-0-1-0 and IR49830-7-1-2-1-3. There is much evidence that transcriptional studies are helpful in finding the genes which can prove useful in generating more resistant crops to both biotic and abiotic stresses (Redden 2013; Yumurtaci 2015; Delp et al. 2009; Kuśnierczyk et al. 2008; Dubey et al. 2013; Poelman et al. 2010).

Breeders are utilising this phenomic screening and genome expression information to look for candidate genes to improve the crop quality (Ashkani et al. 2015; Ray & Satya

2014; Pérez-de-Castro et al. 2012), but still large numbers of crop species remain unscreened for the levels of resistance present in them against biotic and abiotic stresses. To screen all the crop species against all stresses is almost an impossible task. It will require a huge investment in terms of resources and time to screen each and every accession for each species available in field or laboratory. So quick, better and reliable methods are needed to know the status of resistance against insects in plants. The use of transcriptomic data can potentially be exploited to predict the tolerance or resistance status of genotypes of unknown status by computational model-based comparisons with the genotypes where tolerance or resistance status is known (Gavaghan et al. 2002). More recently whole genome expression profiling is a highly accepted approach in classifying and performing class prediction in case of many human diseases. This approach is very popular especially in the case of prediction of various human cancers and neurodegenerative diseases (Bucca et al. 2004; Reis-Filho & Pusztai 2011; Cooper-Knock et al. 2012). Trevino et al. (2011) successfully demonstrated the use of gene expression data and statistical modelling in the prediction of prostate cancer and the use of this approach as a diagnostic tool. Proteomic alteration and transcriptomics data from a study of metastatic and clinically localised prostate cancer samples are used to develop disease predictor model (Varambally et al. 2005). Li et al. (2009) used unsupervised machine learning methods to genome-wide gene expression profiles of 159 gliomas, thereby establishing a robust glioma classification model relying only on the molecular data. The classification based algorithm method is reported to predict toxicity between controls and treated liver samples of rats after repeated doses for 149 compounds (Nagata et al. 2014). Despite the potentiality of the transcriptomics and proteomics data to be used for this prediction analysis, so far there is no such report for

prediction of the biotic and abiotic stress responses of crop species. The use of gene expression data as a predictor of resistance or susceptibility to a biotic or abiotic stress can revolutionise discovery of new resistant crops, but it has yet to be used in the field of crop science.

The current chapter describes the development of discriminant analysis prediction model based on gene expression data of the known resistant and susceptible genotypes of Brassica and the prediction of resistance/susceptible status of unknown genotypes using diagnostic and predictive modelling tool of Partek Genomic Suite version 6.6 (Partek Inc., St. Louis, MO, USA), software. This predictive modelling tool used multivariate statistical analysis to classify the samples into distinct classes based on the various variables selected during the model generation. Unlike other prediction models where extensive programming and statistical skills are required, Partek predictive modelling tool is very simple, user-friendly method which was successfully used in developing classification model based on gene expression data in clinically defined dengue patients to predict disease severity (Sun et al. 2013). The gene expression data used in this chapter is same as generated in chapter 4. The same dataset was used in chapter 5 to analyse gene expression changes between CWR and LR in response to aphid feeding and now in this chapter to develop prediction model indicates towards different uses of 'omics' data to answer different research questions. The dataset from 10 Brassica genotypes under non-induced condition i.e. absence of aphid feeding is used to develop a model. The status of genotypes is classified as resistant or susceptible is based on aphid feeding preference in the field as described in chapter 2. This is the preliminary study, in an attempt to develop a prediction model for Brassica genotypes for classification of resistance status. The classification model was developed using

class prediction tool of Partek software, as it was a trial version so some of the functionalities like cross-validation of grouping with PCA analysis was not possible. The model developed is at its infant stage and requires further validations, but opens a new avenue for future research.

6.2.1 Chapter aim

The core aim of the study is the development of a prediction model based on the whole genome expression data available from genotypes of known resistance or susceptibility status to predict the resistance status of unknown genotypes.

The prediction model developed will include 8 out of 10 genotypes at a time, leaving two to be tested. This way the gene expression data from 8 genotypes grouped as resistant or susceptible will be used to check the status of 2 blinded genotypes. Leaving one species out (2 genotypes from same species) will provide an opportunity to test the accuracy of the model independently for 5 times.

6.3 Material and methods

6.3.1 Microarray Data

The raw data (.CEL) files used in this study were obtained from microarray experiment to study the gene expression changes between the resistant and susceptible genotypes as described in chapter 4. The data set comprised of gene expression data of 10 genotypes from 5 Brassica species (*B.incana, B.montana, B.villosa, B.oleracea acephala, B.fruticulosa*). The RNA extracted from the leaf samples was hybridised to Affymetrix Arabidopsis Gene 1.0 ST array (Affymetrix, Santa Clara, CA, USA).The details of the experiment are already provided in chapter 4 (section 4.3). For this study, the gene expression data from non-induced plants i.e. absence of aphid feeding is used. The genotypes are categorised as resistant or susceptible based on the field experiment as

described in chapter 2. The classification was generated independently for *B.incana, B.montana, B.villosa, B.oleracea acephala, B.fruticulosa*.

6.3.2 Classification Method

The class prediction analysis was performed using the model selection tool of Partek Genomics Suite version 6.6 (Partek Inc., St. Louis, MO, USA), software. The 2 level nested, cross-validation approach was used to select the best classification model along with the accuracy estimation for the model to test new data set. At this level the inner cross validation was done to choose the predictor variables whereas an outer cross validation was done to estimate % accuracy for classifier. At level 1 cross –validation, evaluation of multiple models generated from test data was done and the best model to deploy was selected while 2 level cross-validations provided the accuracy estimate. For this study, Partek software's saved specification for model selection, based on discriminant analysis as classification method; ANOVA as variable selection method; nearest centroid and K-nearest neighbour followed by 2-level 5x5 cross-validation was performed to get the best model for each species prediction. At the end of this step, an estimate of overall accuracy level (%) for models was generated. This % estimate indicated the correct rate at which the unknown sample will be predicted even when the best suited model is selected. In next step 1 level-cross validation is performed to select the best model from all the models generated to deploy in order to test unknown samples. At this stage the model with 100% rating is selected from all the available models, 100% is not the actual estimate, exact estimate will always be the % accuracy generated in 2 level estimates. Here we may be presented with more than 1 model with 100% correct rate, so one or all models can be tested. The best classification model selected for this study with highest % accuracy was based following classifiers: 100

variables (top 100 genes with differential gene expression between two classes), K-nearest neighbour with Euclidean distance measure, and 1 neighbour with nearest centroid, Discriminant analysis and support vector machine.

Table 6.1: Details of best Model with the classifiers for prediction resistance status of different Brassica species

Experiment Description	
Input File	Model 1
Variable to Predict	2.Type
# of Predictor Candidates	179298
# of Samples	32
# of Models	124
Random Seed	10001
Data Order	Randomly reordered
Model Selection Criterion	Normalized correct rate
Cross-Validation	2-level Nested
Outer Partitions	5
Inner Partitions	5
Variable Selection	
Variable Selection Method	ANOVA
Examine	1-way ANOVA p-value (Type)
How many groups of variables	19 groups with manually specified sizes 1-10 10-100-10
Classification	
Classification Method	K-Nearest Neighbor
Distance Measure	Absolute Value/City Block
	Euclidean
# of Neighbors (K) : 1	1
Classification Method	Nearest Centroid
Prior Probability	Equal
	Proportional
Classification Method	Discriminant Analysis
Discriminant Function and Prior Probabilities	Linear with Proportional Prior Probabilities
	Quadratic with Equal Prior Probabilities
	Quadratic with Proportional Prior Probabilities
Positive Outcome	Resistant

Table 6.1 shows the generalised best classification model with classifiers. The selected model (based on gene expression from 8 genotypes) was then saved as test file (.pbb) format. To test independent unknown samples (2 remaining genotypes), the saved

(.pbb) model was deployed. The data set format for test and save model should be in the same format to run the test. In this study **TYPE** (Resistant /susceptible) column is used as prediction criteria. This column was defined (resistant or susceptible) in case of selected model whereas it is labelled unknown in the case of test data set. In order to test the unknown sample the saved model was deployed and run. This will invoke an HTML page with details of the test conducted and result of the run.

To test the accuracy and validity of this model selection tool, the whole data set comprising of data from 10 genotypes was tested 5 times independently by leaving one species (2 genotypes) out at a time. For example to test status of *B.fruticulosa,* the saved test model was generated from gene expression data from 8 genotypes (BRA2856, K10373, K9404, BRA1644, K6926, K10259, BRS-0103 & CGN18468) and unknown prediction set contains gene expression data from *B.fruticulosa (*Bol2009-0080 & Bol2009-0081*).* Similarly this test was successfully performed for all species. To test status of *B.incana*, the saved test model was generated from gene expression data from 8 genotypes *(*K9404, BRA1644, K6926, K10259, BRS-0103, CGN18468, Bol2009-0080 & Bol2009-0081) and unknown prediction set contains gene expression data from *B.incana (*BRA2856,K10373*).* To test status of *B.montana*, the saved test model was generated from gene expression data from 8 genotypes (BRA2856, K10373, K6926, K10259, BRS-0103, CGN18468, Bol2009-0080 & Bol2009-0081) and unknown prediction set contains gene expression data from *B.montana (*K9404, BRA1644*).* To test status of *B.villosa,* the saved test model was generated from gene expression data from 8 genotypes (BRA2856, K10373, BRS-0103, CGN18468,K9404, BRA1644 Bol2009-0080 & Bol2009-0081) and unknown prediction set contains gene expression data from *B.villosa (*K6926, K10259). To test status of *B. oleracea acephala,* the saved

test model was generated from gene expression data from 8 genotypes (BRA2856, K10373, K6926, K10259, K9404, BRA1644, Bol2009-0080 & Bol2009-0081) and unknown prediction set contains gene expression data from *B.villosa (*BRS-0103 & CGN18468,).

6.4 Results

Among the fifteen genotypes that were studied transcriptomically (chapter 4), 10 were used in this study to develop and test the prediction model generated using class prediction tool of Partek software. Ten genotypes used in the study were grouped as resistant (n=7) or susceptible (n=3) as per their performance against aphid feeding in the field experiment (chapter 2). They were further subjected to transcriptomics analysis to study the gene expression changes between two groups in the absence of aphid feeding. The microarray gene expression data of resistant and susceptible groups was then used to develop a prediction model. This model was used to validate the suitability of these genotypes to insect feeding (resistant/susceptible). The 'leave one species out at time' approach was adopted to test the status of resistance in genotypes of that species.

6.4.1 Building of Classification model

For the purpose of this study, the focus was on genes which were significantly different in their expression in resistant and susceptible varieties. Based on model selection criteria, the top 100 genes significantly differentially expressing between 2 groups were selected for each model, which separate resistant genotypes from the susceptible ones. These 100 genes were used to build prediction models to test the unknown genotypes. The tables (table 6.2 to 6.5) show the list of genes (annotated) selected for developing the prediction model. The complete list of 100 variables (genes) selected for each model

is provided in supplementary table S6.1 (*B.incana model*), S6.2 (*B.montana model*), S6.3 (B.*villosa model*) S6.4 (*B.fruticulosa model*) and S6.5 (*B.oleracea model*).

6.4.2 Testing and prediction of unknown samples using classification model with 100 selected gene

The gene expression data from the two genotypes left out during model generation is used as test dataset saved in the same format for rest of 8 genotypes used to develop the model. This data set is then tested against the model developed using in-house save the script in Partek software as previously described in section 6.3.3. The saved (.pbb) model was then deployed to create predictions. This procedure resulted in very good prediction for 8 genotypes out of 10. After cross-validation, the accuracy of the model was found to be varying in a range of 98-100% accuracy in predicting resistant versus susceptible genotypes with an error rate of ± 2%. This result indicates gene expression data can be successfully used to create a prediction model. However, as a very small number of samples were used to develop the model, it is difficult to determine whether this model is over-fit.

It was observed that model predicted mismatch for two genotypes. The genotype Bol2009-008 *(B.fruticulosa*) which was found to be resistant in field assessment but model predicted it as susceptible and the second genotype was *B.villosa* K6926 predicted as resistant was categorised as susceptible in the field experiment.

Table 6.2: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.incana* species using the Partek model selection tool.

S.no	Probeset_id	AGI_code	Gene assignment
1	13338945	AT1G11720	NM_101044 // SS3 /// NM_001198036 // SS3 /// AT1G11720.1 // SS3 /// AT1G11720.2 // SS3
2	13351724	AT1G53570	NM_001036104 // MAP3KA /// NM_104235 // MAP3KA /// NM_179472 // MAP3KA /// NM_001198282 // MAP3KA /// NM_001198283 // MAP3KA /// AT1G53570.1 // MAP3KA /// AT1G53570.2 // MAP3KA /// AT1G53570.5 // MAP3KA /// AT1G53570.3 // MAP3KA /// AT1G53570.4 // MAP3KA /// AY140005 // MAP3KA
3	13361237	AT1G79840	NM_106633 // GL2 /// NM_001198514 // GL2 /// AT1G79840.2 // GL2 /// AT1G79840.1 // GL2 /// AF360294 // GL2
4	13361240	AT1G79840	NM_106633 // GL2 /// NM_001198514 // GL2 /// AT1G79840.2 // GL2 /// AT1G79840.1 // GL2 /// AF360294 // GL2
5	13361883	AT1G01140	NM_099996 // CIPK9 /// NM_179240 // CIPK9 /// NM_179239 // CIPK9 /// AT1G01140.1 // CIPK9 /// AT1G01140.2 // CIPK9 /// AT1G01140.3 // CIPK9 /// AF295664 // CIPK9
6	13367998	AT1G14690	NM_101338 // MAP65-7 /// NM_001084068 // MAP65-7 /// AT1G14690.2 // MAP65-7 /// AT1G14690.1 // MAP65-7 /// AY120768 // MAP65-7
7	13368073	AT1G14810	NM_101350 // AT1G14810 /// AT1G14810.1 // AT1G14810 /// AY070759 // AT1G14810
8	13370122	AT1G19660	NM_101822 // AT1G19660 /// NM_001035991 // AT1G19660 /// AT1G19660.2 // AT1G19660 /// AT1G19660.1 // AT1G19660 /// AK317508 // AT1G19660
9	13370199	AT1G19835	NM_101838 // AT1G19835 /// NM_001198113 // AT1G19835 /// AT1G19835.1 // AT1G19835 /// AT1G19835.2 // AT1G19835

10	13374044	AT1G30280	NM_102767 // AT1G30280 /// AT1G30280.1 // AT1G30280 /// AY070092 // AT1G30280
11	13377908	AT1G50490	NM_103932 // UBC20 /// AT1G50490.1 // UBC20 /// AK227382 // UBC20
12	13380818	AT1G58200	NM_202317 // MSL3 /// NM_104601 // MSL3 /// AT1G58200.1 // MSL3 /// AT1G58200.2 // MSL3 /// AY125504 // MSL3
13	13383170	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
14	13383173	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
15	13384286	AT1G67120	NM_105382 // AT1G67120 /// AT1G67120.1 // AT1G67120
16	13384916	AT1G68410	NM_001036175 // AT1G68410 /// NM_105512 // AT1G68410 /// AT1G68410.2 // AT1G68410 /// AT1G68410.1 // AT1G68410 /// AY050881 // AT1G68410
17	13386388	AT1G72430	NM_105902 // AT1G72430 /// AT1G72430.1 // AT1G72430
18	13388178	AT1G76540	NM_106304 // CDKB2;1 /// AT1G76540.1 // CDKB2;1 /// AB047279 // CDKB2;1
19	13390427	AT2G01340	NM_126195 // At17.1 /// AT2G01340.1 // At17.1 /// BT021987 // At17.1
20	13395074	AT2G20570	NM_127617 // GPRI1 /// NM_001202633 // GPRI1 /// AT2G20570.1 // GPRI1 /// AT2G20570.2 // GPRI1 /// AB062489 // GPRI1
21	13395076	AT2G20570	NM_127617 // GPRI1 /// NM_001202633 // GPRI1 /// AT2G20570.1 // GPRI1 /// AT2G20570.2 // GPRI1 /// AB062489 // GPRI1

22	13399152	AT2G30950	NM_179825 // VAR2 /// AT2G30950.1 // VAR2 /// AF135189 // VAR2
23	13401267	AT2G35800	NM_179921 // AT2G35800 /// AT2G35800.1 // AT2G35800 /// AY063948 // AT2G35800
24	13401269	AT2G35800	NM_179921 // AT2G35800 /// AT2G35800.1 // AT2G35800 /// AY063948 // AT2G35800
25	13401305	AT2G35860	NM_179922 // FLA16 /// AT2G35860.1 // FLA16 /// AY093189 // FLA16
26	13406073	AT2G47170	NM_130285 // ARF1A1C /// AT2G47170.1 // ARF1A1C /// M95166 // ARF1A1C
27	13406530	AT2G47980	NM_130365 // SCC3 /// AT2G47980.1 // SCC3 /// AY063915 // SCC3
28	13411231	AT2G21140	NM_127684 // PRP2 /// AT2G21140.1 // PRP2 /// AK117333 // PRP2
29	13412816	AT2G25540	NM_128111 // CESA10 /// AT2G25540.1 // CESA10
30	13413792	AT2G27450	NM_179770 // NLP1 /// NM_128305 // NLP1 /// AT2G27450.2 // NLP1 /// AT2G27450.1 // NLP1 /// AY072113 // NLP1

Table 6.3: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.montana* species using the Partek model selection tool.

S.no	probeset_id	AGI_code	gene_assignment
1	13335913	AT1G04778	NM_001083989 // AT1G04778 /// AT1G04778.1 // AT1G04778
2	13338945	AT1G11720	NM_101044 // SS3 /// NM_001198036 // SS3 /// AT1G11720.1 // SS3 /// AT1G11720.2 // SS3
3	13348001	AT1G35220	NM_103220 // AT1G35220 /// AT1G35220.1 // AT1G35220
4	13356385	AT1G67900	NM_202374 // AT1G67900 /// NM_105460 // AT1G67900 /// NM_001198418 // AT1G67900 /// AT1G67900.2 // AT1G67900 /// AT1G67900.3 // AT1G67900 /// AT1G67900.1 // AT1G67900 /// AY120729 // AT1G67900
5	13361160	AT1G79590	NM_001036224 // SYP52 /// NM_106607 // SYP52 /// AT1G79590.2 // SYP52 /// AT1G79590.1 // SYP52 /// AF355756 // SYP52
6	13361883	AT1G01140	NM_099996 // CIPK9 /// NM_179240 // CIPK9 /// NM_179239 // CIPK9 /// AT1G01140.1 // CIPK9 /// AT1G01140.2 // CIPK9 /// AT1G01140.3 // CIPK9 /// AF295664 // CIPK9
7	13363988	AT1G05850	NM_100466 // POM1 /// AT1G05850.1 // POM1 /// AY034935 // POM1
8	13367998	AT1G14690	NM_101338 // MAP65-7 /// NM_001084068 // MAP65-7 /// AT1G14690.2 // MAP65-7 /// AT1G14690.1 // MAP65-7 /// AY120768 // MAP65-7
9	13368073	AT1G14810	NM_101350 // AT1G14810 /// AT1G14810.1 // AT1G14810 /// AY070759 // AT1G14810
10	13370122	AT1G19660	NM_101822 // AT1G19660 /// NM_001035991 // AT1G19660 /// AT1G19660.2 // AT1G19660 /// AT1G19660.1 // AT1G19660 /// AK317508 // AT1G19660

11	13370199	AT1G19835	NM_101838 // AT1G19835 /// NM_001198113 // AT1G19835 /// AT1G19835.1 // AT1G19835 /// AT1G19835.2 // AT1G19835
12	13373276	AT1G28320	NM_102597 // DEG15 /// AT1G28320.1 // DEG15
13	13373676	AT1G29400	NM_179396 // ML5 /// NM_102680 // ML5 /// AT1G29400.1 // ML5 /// AT1G29400.2 // ML5 /// AY070368 // ML5
14	13376554	AT1G47128	NM_103612 // RD21 /// AT1G47128.1 // RD21 /// AY072130 // RD21
15	13377908	AT1G50490	NM_103932 // UBC20 /// AT1G50490.1 // UBC20 /// AK227382 // UBC20
16	13380818	AT1G58200	NM_202317 // MSL3 /// NM_104601 // MSL3 /// AT1G58200.1 // MSL3 /// AT1G58200.2 // MSL3 /// AY125504 // MSL3
17	13383170	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
18	13383173	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
19	13390427	AT2G01340	NM_126195 // At17.1 /// AT2G01340.1 // At17.1 /// BT021987 // At17.1
20	13395074	AT2G20570	NM_127617 // GPRI1 /// NM_001202633 // GPRI1 /// AT2G20570.1 // GPRI1 /// AT2G20570.2 // GPRI1 /// AB062489 // GPRI1
21	13401267	AT2G35800	NM_179921 // AT2G35800 /// AT2G35800.1 // AT2G35800 /// AY063948 // AT2G35800
22	13401305	AT2G35860	NM_179922 // FLA16 /// AT2G35860.1 // FLA16 /// AY093189 // FLA16

23	13405651	AT2G46140	NM_130176 // AT2G46140 /// AT2G46140.1 // AT2G46140 /// BT004206 // AT2G46140
24	13406073	AT2G47170	NM_130285 // ARF1A1C /// AT2G47170.1 // ARF1A1C /// M95166 // ARF1A1C
25	13406475	AT2G47890	NM_130356 // AT2G47890 /// NM_201983 // AT2G47890 /// AT2G47890.2 // AT2G47890 /// AT2G47890.1 // AT2G47890 /// BT023413 // AT2G47890
26	13406840	AT2G01450	NM_001035864 // MPK17 /// NM_001035863 // MPK17 /// NM_001035862 // MPK17 /// NM_126206 // MPK17 /// AT2G01450.2 // MPK17 /// AT2G01450.3 // MPK17 /// AT2G01450.4 // MPK17 /// AT2G01450.1 // MPK17 /// AK228266 // MPK17
27	13411231	AT2G21140	NM_127684 // PRP2 /// AT2G21140.1 // PRP2 /// AK117333 // PRP2

Table 6.4: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.villosa* species using the Partek model selection tool.

S.no	probeset_id	AGI_code	gene_assignment
1	13342730	AT1G20760	NM_101928 // AT1G20760 /// AT1G20760.1 // AT1G20760
2	13343517	AT1G22770	NM_102124 // GI /// AT1G22770.1 // GI /// AF105064 // GI
3	13351638	AT1G53400	NM_104219 // AT1G53400 /// AT1G53400.1 // AT1G53400 /// AF370164 // AT1G53400

4	13351640	AT1G53400	NM_104219 // AT1G53400 /// AT1G53400.1 // AT1G53400 /// AF370164 // AT1G53400
5	13357744	AT1G71220	NM_105791 // EBS1 /// NM_001198441 // EBS1 /// AT1G71220.1 // EBS1 /// AT1G71220.2 // EBS1 /// AK230327 // EBS1
6	13361160	AT1G79590	NM_001036224 // SYP52 /// NM_106607 // SYP52 /// AT1G79590.2 // SYP52 /// AT1G79590.1 // SYP52 /// AF355756 // SYP52
7	13361883	AT1G01140	NM_099996 // CIPK9 /// NM_179240 // CIPK9 /// NM_179239 // CIPK9 /// AT1G01140.1 // CIPK9 /// AT1G01140.2 // CIPK9 /// AT1G01140.3 // CIPK9 /// AF295664 // CIPK9
8	13364212	AT1G06460	NM_100526 // ACD32.1 /// AT1G06460.1 // ACD32.1 /// DQ403189 // ACD32.1
9	13364215	AT1G06460	NM_100526 // ACD32.1 /// AT1G06460.1 // ACD32.1 /// DQ403189 // ACD32.1
10	13365772	AT1G09730	NM_100845 // AT1G09730 /// NM_001198022 // AT1G09730 /// AT1G09730.1 // AT1G09730 /// AT1G09730.2 // AT1G09730 /// EU877962 // AT1G09730
11	13366227	AT1G10550	NM_100930 // XTH33 /// AT1G10550.1 // XTH33
12	13367998	AT1G14690	NM_101338 // MAP65-7 /// NM_001084068 // MAP65-7 /// AT1G14690.2 // MAP65-7 /// AT1G14690.1 // MAP65-7 /// AY120768 // MAP65-7

13	13369917	AT1G19050	NM_101763 // ARR7 /// AT1G19050.1 // ARR7 /// AB008490 // ARR7
14	13370122	AT1G19660	NM_101822 // AT1G19660 /// NM_001035991 // AT1G19660 /// AT1G19660.2 // AT1G19660 /// AT1G19660.1 // AT1G19660 /// AK317508 // AT1G19660
15	13370199	AT1G19835	NM_101838 // AT1G19835 /// NM_001198113 // AT1G19835 /// AT1G19835.1 // AT1G19835 /// AT1G19835.2 // AT1G19835
16	13373676	AT1G29400	NM_179396 // ML5 /// NM_102680 // ML5 /// AT1G29400.1 // ML5 /// AT1G29400.2 // ML5 /// AY070368 // ML5
17	13374044	AT1G30280	NM_102767 // AT1G30280 /// AT1G30280.1 // AT1G30280 /// AY070092 // AT1G30280
18	13380818	AT1G58200	NM_202317 // MSL3 /// NM_104601 // MSL3 /// AT1G58200.1 // MSL3 /// AT1G58200.2 // MSL3 /// AY125504 // MSL3
19	13383170	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
20	13383173	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
21	13384916	AT1G68410	NM_001036175 // AT1G68410 /// NM_105512 // AT1G68410 /// AT1G68410.2 // AT1G68410 /// AT1G68410.1 // AT1G68410 /// AY050881 // AT1G68410

22	13388178	AT1G76540	NM_106304 // CDKB2;1 /// AT1G76540.1 // CDKB2;1 /// AB047279 // CDKB2;1
23	13389639	AT1G79610	NM_106609 // NHX6 /// AT1G79610.1 // NHX6 /// AY091100 // NHX6
24	13390427	AT2G01340	NM_126195 // At17.1 /// AT2G01340.1 // At17.1 /// BT021987 // At17.1
25	13395074	AT2G20570	NM_127617 // GPRI1 /// NM_001202633 // GPRI1 /// AT2G20570.1 // GPRI1 /// AT2G20570.2 // GPRI1 /// AB062489 // GPRI1
26	13395076	AT2G20570	NM_127617 // GPRI1 /// NM_001202633 // GPRI1 /// AT2G20570.1 // GPRI1 /// AT2G20570.2 // GPRI1 /// AB062489 // GPRI1
27	13395517	AT2G21530	NM_127725 // AT2G21530 /// AT2G21530.1 // AT2G21530 /// AK118718 // AT2G21530
28	13398162	AT2G28290	NM_179785 // SYD /// NM_001084502 // SYD /// NM_179786 // SYD /// AT2G28290.1 // SYD /// AT2G28290.2 // SYD /// AT2G28290.3 // SYD
29	13399152	AT2G30950	NM_179825 // VAR2 /// AT2G30950.1 // VAR2 /// AF135189 // VAR2
30	13401267	AT2G35800	NM_179921 // AT2G35800 /// AT2G35800.1 // AT2G35800 /// AY063948 // AT2G35800

31	13406073	AT2G47170	NM_130285 // ARF1A1C /// AT2G47170.1 // ARF1A1C /// M95166 // ARF1A1C
32	13406530	AT2G47980	NM_130365 // SCC3 /// AT2G47980.1 // SCC3 /// AY063915 // SCC3
33	13407933	AT2G05120	NM_126541 // AT2G05120 /// NM_001202582 // AT2G05120 /// AT2G05120.1 // AT2G05120 /// AT2G05120.2 // AT2G05120
34	13409122	AT2G15890	NM_127149 // MEE14 /// AT2G15890.1 // MEE14 /// AF324713 // MEE14
35	13409123	AT2G15890	NM_001084426 // MEE14 /// NM_127149 // MEE14 /// AT2G15890.1 // MEE14 /// AT2G15890.2 // MEE14 /// AF324713 // MEE14
36	13410028	AT2G18300	NM_127388 // AT2G18300 /// NM_179645 // AT2G18300 /// NM_001202619 // AT2G18300 /// AT2G18300.1 // AT2G18300 /// AT2G18300.2 // AT2G18300 /// AT2G18300.3 // AT2G18300 /// AF412099 // AT2G18300
37	13411231	AT2G21140	NM_127684 // PRP2 /// AT2G21140.1 // PRP2 /// AK117333 // PRP2
38	13411869	AT2G23070	NM_127871 // AT2G23070 /// AT2G23070.1 // AT2G23070 /// AY062631 // AT2G23070
39	13412228	AT2G24120	NM_127973 // SCA3 /// NM_201796 // SCA3 /// AT2G24120.1 // SCA3 /// AT2G24120.2 // SCA3 /// AK229652 // SCA3

40	13412230	AT2G24120	NM_127973 // SCA3 /// NM_201796 // SCA3 /// AT2G24120.1 // SCA3 /// AT2G24120.2 // SCA3 /// AK229652 // SCA3
41	13412241	AT2G24120	NM_127973 // SCA3 /// NM_201796 // SCA3 /// AT2G24120.1 // SCA3 /// AT2G24120.2 // SCA3 /// AK229652 // SCA3

Table 6.5: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.fruticulosa* species using the Partek model selection tool.

S.No	probeset_id	AGI_code	gene_assignment
1	13335913	AT1G04778	NM_001083989 // AT1G04778 /// AT1G04778.1 // AT1G04778
2	13341938	AT1G18730	NM_101731 // NDF6 /// NM_001123835 // NDF6 /// NM_001084090 // NDF6 /// NM_001123834 // NDF6 /// AT1G18730.1 // NDF6 /// AT1G18730.2 // NDF6 /// AT1G18730.3 // NDF6 /// AT1G18730.4 // NDF6 /// AY080701 // NDF6
3	13353700	AT1G60060	NM_104698 // AT1G60060 /// AT1G60060.1 // AT1G60060 /// AY084785 // AT1G60060
4	13353704	AT1G60060	NM_104698 // AT1G60060 /// AT1G60060.1 // AT1G60060 /// AY084785 // AT1G60060
5	13356388	AT1G67900	NM_202374 // AT1G67900 /// NM_105460 // AT1G67900 /// NM_001198418 // AT1G67900 /// AT1G67900.2 // AT1G67900 /// AT1G67900.3 // AT1G67900 /// AT1G67900.1 // AT1G67900 /// AY120729 // AT1G67900
6	13361883	AT1G01140	NM_099996 // CIPK9 /// NM_179240 // CIPK9 /// NM_179239 // CIPK9 /// AT1G01140.1 // CIPK9 /// AT1G01140.2 // CIPK9 /// AT1G01140.3 // CIPK9 /// AF295664 // CIPK9

7	13362791	AT1G03475	NM_100230 // LIN2 /// AT1G03475.1 // LIN2 /// AB044394 // LIN2
8	13363512	AT1G04980	NM_100376 // PDIL2-2 /// AT1G04980.1 // PDIL2-2
9	13366745	AT1G11545	NM_101028 // XTH8 /// AT1G11545.1 // XTH8 /// AK228427 // XTH8
10	13366783	AT1G11684	---
11	13367834	AT1G14410	NM_101308 // WHY1 /// AT1G14410.1 // WHY1 /// AF370156 // WHY1
12	13374044	AT1G30280	NM_102767 // AT1G30280 /// AT1G30280.1 // AT1G30280 /// AY070092 // AT1G30280
13	13375019	AT1G32900	NM_103023 // AT1G32900 /// AT1G32900.1 // AT1G32900 /// AY123983 // AT1G32900
14	13383170	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
15	13383173	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
16	13385513	AT1G69840	NM_105652 // AT1G69840 /// NM_001084333 // AT1G69840 /// NM_202388 // AT1G69840 /// NM_001084334 // AT1G69840 /// NM_179539 // AT1G69840 /// NM_202387 // AT1G69840 /// NM_001198429 // AT1G69840 /// AT1G69840.7 // AT1G69840 /// AT1G69840.3 // AT1G69840 /// AT1G69840.2 // AT1G69840 /// AT1G69840.6 // AT1G69840 /// AT1G69840.4 // AT1G69840 /// AT1G69840.5 // AT1G69840 /// AT1G69840.1 // AT1G69840 /// AY099840 // AT1G69840

17	13386190	AT1G71830	NM_105841 // SERK1 /// AT1G71830.1 // SERK1
18	13386388	AT1G72430	NM_105902 // AT1G72430 /// AT1G72430.1 // AT1G72430
19	13387951	AT1G76110	NM_106260 // AT1G76110 /// AT1G76110.1 // AT1G76110 /// AY099630 // AT1G76110
20	13388178	AT1G76540	NM_106304 // CDKB2;1 /// AT1G76540.1 // CDKB2;1 /// AB047279 // CDKB2;1
21	13388517	AT1G77330	NM_106382 // AT1G77330 /// AT1G77330.1 // AT1G77330 /// BT026444 // AT1G77330
22	13390427	AT2G01340	NM_126195 // At17.1 /// AT2G01340.1 // At17.1 /// BT021987 // At17.1
23	13390961	AT2G02740	NM_001084403 // WHY3 /// NM_126329 // WHY3 /// AT2G02740.1 // WHY3 /// AT2G02740.2 // WHY3 /// BT015336 // WHY3
24	13391436	AT2G04030	NM_126439 // CR88 /// NM_179601 // CR88 /// AT2G04030.1 // CR88 /// AT2G04030.2 // CR88 /// AF436826 // CR88
25	13394435	AT2G18960	NM_127453 // HA1 /// AT2G18960.1 // HA1 /// BT008692 // HA1
26	13399882	AT2G32487	NM_001084525 // AT2G32487 /// AT2G32487.1 // AT2G32487
27	13401267	AT2G35800	NM_179921 // AT2G35800 /// AT2G35800.1 // AT2G35800 /// AY063948 // AT2G35800

28	13401269	AT2G35800	NM_179921 // AT2G35800 /// AT2G35800.1 // AT2G35800 /// AY063948 // AT2G35800
29	13401305	AT2G35860	NM_179922 // FLA16 /// AT2G35860.1 // FLA16 /// AY093189 // FLA16
30	13406073	AT2G47170	NM_130285 // ARF1A1C /// AT2G47170.1 // ARF1A1C /// M95166 // ARF1A1C
31	13407793	AT2G04540	NM_126485 // AT2G04540 /// AT2G04540.1 // AT2G04540 /// AB073746 // AT2G04540
32	13410413	AT2G19560	NM_127514 // EER5 /// AT2G19560.1 // EER5 /// AK118897 // EER5
33	13411231	AT2G21140	NM_127684 // PRP2 /// AT2G21140.1 // PRP2 /// AK117333 // PRP2
34	13413792	AT2G27450	NM_179770 // NLP1 /// NM_128305 // NLP1 /// AT2G27450.2 // NLP1 /// AT2G27450.1 // NLP1 /// AY072113 // NLP1

Table 6.6: List of selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of *B.oleracea acephala* species using the Partek model selection tool.

S.no	probeset_id	AGI_code	gene_assignment
1	13361883	AT1G01140	NM_099996 // CIPK9 /// NM_179240 // CIPK9 /// NM_179239 // CIPK9 /// AT1G01140.1 // CIPK9 /// AT1G01140.2 // CIPK9 /// AT1G01140.3 // CIPK9 /// AF295664 // CIPK9
2	13335913	AT1G04778	NM_001083989 // AT1G04778 /// AT1G04778.1 // AT1G04778
3	13366227	AT1G10550	NM_100930 // XTH33 /// AT1G10550.1 // XTH33

4	13366314	AT1G10740	NM_100950 // AT1G10740 /// NM_202077 // AT1G10740 /// NM_001198027 // AT1G10740 /// NM_001198026 // AT1G10740 /// AT1G10740.3 // AT1G10740 /// AT1G10740.4 // AT1G10740 /// AT1G10740.1 // AT1G10740 /// AT1G10740.2 // AT1G10740 /// AY120749 // AT1G10740
5	13338945	AT1G11720	NM_101044 // SS3 /// NM_001198036 // SS3 /// AT1G11720.1 // SS3 /// AT1G11720.2 // SS3
6	13367434	AT1G13350	NM_101206 // AT1G13350 /// NM_001198053 // AT1G13350 /// AT1G13350.1 // AT1G13350 /// AT1G13350.2 // AT1G13350
7	13370122	AT1G19660	NM_101822 // AT1G19660 /// NM_001035991 // AT1G19660 /// AT1G19660.2 // AT1G19660 /// AT1G19660.1 // AT1G19660 /// AK317508 // AT1G19660
8	13370199	AT1G19835	NM_101838 // AT1G19835 /// NM_001198113 // AT1G19835 /// AT1G19835.1 // AT1G19835 /// AT1G19835.2 // AT1G19835
9	13370635	AT1G20820	---
10	13374044	AT1G30280	NM_102767 // AT1G30280 /// AT1G30280.1 // AT1G30280 /// AY070092 // AT1G30280
11	13348001	AT1G35220	NM_103220 // AT1G35220 /// AT1G35220.1 // AT1G35220
12	13377908	AT1G50490	NM_103932 // UBC20 /// AT1G50490.1 // UBC20 /// AK227382 // UBC20
13	13351724	AT1G53570	NM_001036104 // MAP3KA /// NM_104235 // MAP3KA /// NM_179472 // MAP3KA /// NM_001198282 // MAP3KA /// NM_001198283 // MAP3KA /// AT1G53570.1 // MAP3KA /// AT1G53570.2 // MAP3KA /// AT1G53570.5 // MAP3KA /// AT1G53570.3 // MAP3KA /// AT1G53570.4 // MAP3KA /// AY140005 // MAP3KA
14	13379358	AT1G54100	NM_104287 // ALDH7B4 /// NM_179476 // ALDH7B4 /// AT1G54100.2 // ALDH7B4 /// AT1G54100.1 // ALDH7B4 /// AY091032 // ALDH7B4
15	13380818	AT1G58200	NM_202317 // MSL3 /// NM_104601 // MSL3 /// AT1G58200.1 // MSL3 /// AT1G58200.2 // MSL3 /// AY125504 // MSL3
16	13383170	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
17	13383173	AT1G64670	NM_105142 // BDG1 /// AT1G64670.1 // BDG1 /// BT005382 // BDG1
18	13386388	AT1G72430	NM_105902 // AT1G72430 /// AT1G72430.1 // AT1G72430
19	13386986	AT1G73700	NM_106029 // AT1G73700 /// AT1G73700.1 // AT1G73700

20	13388178	AT1G76540	NM_106304 // CDKB2;1 /// AT1G76540.1 // CDKB2;1 /// AB047279 // CDKB2;1
21	13360672	AT1G78480	NM_106495 // AT1G78480 /// NM_001198497 // AT1G78476 /// AT1G78480.1 // AT1G78480 /// AT1G78476.1 // AT1G78476
22	13361160	AT1G79590	NM_001036224 // SYP52 /// NM_106607 // SYP52 /// AT1G79590.2 // SYP52 /// AT1G79590.1 // SYP52 /// AF355756 // SYP52
23	13390427	AT2G01340	NM_126195 // At17.1 /// AT2G01340.1 // At17.1 /// BT021987 // At17.1
24	13411231	AT2G21140	NM_127684 // PRP2 /// AT2G21140.1 // PRP2 /// AK117333 // PRP2
25	13412416	AT2G24550	NM_128016 // AT2G24550 /// AT2G24550.1 // AT2G24550 /// AF411794 // AT2G24550
26	13413792	AT2G27450	NM_179770 // NLP1 /// NM_128305 // NLP1 /// AT2G27450.2 // NLP1 /// AT2G27450.1 // NLP1 /// AY072113 // NLP1
27	13401267	AT2G35800	NM_179921 // AT2G35800 /// AT2G35800.1 // AT2G35800 /// AY063948 // AT2G35800
28	13401305	AT2G35860	NM_179922 // FLA16 /// AT2G35860.1 // FLA16 /// AY093189 // FLA16
29	13405494	AT2G45790	NM_130142 // PMM /// AT2G45790.1 // PMM /// AY050806 // PMM
30	13405651	AT2G46140	NM_130176 // AT2G46140 /// AT2G46140.1 // AT2G46140 /// BT004206 // AT2G46140
31	13406073	AT2G47170	NM_130285 // ARF1A1C /// AT2G47170.1 // ARF1A1C /// M95166 // ARF1A1C
32	13406530	AT2G47980	NM_130365 // SCC3 /// AT2G47980.1 // SCC3 /// AY063915 // SCC3

For both these accessions, the results from the feeding behaviour study (chapter 3) were also found to be different from field assessment. The *B.villosa* accession was found to be resistance while *B.fruticulosa* was categories as moderately resistant in EPG study (chapter 3). The mismatch in the case of *B.fruticulosa* may be the result of differences in the genetic bases of model where all genotypes used to develop model had genome C whereas the test data had genome B.

Even though 80% of the unknown samples were predicted correct, it is still not the exact match to the field experiment. This might be due to the fact that sample size used to develop the model is insufficient and unbalanced with 7 resistant and 3 susceptible genotypes. Moreover, it was also felt that field classification from chapter 2 may not be the best way for grouping the samples as resistant or susceptible as these genotypes still need further validation for their resistance status and are genuinely intermediate in their resistance status. Even though there is a strong separation of the resistant and susceptible group which is a proof of good model building, the best option to develop a prediction model would have been to use genotypes which are well characterised and can be used as the benchmark for the resistance and susceptibility to aphid feeding in Brassicas. The results for all 5 models are shown in table 6.7.

6.5 Discussion

The results from this study show that gene expression data can be successfully used in the prediction of resistance or susceptibility status in unknown genotypes present in the germplasm. The advances in the molecular technologies have enabled the use of available data in better and more advanced ways. Whole genome microarray expression profiling for prediction of disease classes in humans is a widely accepted approach (K.

Kim et al. 2014; Nagata et al. 2014; Cooper-Knock et al. 2012; Swanson et al. 2012)but not used in plants as yet to screen for resistance and susceptibility to insect pests in crop plants.

Table 6.7 : Details of 5 best Model run with the classifiers for prediction resistance status of different Brassica species

S.no	Species	Seed no.	Accession no/ cultivar name	Result from prediction	Expected result	Match
1	*B.incana*	24	BRA 2856	Resistant	Resistant	Yes
		26	K 10373	Resistant	Resistant	Yes
2	*B.montana*	37	K 9404	Susceptible	Susceptible	Yes
		38	BRA 1644	Resistant	Resistant	Yes
3	*B.villosa*	398	K 6926	Resistant	Susceptible	No
		401	K 10259	Resistant	Resistant	Yes
4	*B.fruticulosa*	453	Bol2009-0080	Susceptible	Resistant	No
		454	Bol2009-0081	Susceptible	Susceptible	Yes
5	*B.oleracea acephala*	54	BRS-0103	Resistant	Resistant	Yes
		116	CGN18468	Resistant	Resistant	Yes

There are several studies where this approach is used to classify the disease state of a human patient, especially in the case of cancer (Mazzanti et al. 2004; Trevino et al. 2011). Cheng et al. (2010) reported the successful use of expression data to identify expression biomarkers associated with the patient's clinical response and to forecast various human disease outcomes. Ford-Lloyd et al. (2011) have urged for a need of new genetic recourses in plant breeding to increase the rate of gene discovery within available germplasm. It is very important that the available germplasm should be screened quickly and efficiently so as to come up with new improved varieties to keep pace with increasing population and food demand. The traditional phenotyping methods are time-consuming and also require more resources whereas the use of new technologies can provide better solutions to these problems. The need is to develop

more user-friendly and reliable methods for potential gene prediction and subsequent models for prediction of resistance status of the plant.

The use of 'omics' technology to develop such models in the case of human disease is advancing very fast. Amar et al. (2015) reported an integrated analysis of numerous heterogeneous gene expression profiles for detecting robust disease-specific biomarkers and proposing drug targets. The use of latest technologies like RNA-seq and its comparison to microarray-based models is currently applied for clinical endpoint prediction (Xu et al. 2016; De la Blétière et al. 2012; Huang et al. 2015; W. Zhang et al. 2015). The plant science is somewhat lacking in coming up tools and methods which can speed up the characterization of vast germplasm recourses.

The current study was an attempt to address this issue, and prove that such methods can be developed and successfully used in case of plants as well.

The model selection tool of Partek Genomics Suite version 6.6 (Partek Inc., St. Louis, MO, USA), software is one such way which can be used in an effective way to classify the material into groups based on a test model designed from the known species to predict the unknown species. The idea behind this approach (i.e., creating the model with the differentially expressed significant genes in resistant vs. susceptible samples and then validating the unknown samples against this model) was that this model can be used by other researchers to directly predict the resistance status of unknown genotypes by validating the transcriptomic response of unknown samples.This tool has been successfully used in studying transcriptomics profile and development of prediction models for human disease condition (Ding et al. 2013; Swanson et al. 2012). Although, in the current study the

model developed could not yield the 100% results as expected and could not be completed validated due to the restricted usage and time constrains of using a trial version of software. The fully functional version of software provides an opportunity to cross-validate model developed by using PCA analysis method, which can provide a more satisfactory and confirmatory result. In addition to this there were other limitations noticed in this model as the small and unbalanced sample size. This issue can be easily addressed in any further study, but was not addressed in current study due to time and financial constrains. The success of any such model depends on the data set used to develop the classification model. So the gene expression data from confirmed, bench-marked cabbage aphid resistant and susceptible Brassica genotypes will be ideal for developing the model in future. The present study is the first to use this approach to develop and test the model based on the gene expression (in the absence of aphid feeding) differences of resistant and susceptible groups with respect to aphid feeding preference in the field. The model selection tool was successfully used to develop the prediction model and test the genotypes grouped as unknown. The use of Partek Genomics Suite version 6.6 has made it very easy to choose the variable and design the model for prediction and number of models can be generated and tested within a short duration of time. Also, as this study is the part of bigger PGR consortium, the use of metabolomic and proteomic data can be further incorporated into this model which can make it a very robust model for prediction of resistance status. As mentioned earlier also, it is again emphasized that the class prediction model used in this study is just a preliminary study and requires more validation before being used on a large scale by future researchers. It opens up avenues for future researchers to explore the use of microarray data to develop prediction models. The latest development in 'omics' field

with NGS, proteomics and metabolomics a huge amount of data is produced and analysed for better understanding but still the large amount of microarray data stored in public databases can be still used and applied in new ways to develop new tools to take plant research on a further advanced stage.

6.6 Conclusion

To conclude, this study indicates that the variation in the gene expression of resistant and susceptible genotypes can be explored and used as the basis for developing a prediction model to identify pest resistant plant genetic material for plant breeding programmes. The 5 models developed in this study using gene expression data from resistant and susceptible genotypes were successful in grouping the test genotypes into resistant or susceptible. However, still the prediction did not match exactly with the field results which were considered as the base of characterizing resistant and susceptible genotypes. Among many reason for this imperfect prediction most important may be the small, unbalanced sample size, use of genotypes from different genome types. The inclusion of more samples of the similar subspecies group to build the model may improve the prediction of insect resistance status of the Brassica genotypes using the transcriptomics data which then may serve as a genomic database tool to screen the Brassica genotypes. The use of advanced molecular, genomic and transcriptomics techniques along with the statistical bioinformatics could provide a better understanding and prediction capability to detect various biotic and abiotic stress responses in plants also.

Table S6.1: List of 100 selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of B.incana species using the Partek model selection tool.

Table S6.2: List of 100 selected genes used to build the prediction model o test suitability to aphid feeding in two genotypes of B.montana species using the Partek model selection tool.

Table S6.3: List of 100 selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of B.villosa species using the Partek model selection tool.

Table S6.4: List of 100 selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of B.fruticulosa species using the Partek model selection tool.

Table S6.5: List of 100 selected genes used to build the prediction model to test suitability to aphid feeding in two genotypes of B.oleracea acephala species using the Partek model selection tool.

CHAPTER 7

General discussion

7.1 Introduction

To accommodate the increased demand for food, a 70% increase in food production is required over the next four decades to feed an ever-increasing population (FAO 2009). Under the threat of food shortages and global climatic changes, it is crucial for us to understand plant defence mechanisms against different biotic and abiotic stresses faced by crop plants (Mba et al. 2012; Barah et al. 2013). The advances in plant genetics remain a key component of global food security and have always provided new knowledge and technologies needed to address these challenges (Ronald 2011). Insect pests often result in a huge loss of crop yield and quality (Brown et al. 1999; Schoonhoven et al. 2005). Aphids are most serious insect pests of Brassica crops and reported to cause 30% to 80% crop yield loss annually worldwide (Dedryver et al. 2010; Razaq et al. 2011; Isik & Görür 2009). Cabbage Aphid (*B.brassicae*) is a specialist, Brassica feeding insect which causes severe leaf fouling due to its tendencies to colonise, transmit plant pathogenic viruses like cauliflower and turnip mosaic viruses and causing huge crop losses (Bush et al. 2006; Harvey et al. 2007; García-Palacios et al. 2013; Martinière et al. 2009). Currently, insecticides and pesticides are used to control this insect pest. However, due to number of growing concerns like insect resistance to these chemicals, pesticide residues in food and ill effects on humans and environment, decline in number of effective insecticides, as a result of non-availability due to more stringent health and safety criteria as part of European pesticide legislation (Directive EC1107/09), alternative methods of pest control are highly needed. Plant phenomics is a somewhat new term (Finkel 2009; Schilling et al. 1999) used to define the phenotypic features of a plant in relation to its functions (Saint Pierre 2012). The plant breeders are constantly looking for traits within the crop species to develop new

insect resistant varieties of crops. Therefore, identification of insect resistant genotypes present in the germplasm is often considered as the starting point of the breeding process. High throughput phenotyping along with phenomics tools helps to determine useful traits like insect resistance which can be used for breeding new varieties. For example, to study variation in levels of antixenosis and antibiosis in hundreds of ecotypes of Arabidopsis, use of automated video tracking of the green peach aphid on leaf discs of Arabidopsis thaliana is reported by Kloth et al. (2012). The development of high-throughput SNP discovery in common beans, cassava using next generation sequencing is already reported by many researchers (Hyten et al. 2010; Varshney et al. 2009; Varshney et al. 2006; Ferguson et al. 2012). Pawełkowicz et al. (2016) reported the use of "omics" techniques in research conducted in cucumber plant breeding programmes on the ripening process, phloem transport, disease resistance, cold tolerance and fruit quality traits.

In this study, a combination of phenomics and transcriptomics tools was used to determine insect resistant plant material and candidate genes for insect resistance. The field assessment (Chapter 2), feeding behaviour assessment (Chapter 3) and gene expression analysis (Chapter 4 & 5) of Brassica accessions (both CWR and LR) were conducted to identify the insect resistant (non-preferred) material in response to aphid feeding. The integrated approach of using field assessment (Chapter 2) of 200 Brassica accessions to identify the cabbage aphid feeding preference on these accessions and then using 15 accessions to further study feeding behaviour using EPG technique (Chapter 3) and gene expression changes (chapter 4 & 5) in response to presence and absence of aphid feeding in resistant (preferred) and susceptible (non-preferred) accessions has generated very useful information in terms of presences of variable

levels of resistance to cabbage aphid feeding and, also revealed candidate genes which can form the basis for further research investigations.

7.2 Screening of Brassica Germplasm Based on Field Assessment to Identify Insect Resistance

In this study, 200 Brassica accessions (CWR and LR) with no or very limited information regarding the level of resistance against cabbage aphid were tested in the field (Chapter 2). The field based germplasm phenotyping is considered more realistic as plants are present in their natural habitat (Pierre 2012). This study provided vital information about differences in colonisation levels of cabbage aphid on Brassica genotypes used which may be due to host-specific resistance in Brassica germplasm. The presence of resistance against aphid resistance in CWR is already reported in the literature (Singh et al. 1994; Ellis et al. 1996; Ellis et al. 1998). In this study, the accessions from B.incana, B.fruticulosa, and B.villosa were noted to have low the level of aphid infestation as compared to LR (B.oleracea) species. The CWR belonging to B.incana, B.fruticulosa and B.villosa are reported to possess resistance to cabbage aphid (Ellis et al. 2000; Nayidu et al. 2014; Nayidu et al. 2015). Singh et al. (1994) reported the presence of resistance to cabbage aphid through antibiosis in wild species, B.fruticulosa, B.insularis and B.villosa. Antixenosis, antibiosis and tolerance are three categories to define the host plant resistance against insect pests. Antixenosis or non-preference arises from host plant's own characters which make them unattractive to insect pests. Antibiosis is a type of resistance in which host plant causes injury; reduce reproduction or death of pest. It is often noticed that both resistant and susceptible plants respond in a similar way to insect pest but resistant varieties response is more quick and dramatic. Tolerance is the third category to define resistance, where host

plant develops a capacity to grow and yield well despite damage caused by insect pests. The tolerant plants are often described to have better wound healing and disease fighting capacity (Leon 2001). The resistance mechanism in the selected accessions was further explored by way of studying feeding behaviour by means of EPG (Chapter 3). Although, field screening is highly used technique to characterise resistant and susceptible accessions, but in our study it not used to its maximum potential as it was used to narrow down the number of accessions and selection Brassica accessions to be investigated for the feeding behaviour assessment (Chapter 3) and gene expression studies (Chapter 4) as it was not possible to test as many as 200 accessions for subsequent experiments due to financial and time constrains.

7.3 Intraspecific and interspecific variation in feeding performance of cabbage aphid (*B.brassicae*) in Brassicas

The EPG technique is a widely used method to study the detailed feeding behaviour of phloem feeding insects and localisation of resistance factors in plant tissue (Tjallingii, W. F. and Mayoral 1992). EPG studies have produced reliable information concerning the location of the resistance mechanisms (Verdugo et al. 2012). For example, the resistance mechanism against M.persicae in the "Malo Konare" peach cultivar was located in the vascular system and was found to provide antibiosis resistance (Monet et al. 1998; Sauge et al. 1998; Sauge 1998a). Broekgaarden et al. (2012) reported antibiosis based phloem specific resistance in B.oleracea cultivars against the whitefly. EPG studies are the highly preferred way of studying feeding behaviour of insect pests to compare interspecies and intraspecies variations as well (Ghaffar et al. 2011; Gabryś & Pawluk 1999; Soffan & Aldawood 2015). The electronic monitoring of feeding behaviour suggests that cabbage aphid is capable of feeding on a wide range of Brassica

species. The results from this study (chapter 3) revealed the presence of variable levels of feeding suitability among Brassica accessions as a host. For the EPG results, varying resistance or susceptibility statuses were defined according to the plant factor involved (Table 3.2). Because EPG parameters can be useful in identifying the tissues containing putative resistance factors (van Helden & Tjallingii 1993; Tjallingii 1995), a detailed analysis of the EPG results may help elucidate the location of the resistance mechanisms. The results from our study indicated the presence of antibiosis component of resistance mechanism rather than antixenosis, as cabbage aphid's presence was noted on all the accessions tested in this study. The analysis of waveforms Non-Penetration, Pathway, xylem, derailed stylet, E2 and time to E2 were used to classify three tier responses to aphid feeding by plants. The plants offer resistance to phloem feeding insects at surface level, pathway level and in phloem. EPG analysis of non-penetration pathway (Figure 3.3a) indicated the presence of surface resistance in two accessions BRA2856 (B.incana) and K6926 (B.villosa), as aphids could not penetrate the leaf tissue and spent all 6 h in non-penetration phase. This strongly suggests that some sort of resistance factor is present on the surface of these two accessions. The presence of resistance due to the high density of trichomes is reported in B.villosa (Nayidu et al. 2014; Nayidu et al. 2015). It is particularly interesting to note that this accession was found susceptible in field assessment (chapter 2), suggesting that despite the initial inability to feed (for first 6h), aphids are likely able to develop induced susceptibility under natural environment, as shown in other aphid-plant systems (Karban & Baldwin 1997; Prado & Tjallingii 1997). The analysis of time to E2 and E2 parameters (Table 3.3) showed that B.oleracea accession (BRA 2401) had more sustained phloem feeding as compared B.incana (K10373). It was also noted that difference in response to aphid

feeding, exists between accessions of the same species as well as different species which makes the analysis of insect defence system even more complex. The analysis by Electrical Penetration Graph of aphid probing behaviour during penetration of stylets into the plant tissues showed that resistance was mainly linked to difficulties either to reach the phloem or to initiate sap uptake or to sustain food ingestion. This study clearly demonstrates the existence of different mechanisms of resistance among the genotypes tested, and therefore their individual importance in a Brassica breeding programme for resistance.

7.4 Gene expression study: An approach to identify genes of interest

Microarrays are the most extensively used transcriptomics tool to investigate the gene expression changes of plants in response to biotic or abiotic stresses (Duggan et al. 1999; Lee et al. 2008; Meyers et al. 2004; Couldridge et al. 2007). Analysis of the expression of thousands of genes in a single experiment has made the microarray unique and highly preferred technology for studying transcriptional changes in plants in response to insect feeding. This study focused on the comparison of gene expression changes of insect resistant (non-preferred) and susceptible (preferred) accessions in response to presence or absence of aphid feeding. The selection of both induced (presence of aphid feeding) and non-induced (absence of aphid feeding) conditions to compare resistant vs. susceptible accessions provided the opportunity to look at both constitutive gene expression and induced transcriptional changes in response to aphid feeding. Constitutive gene expressions are often considered more durable, as they are always present and do not change easily (Karban & Baldwin 1997).

This study adopted a unique and novel approach to evaluate gene expression data from induced (presence of aphid feeding) and non-induced (absence of aphid feeding) experiments based on different group classifications (chapter 4). The whole data was analysed four times, based on resistant and susceptible grouping of Brassica accessions according to field performance (chapter 2) and feeding behaviour assessment (chapter 3). This approach provided valuable results in terms of localising the gene expression changes with respect to surface, pathway or phloem based resistance in addition to overall changes based on the field experiment.

It was observed that in field based comparison of resistant and susceptible accessions, more genes were regulated constitutively. Fifty-four significant gene probes were found expressing differentially where 43 were down-regulated and 11 up-regulated. In contrast, aphid feeding for 24h induced significant changes in gene expression of only 7 gene probes (4 up-regulated and 3 down-regulated). Overall a very small number of gene probes (section 4.4.1) showed changes in gene expression between resistant and susceptible accessions in response to presence or absence of aphid feeding. The occurrence of a higher number of constitutively expressed genes in comparison to only 7 significant differentially expressed genes in response to aphid feeding could most likely be due to the taxonomic differences as genotypes belonging to different species were used in the study and may have nothing to do with the resistance against aphids. Moreover, another factor for getting a low number of significant genes may be due to the use of Arabidopsis array used in this study, where we may be losing on Brassica specific genes which is one drawback of using an interspecies array.

The results from EPG based classifications (section 4.4.2) showed some major changes in gene expression patterns of gene probes which may be linked to the location of

resistance factors at surface, pathway or phloem. This study clearly indicated that different set of gene probes in response to aphid feeding were induced or repressed in all three comparisons based on EPG assessment. The significant gene probes like NPC6 and PDCB3 were found expressing in Non-penetration based assessment and can be linked directly to surface resistance whereas PP2-A10 – At1g10155; At3g56240; At5g09650 were found significant in their differential gene expression in phloem based classification. The Gene Ontology (GO) enrichment analysis showed that the gene probes having significant transcriptional changes mainly belong to GO terms response to biotic stress, response to insect, defence response, signalling, transport activity, cell wall, transcription factors, protein kinases and hormone induced pathways. In addition to these, a large set of gene probes encoding hypothetical proteins with unassigned functions were found. The proteins of unidentified functions have been identified in other similar studies as well (Kuśnierczyk et al. 2008; Broekgaarden et al. 2007). The up-regulation of these hypothetical proteins in response to aphid feeding indicates towards their possible role in defence mechanisms and need further investigation. Also due to lack of available annotation, a large set of significant gene probes could not be analysed which may be playing important roles in defence against insects

Comparative transcriptional profiling is reported to be successfully used to profile contrasting genotypes in the aphid-plant system previously also (Walia et al. 2005; Voelckel & Baldwin 2004; Moran & Thompson 2001; Moran et al. 2002; Thompson & Goggin 2006). The variation in gene expression between induced and non-induced plants were found in this study. This is also reported in other studies where plant response to sap feeding insects is investigated (De Vos et al. 2005; Broekgaarden et al. 2008; Pelgrom et al. 2014). The insect resistance in B.villosa due to the presence of

trichomes is studied using transcriptional response (Nayidu et al. 2015). The reports about involvement of genes enriched under GO categories like cell wall modifications, response to stress, response to jasmonic acid and ethylene are always linked to play role in resistance against herbivores (Broekgaarden, Snoeren, et al. 2011; Delp et al. 2009; Appel et al. 2012; Mewis et al. 2006; Reymond 2000).The gene probe AT3G24190, AT1g01141and AT5G1218 belong to SNF1/AMPK/SnRK1/CPK protein kinases family, which typically respond to biotic stress in plants (Schwachtje & Baldwin 2008; Crozet et al. 2014; Polge et al. 2008). These were found to be down-regulated in our study. Appel et al. (2014) reported down-regulation of SnRk1, (member of protein kinases family) in Arabidopsis in response to aphid feeding. The result from transcriptional analysis of Brassica genotypes in this study provides strong evidence that aphid faces resistance at different levels in response to aphid feeding. The genes like NPC6 ,PDCB3 PP2-A10, IIL1, At3g56240; At5g09650, are strong candidates to be tested for their specific role against cabbage aphid resistance and testing the hypothesis using more advanced "omics" technology that they are source of surface level resistance and for phloem based resistance in plants against aphid feeding

7.5 Impact of domestication: Variation in gene expression patterns in crop wild relative (CWR) and landrace (LR) genotypes

This study (Chapter 5) explored the differences in the gene expression between resistant CWR and resistant LR of Brassica. Both CWR and LR are regarded as important resources for crop improvement and may possess resistance against insect feeding. These results showed that gene expression is highly regulated and aphid feeding induced 216 gene probes in contrast to only 75 non-induced plants. The significantly differentially expressed genes in categories like response to stress, response to stimulus,

signalling, developmental processes can be good indicators of aphid resistance or tolerance in the CWR as compared to LR and may prove useful in the development of aphid resistant Brassica varieties in future. Several photosynthesis- and/or development related genes were repressed upon aphid feeding (Table 5.2 and TableS5.6) which indicated towards the likelihood that plants have reallocated resources for defence at the expense of growth and/or photosynthesis as reported earlier (Bilgin et al. 2010; Wu & Baldwin 2010; Herms & Mattson 1992). The number of stress related genes induced in CWR in response to aphid feeding was more than in the absence of aphid feeding. The gene probes like PTR3, NAC083, TIFY7, RAP210 and EDR6 found up-regulated in CWR plants are strong candidates for resistance against cabbage aphid. The further investigation of these candidates is highly recommended in order to confirm their role and further use in the breeding new more resistant varieties. In the first instance, an immediate way forward will be to do RT-PCR to confirm their function followed by mapping these genes against known QTL's. In addition, doing the knock-in/knock-out studies using these candidates in model plant Arabidopsis can help in confirming the role of these candidate genes.

The phenotyping of all available accessions in plant genetic resources to test for aphid resistance is a nearly impossible task. It will demand a huge investment in terms of time, man-power and other resources. To overcome this problem plant researchers are working towards the development of new high-throughput techniques using "omics" tools to screen germplasm for their resistance status. It was attempted in this study (chapter 6) for the first time, to use gene expression data generated from microarray experiment (chapter 4) to develop a class prediction model based on already defined genotype classifications (resistant or susceptible) to predict genotypes of unknown

status. This approach is successfully used in disease prediction in humans especially in case of cancer (Bucca et al. 2004; Reis-Filho & Pusztai 2011; Cooper-Knock et al. 2012). Trevino et al. (2011) successfully demonstrated the use of gene expression data and statistical modelling in prediction of prostate cancer and use of this approach as a diagnostic tool. The results from this study showed that this approach can be used in plant studies also. Currently, this approach is in its infant stage and requires a lot of further validations before it can be successfully used as a prediction model. Nevertheless, the results from this study are promising.

7.6 Limitations and future perspective

Although efforts were made to design a perfect study, due to constraints in available resources and time scale the scope of this research project was limited. There were a few limitations noticed during the course of study. Generally, field trials are conducted 2 or 3 times before drawing any conclusions (as they are not easily reproducible) to account for environmental variations and other climatic changes. In this study, field assessment (chapter 2) was conducted only once (in 2011) due to unavailability of field space and available funds. This may be considered as a limitation by some, but the main aim of conducting the field trial to narrow down the number of accessions for feeding behaviour (chapter 3) and gene expression studies (Chapter 4 & 5) was achieved. There were some contradictions noticed in the results from field assessment (chapter 2) and feeding behaviour assessment (chapter 3) about the level of resistance in some accessions for example K6926 which was found susceptible in the field but highly resistant in feeding assessment which may be due to initial resistance offered by plant, but then eventually plant became susceptible to aphid feeding (Karban & Baldwin 1997; Prado & Tjallingii 1997) and accession 57071 which was found resistant in field

(chapter 2) but had sustained phloem feeding (table 3.2) in feeding behaviour assessment (chapter 3). In explanation to this, it is hypothesised that cabbage aphid induced a delayed resistance response to feeding in this particular accession. This type of response is also noticed in other aphid-plants studies also (Will & van Bel 2006; Aidemark et al. 2009; Will et al. 2007; Knoblauch 1998). It is worth noting that the lack of agreement between field assessment and EPG experiments may be because of the large difference in the time windows involved between in two experiments. It is not uncommon to notice disagreement between studies.

The gene expression study was conducted using *Arabidopsis* chips to test Brassica crops. Although use of model plant arrays to study non model plants has been successfully used previously (Broekgaarden et al. 2007; Broekgaarden et al. 2008; Lee et al. 2004), it may also lead to loss of species specific information. The presence of low numbers of significant differentially expressing genes in some comparisons in (chapter 4) in this study may be result of this cross-species array use, where we may have lost the Brassica specific genes. In addition the candidate gene selection from microarray data is based on mRNA abundance but do not account for post translational modifications. So the further proteomic and/or metabolomic analysis of these candidate genes can be very useful to confirm the stress response role of these genes. The use of qRT-PCR in validation of gene expression data is the most accepted method. Microarray data only suggest the candidate genes which are helpful in creating hypothesis around the putative function of that gene. The function of these candidates should be confirmed by using knockouts, T-DNA insertion lines, tilling mutants and using transgenic approaches. The use of sequence based techniques like RNA-seq can be used to get more precise estimation of gene expression.

As this study was part of a bigger PGR secure project, the results from this study i.e. feeding behaviour assessment and candidate genes, along with the results from RNA-seq analysis of B.oleracea and its relatives conducted by another partner, resulted in development of publicly available, the 90k Affymetrix Axiom array containing 40,000 SNPs selected from a set of broccoli varieties, 21,000 polymorphic SNPs from a set of heading cabbages, 4200 already validated B. oleracea SNPs and 5000 SNPs that are polymorphic between B.oleracea and the wild relative B.incana, as well as 5000 that are polymorphic between B.oleracea and B.montana. The array also contains 5000 SNPs that are polymorphic within B.fruticulosa. This array will be very useful in a number of applications including QTL mapping in B. oleracea and CWR, association mapping in B.oleracea, as well as relationship analysis among species, subspecies, varieties and landraces.

To conclude, the recent advances in 'omics' technology and development of bioinformatics tools is providing more opportunities and opening avenue of system biology to a better understating of underlying biological and molecular processes involved in insect resistance in plants. Identification of genes networks playing role in resistance mechanisms by sensing the stress, activating the response to it, will be very helpful in developing an effective breeding programme. The findings from this study will contribute to ongoing and future research in investigating cabbage aphid resistance in Brassica for a broad spectrum and durable insect resistant cultivars. The development of crop plants which can cope with biotic and abiotic stresses will significantly contribute in achieving global food security.

CHAPTER 8

References

Agarwal, R. and Narayan, J. (2015). Unraveling the impact of bioinformatics and omics in Agriculture. *Int Journal of Plant Biology and Reasearch*, 3(2), p.1039.

Ahmad, M. and Akhtar, S. (2013). Development of insecticide resistance in field populations of Brevicoryne brassicae (Hemiptera: Aphididae) in Pakistan. *Journal of economic entomology*, *106*(2), pp.954-958.

Ahmad, M., Arif, M.I. and Ahmad, M. (2007). Occurrence of insecticide resistance in field populations of Spodoptera litura (Lepidoptera: Noctuidae) in Pakistan. *Crop Protection*, *26*(6), pp.809-817.

Ahuja, I., de Vos, R.C., Bones, A.M. and Hall, R.D. (2010). Plant molecular stress responses face climate change. *Trends in plant science*, *15*(12), pp.664-674.

Ahuja, I., Rohloff, J. and Bones, A.M. (2009). Defence mechanisms of Brassicaceae: implications for plant-insect interactions and potential for integrated pest management. A review. *Agronomy for Sustainable Development*, *30*(2), pp.311-348.

Aidemark, M., Andersson, C.J., Rasmusson, A.G. and Widell, S. (2009). Regulation of callose synthase activity in situ in alamethicin-permeabilized Arabidopsis and tobacco suspension cells. *BMC plant biology*, *9*(1), p.27.

Ali, J.G. and Agrawal, A.A. (2012). Specialist versus generalist insect herbivores and plant defense. *Trends in plant science*, *17*(5), pp.293-302.

Al-Shehbaz, I.A. (2012). A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon*, *61*(5), pp.931-954.

Alvarez, A.E., Tjallingii, W.F., Garzo, E., Vleeshouwers, V., Dicke, M. and Vosman, B. (2006). Location of resistance factors in the leaves of potato and wild tuber-bearing Solanum species to the aphid Myzus persicae.*Entomologia Experimentalis et Applicata*, *121*(2), pp.145-157.

Amar, D., Hait, T., Izraeli, S. and Shamir, R. (2015). Integrated analysis of numerous heterogeneous gene expression profiles for detecting robust disease-specific biomarkers and proposing drug targets. *Nucleic acids research*, *43*(16), pp.7779-7789.

Appel, H.M., Fescemyer, H., Ehlting, J., Weston, D., Rehrig, E., Joshi, T. et.al. (2014). Transcriptional responses of Arabidopsis thaliana to chewing and sucking insect herbivores. *Frontiers in plant science*, *5*, p.565.

Appel, H.M., Arnold, T.M. and Schultz, J.C. (2012). Effects of jasmonic acid, branching and girdling on carbon and nitrogen transport in poplar. *New Phytologist*, *195*(2), pp.419-426.

Ashkani, S., Rafii, M.Y., Shabanimofrad, M., Miah, G., Sahebi, M., Azizi, P. et.al. (2015). Molecular Breeding Strategy and Challenges Towards Improvement of Blast Disease Resistance in Rice Crop. *Frontiers in plant science*, *6*.p886.

Augenlicht, L.H. and Kobrin, D. (1982). Cloning and screening of sequences expressed

in a mouse colon tumor. *Cancer research*, *42*(3), pp.1088-1093.

Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S. et.al. (2005). Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis. *Genome research*, *15*(4), pp.487-495.

Backus, E.A., Rangasamy, M., Stamm, M., Mcauslane, H.J. and Cherry, R. (2013). Waveform library for chinch bugs (Hemiptera: Heteroptera: Blissidae): characterization of electrical penetration graph waveforms at multiple input impedances. *Annals of the Entomological Society of America*, *106*(4), pp.524-539.

Bailey-Serres, J., Fukao, T., Ronald, P., Ismail, A., Heuer, S. and Mackill, D. (2010). Submergence tolerant rice: SUB1's journey from landrace to modern cultivar. *Rice*, *3*(2-3), pp.138-147.

Barah, P. and Bones, A.M. (2015). Multidimensional approaches for studying plant defence against insects: from ecology to omics and synthetic biology.*Journal of Experimental Botany*, *66*(2), pp.479-493.

Barah, P., Winge, P., Kusnierczyk, A., Tran, D.H. and Bones, A.M. (2013). Molecular signatures in Arabidopsis thaliana in response to insect attack and bacterial infection. *PLoS One*, *8*(3), p.e58987.

Barh, D., Khan, M.S. and Davies, E. (2015). *PlantOmics: The Omics of Plant Science*. Springer.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M. et.al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic acids research*, *41*(D1), pp.D991-D995.

Barthel, A., Kopka, I., Vogel, H., Zipfel, P., Heckel, D.G. and Groot, A.T. (2014). Immune defence strategies of generalist and specialist insect herbivores. *Proceedings of the Royal Society of London B: Biological Sciences*, *281*(1788), p.20140897.

Becher, M., Talke, I.N., Krall, L. and Krämer, U. (2004). Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator Arabidopsis halleri. *The Plant Journal*, *37*(2), pp.251-268.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N.J. and Nekrasov, V. (2015). Editing plant genomes with CRISPR/Cas9. *Current opinion in biotechnology*, *32*, pp.76-84.

Beneteau, J., Renard, D., Marché, L., Douville, E., Lavenant, L., Rahbé, Y. et.al. (2010). Binding properties of the N-acetylglucosamine and high-mannose N-glycan PP2-A1 phloem lectin in Arabidopsis. *Plant physiology*, *153*(3), pp.1345-1361.

Bennett, R.N. and Wallsgrove, R.M. (1994). Secondary metabolites in plant defence mechanisms. *New Phytologist*, *127*(4), pp.617-633.

Benrey, B., Callejas, A., Rios, L., Oyama, K. and Denno, R.F. (1998). The Effects of Domestication ofBrassicaandPhaseoluson the Interaction between Phytophagous Insects and Parasitoids. *Biological Control*, *11*(2), pp.130-140.

Bhatia, V., Uniyal, P.L. and Bhattacharya, R. (2011). Aphid resistance in Brassica crops: challenges, biotechnological progress and emerging possibilities. *Biotechnology advances*, *29*(6), pp.879-888.

Bilgin, D.D., Zavala, J.A., Zhu, J.I.N., Clough, S.J., Ort, D.R. and DeLUCIA, E. (2010). Biotic stress globally downregulates photosynthesis genes.*Plant, cell & environment*, *33*(10), pp.1597-1613.

Blackman, R. and Eastop, V.F. (2000). *Aphids on the World's Crops: An Identification and Information Guide, 2nd E*. John Wiley & Sons.

Bolouri, M.M. and Van den Ende, W. (2012). Sugars and plant innate immunity. *Journal of experimental botany*, *63*(11), pp.3989-3998.

Bones, A.M. and Rossiter, J.T. (2006). The enzymic and chemically induced decomposition of glucosinolates. *Phytochemistry*, *67*(11), pp.1053-1067.

Bortesi, L. and Fischer, R. (2014). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology advances*, *33*(1), pp.41-52.

Brar, D.S., Khush, G.S. (2002). Transferring genes from wild species into rice. In Quantitative Genetics, Genomics and Plant Breeding. Kang, M.S. (ed.). Wallingford, UK: CAB International, pp. 197–217

Brar, D.S. and Khush, G.S. (1997). Alien introgression in rice. *Plant molecular biology*, *35*(1-2), pp.35-47.

Brien, C.J., Berger, B., Rabie, H. and Tester, M. (2013). Accounting for variation in designing greenhouse experiments with special reference to greenhouses containing plants on conveyor systems. *Plant methods*, *9*(1), p.1-22.

Broekgaarden, C., Riviere, P., Steenhuis, G., del sol Cuenca, M., Kos, M. and Vosman, B. (2012). Phloem-specific resistance in Brassica oleracea against the whitefly Aleyrodes proletella. *Entomologia Experimentalis et Applicata*, *142*(2), pp.153-164.

Broekgaarden, C., Voorrips, R.E., Dicke, M. and Vosman, B. (2011). Transcriptional responses of Brassica nigra to feeding by specialist insects of different feeding guilds. *Insect Science*, *18*(3), pp.259-272.

Broekgaarden, C., Poelman, E.H., Steenhuis, G., Voorrips, R.E., Dicke, M. and Vosman, B. (2008). Responses of Brassica oleracea cultivars to infestation by the aphid Brevicoryne brassicae: an ecological and molecular approach. *Plant, cell & environment*, *31*(11), pp.1592-1605.

Broekgaarden, C., Poelman, E.H., Steenhuis, G., Voorrips, R.E., Dicke, M. and Vosman, B. (2007). Genotypic variation in genome-wide transcription profiles induced by insect feeding: Brassica oleracea–Pieris rapae interactions. *BMC*

Genomics, *8*, p.239.

Brown, A.F., Yousef, G.G., Chebrolu, K.K., Byrd, R.W., Everhart, K.W., Thomas, A. et.al. (2014). High-density single nucleotide polymorphism (SNP) array mapping in Brassica oleracea: identification of QTL associated with carotenoid variation in broccoli florets. *Theoretical and Applied Genetics*, *127*(9), pp.2051-2064.

Brown, J., McCaffrey, J.P., Harmon, B.L., Davis, J.B., Brown, A.P. and Erickson, D.A., (1999). Effect of late season insect infestation on yield, yield components and oil quality of Brassica napus, B. rapa, B. juncea and Sinapis alba in the Pacific Northwest region of the United States. *The Journal of Agricultural Science*, *132*(03), pp.281-288.

Bruce, T.J. (2015). Interplay between insects and plants: dynamic and complex interactions that have coevolved over millions of years but act in milliseconds. *Journal of Experimental Botany*, *66*(2), pp.455-465.

Bucca, G., Carruba, G., Saetta, A., Muti, P., Castagnetta, L. and Smith, C.P. (2004). Gene expression profiling of human cancers. *Annals of the New York Academy of Sciences*, *1028(1)*, pp.28-37.

Burger, J.C., Chapman, M.A. and Burke, J.M. (2008). Molecular insights into the evolution of crop plants. *American Journal of Botany*, *95*(2), pp.113-122.

Burton, G.W. and Wilson, J.P. (1995). Registration of Tift 65 parental inbred line of pearl millet. *Crop Science*, *35*(4), pp.1244-1244.

Bush, J., Jander, G. and Ausubel, F.M. (2006). Prevention and control of pests and diseases. *Methods in molecular biology (Clifton, N.J.)*, 323, pp.13–25.

Cabrera-Bosquet, L., Crossa, J., von Zitzewitz, J., Serret, M.D. and Luis Araus, J. (2012). High-throughput Phenotyping and Genomic Selection: The Frontiers of Crop Breeding ConvergeF. *Journal of integrative plant biology*,*54*(5), pp.312-320.

Calatayud, P.A., Rahbé, Y., Tjallingii, W.F., Tertuliano, M. and Rü, B. (1994). Electrically recorded feeding behaviour of cassava mealybug on host and non-host plants. *Entomologia experimentalis et Applicata*, *72*(3), pp.219-232.

Cardone, M., Mazzoncini, M., Menini, S., Rocco, V., Senatore, A., Seggiani, M. et.al. (2003). Brassica carinata as an alternative oil crop for the production of biodiesel in Italy: agronomic evaluation, fuel production by transesterification and characterization. *Biomass and Bioenergy*, *25*(6), pp.623-636.

Carlsson, J., Lagercrantz, U., Sundström, J., Teixeira, R., Wellmer, F., Meyerowitz, E.M. et. al. (2007). Microarray analysis reveals altered expression of a large number of nuclear genes in developing cytoplasmic male sterile Brassica napus flowers. *The Plant Journal*, *49*(3), pp.452-462.

Carr, D.E. and Eubanks, M.D. (2002). Inbreeding alters resistance to insect herbivory and host plant quality in Mimulus guttatus (Scrophulariaceae).*Evolution*, *56*(1), pp.22-30.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A., Tang, H., Wang, X. et.al. (2014). Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome.*Science*, *345*(6199), pp.950-953.

Chaudhary, B. (2013). Plant Domestication and Resistance to Herbivory.*International Journal of Plant Genomics*, *2013*. p.572784.

Chen, X., Zhang, Z., Visser, R.G., Broekgaarden, C. and Vosman, B. (2013). Overexpression of IRM1 enhances resistance to aphids in Arabidopsis thaliana. *PloS one*, *8*(8), p.e70914.

Chen, X., Vosman, B., Visser, R.G., van der Vlugt, R.A. and Broekgaarden, C. (2012). High throughput phenotyping for aphid resistance in large plant collections. *Plant Methods*, *8*(1), p.33.

Cheng, F., Wu, J. and Wang, X. (2014). Genome triplication drove the diversification of Brassica plants. *Horticulture research*, *1*, p.14024.

Cheng, F., Liu, S., Wu, J., Fang, L., Sun, S., Liu, B. et.al. (2011). BRAD, the genetics and genomics database for Brassica plants. *BMC plant biology*, *11*(1), p.136.

Cheng, F., Cho, S.H. and Lee, J.K. (2010). Multi-gene expression-based statistical approaches to predicting patients' clinical outcomes and responses. *Statistical Methods in Molecular Biology*, pp.471-484.

Chetelat, R.T., DeVerna, J.W. and Bennett, A.B. (1995). Introgression into tomato (Lycopersicon esculentum) of the L. chmielewskii sucrose accumulator gene (sucr) controlling fruit sugar composition. *Theoretical and applied genetics*, *91*(2), pp.327-333.

Cole, R.A. (1997a). Comparison of feeding behaviour of two Brassica pests Brevicoryne brassicae and Myzus persicae on wild and cultivated Brassica species. *Entomologia Experimentalis et Applicata*, *85*(2), pp.135-143.

Cole, R.A. (1997b). The relative importance of glucosinolates and amino acids to the development of two aphid pests Brevicoryne brassicae and Myzus persicae on wild and cultivated brassica species. *Entomologia Experimentalis et Applicata*, *85*(2), pp.121-133.

Cole, R.A. (1994). Locating a resistance mechanism to the cabbage aphid in two wild Brassicas. *Entomologia Experimentalis et Applicata*, *71*(1), pp.23-31.

Collins, R.M., Afzal, M., Ward, D.A., Prescott, M.C., Sait, S.M., Rees, H.H. et.al. (2010). Differential proteomic analysis of Arabidopsis thaliana genotypes exhibiting resistance or susceptibility to the insect herbivore, Plutella xylostella. *PLoS One*, *5*(4), p.e10103.

Cooper-Knock, J., Kirby, J., Ferraiuolo, L., Heath, P.R., Rattray, M. and Shaw, P.J. (2012). Gene expression profiling in human neurodegenerative disease. *Nature Reviews Neurology*, *8*(9), pp.518-530.

Cornelis, M.C., El-Sohemy, A. and Campos, H. (2007). GSTT1 genotype modifies the

association between cruciferous vegetable intake and the risk of myocardial infarction. *The American journal of clinical nutrition*, *86*(3), pp.752-758.

Couldridge, C., Newbury, H.J., Ford-Lloyd, B., Bale, J. and Pritchard, J. (2007). Exploring plant responses to aphid feeding using a full Arabidopsis microarray reveals a small number of genes with significantly altered expression. *Bulletin of entomological research*, *97*(05), pp.523-532.

Cox, G.M. and Cochran, W.G. (1946). Designs of greenhouse experiments for statistical analysis. *Soil Science*, *62*(1), pp.87-98.

Confraria, A., Elias, C.A., Rodrigues, A., Crozet, P., Adamo, M., Margalha, L. et.al. (2014). Mechanisms of regulation of SNF1/AMPK/SnRK1 protein kinases. *Frontiers in plant science*, 5, p.190.

Dedryver, C.A., Le Ralec, A. and Fabre, F. (2010). The conflicting relationships between aphids and men: a review of aphid damage and control strategies. *Comptes rendus biologies*, *333*(6), pp.539-553.

Del Carmen Martínez-Ballesta, M., Moreno, D.A. and Carvajal, M. (2012). The physiological importance of glucosinolates on plant response to abiotic stress in brassica. *International journal of molecular sciences*, *14*(6), pp.11607-11625.

De La Blétiere, D.R., Blanchet, O., Cornillet-Lefèbvre, P., Coutolleau, A., Baranger, L., Geneviève, F., Luquet, I. et.al. (2012). Routine use of microarray-based gene expression profiling to identify patients with low cytogenetic risk acute myeloid leukemia: accurate results can be obtained even with suboptimal samples. *BMC medical genomics*, *5*(1), p.1.

De Ilarduya, O.M., Xie, Q. and Kaloshian, I. (2003). Aphid-induced defense responses in Mi-1-mediated compatible and incompatible tomato interactions. *Molecular Plant-Microbe Interactions*, *16*(8), pp.699-708.

Delp, G., Gradin, T., Åhman, I. and Jonsson, L.M. (2009). Microarray analysis of the interaction between the aphid Rhopalosiphum padi and host plants reveals both differences and similarities between susceptible and partially resistant barley lines. *Molecular Genetics and Genomics*, *281*(3), pp.233-248.

De Villiers, S.M. and Hoisington, D.A. (2011). The trends and future of biotechnology crops for insect pest control. *African Journal of Biotechnology*,*10*(23), pp.4677-4681.

De Vos, M., Van Oosten, V.R., Van Poecke, R.M., Van Pelt, J.A., Pozo, M.J., Mueller, M.J. et.al. (2005). Signal signature and transcriptome changes of Arabidopsis during pathogen and insect attack. Molecular plant-microbe interactions, 18(9), pp.923-937.

De Vos, M., Kim, J.H. and Jander, G. (2007). Biochemistry and molecular biology of Arabidopsis–aphid interactions. Bioessays, 29(9), pp.871-883.

Devoto, A., Ellis, C., Magusin, A., Chang, H.S., Chilcott, C., Zhu, T. et.al. (2005).

Expression profiling reveals COI1 to be a key regulator of genes involved in wound-and methyl jasmonate-induced secondary metabolism, defence, and hormone interactions. *Plant molecular biology*,*58*(4), pp.497-513.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. *Nature*, *418*(6898), pp.700-707.

Diaz-Montano, J., Reese, J.C., Louis, J., Campbell, L.R. and Schapaugh, W.T. (2007). Feeding behavior by the soybean aphid (Hemiptera: Aphididae) on resistant and susceptible soybean genotypes. *Journal of Economic Entomology*, *100*(3), pp.984-989.

Dinant, S., Clark, A.M., Zhu, Y., Vilaine, F., Palauqui, J.C., Kusiak, C. et.al. (2003). Diversity of the superfamily of phloem lectins (phloem protein 2) in angiosperms. *Plant Physiology*, *131*(1), pp.114-128.

Dinant, S., Bonnemain, J.L., Girousse, C. and Kehr, J. (2010). Phloem sap intricacy and interplay with aphid feeding. *Comptes rendus biologies*, *333*(6), pp.504-515.

Ding, L.H., Park, S., Peyton, M., Girard, L., Xie, Y., Minna, J.D. et.al. (2013). Distinct transcriptome profiles identified in normal human bronchial epithelial cells after exposure to γ-rays and different elemental particles of high Z and energy. *BMC genomics*, *14*(1), p. 372.

Dixon, A.F.G. (1998) Aphid Ecology: An Optimization Approach. (2nd ed). Chapman & Hall, London. 300 pp

Doebley, J.F., Gaut, B.S. and Smith, B.D. (2006). The molecular genetics of crop domestication. *Cell*, *127*(7), pp.1309-1321.

Dong, X., Yi, H., Lee, J., Nou, I.S., Han, C.T. and Hur, Y. (2015). Global gene-expression analysis to identify differentially expressed genes critical for the heat stress response in Brassica rapa. *PloS one*, *10*(6), p.e0130451.

Dubey, N.K., Goel, R., Ranjan, A., Idris, A., Singh, S.K., Bag, S.K. et.al (2013). Comparative transcriptome analysis of Gossypium hirsutum L. in response to sap sucking insects: aphid and whitefly. *BMC genomics*, *14*(1), p.241.

Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. and Trent, J.M. (1999). Expression profiling using cDNA microarrays. *Nature genetics*, *21*, pp.10-14.

Dunwell, J.M., Moya-León, M.A. and Herrera, R. (2001). Transcriptome analysis and crop improvement:(A review). *Biological research*, *34*(3-4), pp.153-164.

Ekström, A.M., Serafini, M., Nyren, O., Wolk, A., Bosetti, C. and Bellocco, R. (2011). Dietary quercetin intake and risk of gastric cancer: results from a population-based study in Sweden. *Annals of Oncology*, *22*(2), pp.438-443.

Ellis, P.R., Pink, D.A.C., Phelps, K., Jukes, P.L., Breeds, S.E. and Pinnegar, A.E. (1998). Evaluation of a core collection of Brassica oleracea accessions for resistance to Brevicoryne brassicae, the cabbage aphid.*Euphytica*, *103*(2), pp.149-160.

Ellis, P.R., Singh, R., Pink, D.A.C., Lynn, J.R. and Saw, P.L. (1996). Resistance to Brevicoryne brassicae in horticultural brassicas. *Euphytica*,*88*(2), pp.85-96.

Ellis, P.R., Kift, N.B., Pink, D.A.C., Jukes, P.L., Lynn, J. and Tatchell, G.M. (2000). Variation in resistance to the cabbage aphid (Brevicoryne brassicae) between and within wild and cultivated Brassica species. *Genetic Resources and Crop Evolution*, *47*(4), pp.395-401.

Elzinga, D.A., De Vos, M. and Jander, G. (2014). Suppression of plant defenses by a Myzus persicae (green peach aphid) salivary effector protein.*Molecular Plant-Microbe Interactions*, *27*(7), pp.747-756.

Elzinga, D.A. and Jander, G. (2013). The role of protein effectors in plant–aphid interactions. *Current opinion in plant biology*, *16*(4), pp.451-456.

Evans, L.T. (1996). *Crop evolution, adaptation and yield*. Cambridge university press.

FAO. (2008). Biofuels:Prospects,Risks and opportunity. The state of food and agriculture. FAO, Rome, p.129.

FAO, U., 2009. How to Feed the World in 2050. In *Rome: High-Level Expert Forum*.

Ferguson, M.E., Hearne, S.J., Close, T.J., Wanamaker, S., Moskal, W.A., Town, C.D. et.al. (2012). Identification, validation and high-throughput genotyping of transcribed gene SNPs in cassava. *Theoretical and Applied Genetics*, *124*(4), pp.685-695.

Finkel, E. (2009). Imaging. With'phenomics,'plant scientists hope to shift breeding into overdrive. *Science (New York, NY)*, *325*(5939), p.380-81.

Finley, J.W. (2003). The antioxidant responsive element (ARE) may explain the protective effects of cruciferous vegetables on cancer. *Nutrition reviews*,*61*(7), p.250.

Ford-Lloyd, B.V., Schmidt, M., Armstrong, S.J., Barazani, O., Engels, J., Hadas, R. et.al. (2011). Crop wild relatives—undervalued, underutilized and under threat?.*BioScience*, *61*(7), pp.559-565.

Foyer, C.H., Verrall, S.R. and Hancock, R.D. (2015). Systematic analysis of phloem-feeding insect-induced transcriptional reprogramming in Arabidopsis highlights common features and reveals distinct responses to specialist and generalist insects. *Journal of Experimental Botany*, *66*(2), pp.495-512.

Fukao, T., Harris, T. and Bailey-Serres, J. (2009). Evolutionary analysis of the Sub1 gene cluster that confers submergence tolerance to domesticated rice. *Annals of Botany*, *103*(2), pp.143-150.

Furbank, R.T. and Tester, M. (2011). Phenomics–technologies to relieve the phenotyping bottleneck. *Trends in plant science*, *16*(12), pp.635-644.

Fürstenberg-Hägg, J., agrobelny, M. and Bak, S. (2013). Plant defense against insect herbivores. *International journal of molecular sciences*, *14*(5), pp.10242-10297.

Gabryś, B. and Pawluk, M. (1999). Acceptability of different species of Brassicaceae as hosts for the cabbage aphid. In *Proceedings of the 10th International Symposium on Insect-Plant Relationships* (pp. 105-109). Springer Netherlands.

Gaeta, R.T., Yoo, S.Y., Pires, J.C., Doerge, R.W., Chen, Z.J. and Osborn, T.C. (2009). Analysis of gene expression in resynthesized Brassica napus allopolyploids using Arabidopsis 70mer oligo microarrays. *PLoS One*, *4*(3), p.e4760.

Gao, J., Yu, X., Ma, F. and Li, J. (2014). RNA-seq analysis of transcriptome and glucosinolate metabolism in seeds and sprouts of broccoli (Brassica oleracea var. italic). *PloS one*, *9*(2), p.e88804.

Gao, L.L., Klingler, J.P., Anderson, J.P., Edwards, O.R. and Singh, K.B. (2008). Characterization of pea aphid resistance in Medicago truncatula. *Plant Physiology*, *146*(3), pp.996-1009.

García-Palacios, P., Milla, R., Delgado-Baquerizo, M., Martín-Robles, N., Álvaro-Sánchez, M. and Wall, D.H. (2013). Side-effects of plant domestication: ecosystem impacts of changes in litter quality. *New Phytologist*, *198*(2), pp.504-513.

Garthwaite, D. (2012). Changes in Pesticide Use in Vegetables Since 2007–A Review of the Recent 2011 FERA Survey of Pesticide Use in Edible Protected and Outdoor Vegetable Crops. *Outlooks on Pest Management*,*23*(6), pp.248-252.

Garzo, E., Moreno, A., Hernando, S., Mariño, V., Torne, M., Santamaria, E. et.al. (2015). Electrical penetration graph technique as a tool to monitor the early stages of aphid resistance to insecticides. *Pest management science*, 72(4), pp.707–18.

Gase, K. and Baldwin, I.T. (2012). Transformational tools for next-generation plant ecology: manipulation of gene expression for the functional analysis of genes. *Plant Ecology & Diversity*, *5*(4), pp.485-490.

Gavaghan, C.L., Wilson, I.D. and Nicholson, J.K. (2002). Physiological variation in metabolic phenotyping and functional genomic studies: use of orthogonal signal correction and PLS-DA. *FEBS letters*, *530*(1), pp.191-196.

Ghaffar, M.B.A., Pritchard, J. and Ford-Lloyd, B. (2011). Brown planthopper (N. lugens Stal) feeding behaviour on rice germplasm as an indicator of resistance. *PLoS One*, *6*(7), p.e22137.

Gill, H.., Harsh, G. and Gillett-Kaufman, J.L. (2013). Cabbage aphid - Brevicoryne brassicae Linnaeus. *(EENY-577), University of Florida, Gainesville, FL*.

Giri, A.P., Wünsche, H., Mitra, S., Zavala, J.A., Muck, A., Svatoš, A. et.al. (2006). Molecular interactions between the specialist herbivore Manduca sexta (Lepidoptera, Sphingidae) and its natural host Nicotiana attenuata. VII. Changes in the plant's proteome. *Plant Physiology*, *142*(4), pp.1621-1641.

Goggin, F.L., Lorence, A. and Topp, C.N. (2015). Applying high-throughput phenotyping to plant–insect interactions: picturing more resistant crops.*Current*

Opinion in Insect Science, *9*, pp.69-76.

Goggin, F.L. (2007). Plant–aphid interactions: molecular and ecological perspectives. *Current opinion in plant biology*, *10*(4), pp.399-408.

Granado, F., Olmedilla, B. and Blanco, I. (2003). Nutritional and clinical relevance of lutein in human health. *British Journal of Nutrition*, *90*(03), pp.487-502.

Grubb, C.D. and Abel, S. (2006). Glucosinolate metabolism and its control.*Trends in plant science*, *11*(2), pp.89-100.

Guerrieri, E. and Digilio, M.C. (2008). Aphid-plant interactions: a review.*Journal of Plant Interactions*, *3*(4), pp.223-232.

Gurr, S.J. and Rushton, P.J. (2005). Engineering plants with increased disease resistance: what are we going to express?. *TRENDS in Biotechnology*, *23*(6), pp.275-282.

Hajjar, R. and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica*,*156*(1-2), pp.1-13.

Hammond, J.P., Broadley, M.R., Craigon, D.J., Higgins, J., Emmerson, Z.F., Townsend, H.J. et.al. (2005). Using genomic DNA-based probe-selection to improve the sensitivity of high-density oligonucleotide arrays when applied to heterologous species. *Plant Methods*, *1*(1), p.1.

Hancock, R.D., Hogenhout, S. and Foyer, C.H. (2015). Mechanisms of plant–insect interaction. *Journal of experimental botany*, *66*(2), pp.421-424.

Hardie, J. (1989). Spectral specificity for targeted flight in the black bean aphid, Aphis fabae. *Journal of Insect Physiology*, *35*(8), pp.619-626.

Harvey, J.A., Witjes, L.M., Benkirane, M., Duyts, H. and Wagenaar, R. (2007). Nutritional suitability and ecological relevance of Arabidopsis thaliana and Brassica oleracea as foodplants for the cabbage butterfly, Pieris rapae.*Plant Ecology*, *189*(1), pp.117-126.

Helden, M. and Tjallingii, W.F. (1993). Tissue localisation of lettuce resistance to the aphid Nasonovia ribisnigri using electrical penetration graphs. *Entomologia Experimentalis et Applicata*, *68*(3), pp.269-278.

Herms, D.A. and Mattson, W.J. (1992). The dilemma of plants: to grow or defend. *Quarterly review of biology*, pp.283-335.

Hewer, A., Becker, A. and van Bel, A.J. (2011). An aphid's Odyssey–the cortical quest for the vascular bundle. *The Journal of experimental biology*,*214*(22), pp.3868-3879.

Hines, R.L. and Hutchison, W.D. (2013). Cabbage aphids. VegEdge, vegetable IPM resource for the Midwest. *University of Minnesota, Minneapolis, MN.(2 October 2013).*

Hong, S.P. and Carlson, M. (2007). Regulation of snf1 protein kinase in response to environmental stress. *Journal of Biological Chemistry*, *282*(23), pp.16838-16845.

Howe, G.A. and Jander, G. (2008). Plant immunity to insect herbivores.*Annu. Rev. Plant Biol.*, *59*, pp.41-66.

Hu, J., Sadowski, J., Osborn, T.C., Landry, B.S. and Quiros, C.F. (1998). Linkage group alignment from four independent Brassica oleracea RFLP maps. *Genome*, *41*(2), pp.226-235.

Huang, F., Tjallingii, W.F., Zhang, P., Zhang, J., Lu, Y. and Lin, J. (2012). EPG waveform characteristics of solenopsis mealybug stylet penetration on cotton. *Entomologia Experimentalis et Applicata*, *143*(1), pp.47-54.

Huang, H.L., Wu, Y.C., Su, L.J., Huang, Y.J., Charoenkwan, P., Chen, W.L. et.al. (2015). Discovery of prognostic biomarkers for predicting lung cancer metastasis using microarray and survival data. *BMC bioinformatics*, *16*(1), p.54.

Hui, D., Iqbal, J., Lehmann, K., Gase, K., Saluz, H.P. and Baldwin, I.T. (2003). Molecular Interactions between the specialist herbivoremanduca sexta (Lepidoptera, Sphingidae) and its natural host nicotiana attenuata: V. Microarray analysis and further characterization of large-scale changes in herbivore-induced mRNAs. *Plant physiology*, *131*(4), pp.1877-1893.

Hyten, D.L., Song, Q., Fickus, E.W., Quigley, C.V., Lim, J.S., Choi, I.Y. et.al (2010). High-throughput SNP discovery and assay development in common bean. *BMC genomics*,*11*(1), p.475.

Hyten, D.L., Song, Q., Zhu, Y., Choi, I.Y., Nelson, R.L., Costa, J.M. et.al. (2006). Impacts of genetic bottlenecks on soybean genome diversity. *Proceedings of the National Academy of Sciences*, *103*(45), pp.16666-16671.

Ishida, M., Hara, M., Fukino, N., Kakizaki, T. and Morimitsu, Y. (2014). Glucosinolate metabolism, functionality and breeding for the improvement of Brassicaceae vegetables. *Breeding science*, *64*(1), p.48-59.

Ishikawa, K., Yoshimura, K., Ogawa, T. and Shigeoka, S. (2010). Distinct regulation of Arabidopsis ADP-ribose/NADH pyrophosphohydrolases, AtNUDX6 and 7, in biotic and abiotic stress responses. *Plant signaling & behavior*, *5*(7), pp.839-841.

Işık, M. and Görür, G. (2009). Aphidicidial activity of seven essential oils against the cabbage aphid, Brevicoryne brassicae L.(Hemiptera: Aphididae).*Munis Entomology and Zoology*, *4*(2), pp.424-431.

Izzah, N.K., Lee, J., Jayakodi, M., Perumal, S., Jin, M., Park, B.S. et. al. (2014). Transcriptome sequencing of two parental lines of cabbage (Brassica oleracea L. var. capitata L.) and construction of an EST-based genetic map. *BMC genomics*, *15*(1), p.1.

Janssen, J.A.M., Tjallingii, W.F. and Lenteren, J.V. (1989). Electrical recording and ultrastructure of stylet penetration by the greenhouse whitefly.*Entomologia*

Experimentalis et Applicata, *52*(1), pp.69-81.

Jaouannet, M., Morris, J.A., Hedley, P.E. and Bos, J.I. (2015). Characterization of Arabidopsis transcriptional responses to different aphid species reveals genes that contribute to host susceptibility and non-host resistance. *PLoS Pathog*, *11*(5), p.e1004918.

Jaouannet, M., Rodriguez, P.A., Thorpe, P., Lenoir, C.J., MacLeod, R., Escudero-Martinez, C. and Bos, J.I. (2014). Plant immunity in plant-aphid interactions. *Front. Plant Sci*, *5*(663), pp.10-3389.

Jones, A.M.E., Winge, P., Bones, A.M., Cole, R. and Rossiter, J.T. (2002). Characterization and evolution of a myrosinase from the cabbage aphid Brevicoryne brassicae. *Insect biochemistry and molecular biology*, *32*(3), pp.275-284.

Kagale, S., Robinson, S.J., Nixon, J., Xiao, R., Huebert, T., Condie, J. et.al. (2014). Polyploid evolution of the Brassicaceae during the Cenozoic era. *The Plant Cell*, *26*(7), pp.2777-2791.

Karban, R. and Baldwin, I.T. (2007). *Induced responses to herbivory*. University of Chicago Press.

Karim, S., Lundh, D., Holmström, K.O., Mandal, A. and Pirhonen, M., 2005. Structural and functional characterization of AtPTR3, a stress-induced peptide transporter of Arabidopsis. *Journal of molecular modeling*, *11*(3), pp.226-236.

Kaul, S., Koo, H.L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L.J. et.al. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana.*nature*, *408*(6814), pp.796-815.

Kazana, E., Pope, T.W., Tibbles, L., Bridges, M., Pickett, J.A., Bones, A.M. et.al (2007). The cabbage aphid: a walking mustard oil bomb. *Proceedings of the Royal Society of London B: Biological Sciences*, *274*(1623), pp.2271-2277.

Kehr, J. (2006). Phloem sap proteins: their identities and potential roles in the interaction between plants and phloem-feeding insects. *Journal of Experimental Botany*, *57*(4), pp.767-774.

Kempema, L.A., Cui, X., Holzer, F.M. and Walling, L.L. (2007). Arabidopsis transcriptome changes in response to phloem-feeding silverleaf whitefly nymphs. Similarities and distinctions in responses to aphids. *Plant Physiology*, *143*(2), pp.849-865.

Kerchev, P.I., Fenton, B., Foyer, C.H. and Hancock, R.D. (2012). Plant responses to insect herbivory: interactions between photosynthesis, reactive oxygen species and hormonal signalling pathways. *Plant, cell & environment*, *35*(2), pp.441-453.

Kessing, J.L.M. and RFL, M. (1991). Brevicoryne brassicae. *Crop Knowledge Master*. Department of Entomology, Honolulu,Hawaii.

Kessler, A. and Halitschke, R. (2007). Specificity and complexity: the impact of

herbivore-induced plant responses on arthropod community structure.*Current opinion in plant biology*, *10*(4), pp.409-414.

Kessler, A., Halitschke, R. and Baldwin, I.T. (2004). Silencing the jasmonate cascade: induced plant defenses and insect populations. *Science,305*(5684), pp.665-668.

Kessler, A. and Baldwin, I.T.(2002). Plant responses to insect herbivory: the emerging molecular analysis. *Annual review of plant biology*, *53*(1), pp.299-328.

Kettles, G.J., Drurey, C., Schoonbeek, H.J., Maule, A.J. and Hogenhout, S.A. (2013). Resistance of Arabidopsis thaliana to the green peach aphid, Myzus persicae, involves camalexin and is regulated by microRNAs. *New Phytologist*, *198*(4), pp.1178-1190.

Khan, S.A., Marimuthu, M., Predeesh, C., Aguirre-Rojas, L.M., Reese, J.C. and Smith, C.M. (2015). Electrical Penetration Graph Recording of Russian Wheat Aphid (Hemiptera: Aphididae) Feeding on Aphid-Resistant Wheat and Barley. *Journal of economic entomology*, *108*(5), pp.2465-2470.

Katam, K., Jones, K.A. and Sakata, K. (2015). Advances in proteomics and bioinformatics in agriculture research and crop improvement. *Journal of Proteomics & Bioinformatics*, 8, pp.039-048.

Kim, H.A., Lim, C.J., Kim, S., Choe, J.K., Jo, S.H., Baek, N. and Kwon, S.Y. (2014). High-throughput sequencing and de novo assembly of Brassica oleracea var. Capitata L. for transcriptome analysis. *PLoS One*, *9*(3), p.e92087.

Kim, J.H., Lee, B.W., Schroeder, F.C. and Jander, G. (2008). Identification of indole glucosinolate breakdown products with antifeedant effects on Myzus persicae (green peach aphid). *The Plant Journal*, *54*(6), pp.1015-1026.

Kim, K., Bolotin, E., Theusch, E., Huang, H., Medina, M.W. and Krauss, R.M. (2014). Prediction of LDL cholesterol response to statin using transcriptomic and genetic variation. *Genome biology*, *15*(9), p.460.

King, G.J. (2005). Brassica: Harvesting the genome, diversity and products.*The Multinational Brassica Genome Project. BBSRC, Rothamstead Research Centre, Norwich, UK*, *1*, pp.1-18.

Klingler, J., Powell, G., Thompson, G.A. and Isaacs, R. (1998). Phloem specific aphid resistance in Cucumis melo line AR 5: effects on feeding behaviour and performance of Aphis gossypii. *Entomologia Experimentalis et Applicata*, *86*(1), pp.79-88.

Kloth, K.J., Thoen, M.P., Bouwmeester, H.J., Jongsma, M.A. and Dicke, M. (2012). Association mapping of plant resistance to insects. *Trends in plant science*, *17*(5), pp.311-319.

Knoblauch, M. and van Bel, A.J. (1998). Sieve tubes in action. *The Plant Cell*, *10*(1), pp.35-50.

Kojima, S., Bohner, A., Gassert, B., Yuan, L. and Wirén, N.V. (2007). AtDUR3

represents the major transporter for high-affinity urea transport across the plasma membrane of nitrogen-deficient Arabidopsis roots. *The Plant Journal*, *52*(1), pp.30-40.

Korth, K.L. (2003). Profiling the response of plants to herbivorous insects. *Genome biology*, 4(7), p.221.

Kumar, S., Atri, C., Sangha, M.K. and Banga, S.S., 2011. Screening of wild crucifers for resistance to mustard aphid, Lipaphis erysimi (Kaltenbach) and attempt at introgression of resistance gene (s) from Brassica fruticulosa to Brassica juncea. *Euphytica*, *179*(3), pp.461-470.

Kuśnierczyk, A., Winge, P., Jørstad, T.S., Troczyńska, J., Rossiter, J.T. and Bones, A.M. (2008). Towards global understanding of plant defence against aphids-timing and dynamics of early Arabidopsis defence responses to cabbage aphid (Brevicoryne brassicae) attack. *Plant, cell & environment*,*31*(8), p.1097.

Lagercrantz, U. and Lydiate, D.J. (1996). Comparative genome mapping in Brassica. *Genetics*, *144*(4), pp.1903-1910.

Lange, E.S., Balmer, D., Mauch-Mani, B. and Turlings, T.C. (2014). Insect and pathogen attack and resistance in maize and its wild ancestors, the teosintes. *New Phytologist*, *204*(2), pp.329-341.

Lee, H.S., Wang, J., Tian, L., Jiang, H., Black, M.A., Madlung, A. et.al.(2004). Sensitivity of 70-mer oligonucleotides and cDNAs for microarray analysis of gene expression in Arabidopsis and its related species. *Plant biotechnology journal*, *2*(1), pp.45-57.

Lee, S.C., Lim, M.H., Kim, J.A., Lee, S.I., Kim, J.S., Jin, M. et.al. (2008). Transcriptome analysis in Brassica rapa under the abiotic stresses using Brassica 24K oligo microarray. *Mol Cells*, *26*(6), pp.595-605.

Lei, H., Tjallingii, W.F. and Lenteren, J.V. (1998). Probing and feeding characteristics of the greenhouse whitefly in association with host-plant acceptance and whitefly strains. *Entomologia Experimentalis et Applicata*,*88*(1), pp.73-80.

León, J., Rojo, E. and Sánchez-Serrano, J.J., 2001. Wound signalling in plants. *Journal of Experimental Botany*, *52*(354), pp.1-9.

Lester, R.N. (1989). Evolution under domestication involving disturbance of genic balance. *Euphytica*, 44, pp.125–132.

Lewis, W.J., Van Lenteren, J.C., Phatak, S.C. and Tumlinson, J.H. (1997). A total system approach to sustainable pest management. *Proceedings of the National Academy of Sciences*, *94*(23), pp.12243-12248.

Li, A., Walling, J., Ahn, S., Kotliarov, Y., Su, Q., Quezado, M., Oberholtzer, J.C. et.al (2009). Unsupervised analysis of transcriptomic profiles reveals six glioma subtypes. *Cancer research*,*69*(5), pp.2091-2099

Liang, D., Liu, M., Hu, Q., He, M., Qi, X., Xu, Q. et.al. (2015). Identification of

differentially expressed genes related to aphid resistance in cucumber (Cucumis sativus L.). *Scientific reports*, *5*, p.9645.

Liechti, R. and Farmer, E.E. (2002). The jasmonate pathway. *Science*,*296*(5573), pp.1649-1650.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I.A. et.al. (2014). The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. *Nature communications*, *5*. p.3930.

Lockhart, D.J. and Winzeler, E.A. (2000). Genomics, gene expression and DNA arrays. *Nature*, 405(6788), pp.827–836.

London, S.J., Yuan, J.M., Chung, F.L., Gao, Y.T., Coetzee, G.A., Ross, R.K. et.al. (2000). Isothiocyanates, glutathione S-transferase M1 and T1 polymorphisms, and lung-cancer risk: a prospective study of men in Shanghai, China. *The Lancet*, *356*(9231), pp.724-729.

Louis, J. and Shah, J. (2015). Plant defence against aphids: the PAD4 signalling nexus. *Journal of experimental botany*, *66*(2), pp.449-454.

Louis, J., Singh, V. and Shah, J. (2012). Arabidopsis thaliana-aphid interaction. *The Arabidopsis Book*, p.e0159.

Lowe, A.J., Moule, C., Trick, M. and Edwards, K.J. (2004). Efficient large-scale development of microsatellites for marker and mapping applications in Brassica crop species. *Theoretical and applied genetics*, *108*(6), pp.1103-1112.

Lu, Y., Huggins, P. and Bar-Joseph, Z. (2009). Cross species analysis of microarray expression data. *Bioinformatics*, *25*(12), pp.1476-1483.

Mamun, M.S.A., Ali, M.H., Ferdous, M.M., Rahman, M.A. and Hossain, M.A. (2010). Assessment of several mustard varieties resistance to mustard aphid, Lipaphis erysimi (Kalt.). *J. Soil Nature*, *4*(1), pp.34-38.

Martinière, A., Zancarini, A. and Drucker, M. (2009). Aphid transmission of Cauliflower mosaic virus: the role of the host plant. *Plant signaling & behavior*, *4*(6), pp.548-550.

Matis, J.H., Kiffe, T.R., Matis, T.I., Jackman, J.A. and Singh, H. (2007). Population size models based on cumulative size, with application to aphids.*ecological modelling*, *205*(1), pp.81-92.

Maxted, N., Kell, S.P. and Ford-Lloyd, B.V. (2008). Crop wild relative conservation and use: establishing the context. In: Maxted, N., Dulloo, E., Ford-Lloyd, B., Kell, S.P. and Iriondo, J.M. (eds.), Crop Wild Relative Conservation and Use. CABI Publishing, Wallingford. Pp. 3–30.

Mazzanti, C., Zeiger, M.A., Costourous, N., Umbricht, C., Westra, W.H., Smith, D., Somervell, H., Bevilacqua, G., Alexander, H.R. and Libutti, S.K., 2004. Using gene expression profiling to differentiate benign versus malignant thyroid tumors. *Cancer Research*, *64*(8), pp.2898-2903.

Mba, C., Guimaraes, E.P. and Ghosh, K. (2012). Re-orienting crop improvement for the changing climatic conditions of the 21st century.*Agriculture & Food Security*, *1*(1), p.1.

Metcalf, R.L. (1989). Insect resistance to insecticides. *Pesticide science*,*26*(4), pp.333-358.

Mewis, I., Tokuhisa, J.G., Schultz, J.C., Appel, H.M., Ulrichs, C. and Gershenzon, J. (2006). Gene expression and glucosinolate accumulation in Arabidopsis thaliana in response to generalist and specialist herbivores of different feeding guilds and the role of defense signaling pathways. *Phytochemistry*, *67*(22), pp.2450-2462.

Mewis, I., Appel, H.M., Hom, A., Raina, R. and Schultz, J.C. (2005). Major signaling pathways modulate Arabidopsis glucosinolate accumulation and response to both phloem-feeding and chewing insects. *Plant Physiology*,*138*(2), pp.1149-1162.

Meyers, B.C., Galbraith, D.W., Nelson, T. and Agrawal, V. (2004). Methods for transcriptional profiling in plants. Be fruitful and replicate. *Plant Physiology*, *135*(2), pp.637-652.

Miles, P.W. (1999). Aphid saliva. *Biological Reviews of the Cambridge Philosophical Society*, *74*(01), pp.41-85.

Mithöfer, A. and Boland, W. (2012). Plant defense against herbivores: chemical aspects. *Annual review of plant biology*, *63*, pp.431-450.

Mochida, K. and Shinozaki, K. (2011). Advances in omics and bioinformatics tools for systems analyses of plant functions. *Plant and Cell Physiology*,*52*(12), pp.2017-2038.

Monet, R., Guye, A. and Massonie, G. (1997). Breeding for resistance to green aphid Myzus persicae Sulzer in the peach. In *IV International Peach Symposium 465* (pp. 171-176).

Mooney, K.A., Halitschke, R., Kessler, A. and Agrawal, A.A. (2010). Evolutionary trade-offs in plants mediate the strength of trophic cascades.*Science*, *327*(5973), pp.1642-1644.

Moran, P.J., Cheng, Y., Cassell, J.L. and Thompson, G.A. (2002). Gene expression profiling of Arabidopsis thaliana in compatible plant-aphid interactions. *Archives of insect biochemistry and physiology*, *51*(4), pp.182-203.

Moran, P.J. and Thompson, G.A. (2001). Molecular responses to aphid feeding in Arabidopsis in relation to plant defense pathways. *Plant physiology*, 125(2), pp.1074–1085.

Morkunas, I. and Ratajczak, L. (2014). The role of sugar signaling in plant defense responses against fungal pathogens. *Acta Physiologiae Plantarum*,*36*(7), pp.1607-1619.

Moy, K.A., Yuan, J.M., Chung, F.L., Wang, X.L., Van Den Berg, D., Wang, R. et.al. (2009). Isothiocyanates, glutathione S-transferase M1 and T1 polymorphisms and

gastric cancer risk: A prospective study of men in Shanghai, China. *International Journal of Cancer*, *125*(11), pp.2652-2659.

Müller, C. and Riederer, M. (2005). Plant surface properties in chemical ecology. *Journal of chemical ecology*, *31*(11), pp.2621-2651.

Nagaharu, U. (1935). Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. *Jap J Bot*, *7*, pp.389-452.

Nagata, K., Washio, T., Kawahara, Y. and Unami, A. (2014). Toxicity prediction from toxicogenomic data based on class association rule mining.*Toxicology Reports*, *1*, pp.1133-1142.

Nalam, V.J., Keeretaweep, J., Sarowar, S. and Shah, J. (2012). Root-derived oxylipins promote green peach aphid performance on Arabidopsis foliage.*The Plant Cell*, *24*(4), pp.1643-1653.

Narusaka, M., Abe, H., Kobayashi, M., Kubo, Y. and Narusaka, Y. (2006). Comparative analysis of expression profiles of counterpart gene sets between Brassica rapa and Arabidopsis thaliana during fungal pathogen Colletotrichum higginsianum infection. *Plant Biotechnology*, *23*(5), pp.503-508.

Nayidu, N.K., Tan, Y., Taheri, A., Li, X., Bjorndahl, T.C., Nowak, J. et.al. (2014). Brassica villosa, a system for studying non-glandular trichomes and genes in the Brassicas. *Plant molecular biology*, *85*(4-5), pp.519-539.

Nayidu, N.K., Bonham-Smith, P. and Gruber, M. (2015). Brassica Villosa a Potential Tool to Improve the Insect or Disease Resistance of Brassica Crop Species. *Transcriptomics*, *3*(114), p.2.

Nevo, E. (2014). Evolution of wild emmer wheat and crop improvement.*Journal of Systematics and Evolution*, *52*(6), pp.673-696.

Nishio, T. (2014). Genomics and breeding of Brassicaceae crops. *Breeding science*, *64*(1), p.1.

Nishizawa, T., Tamaoki, M., Kaneko, Y., Aono, M., Kubo, A., Saji, H. et.al. (2012). High-throughput capture of nucleotide sequence polymorphisms in three Brassica species (Brassicaceae) using DNA microarrays. *American journal of botany*, *99*(3), pp.e94-e96.

Nottingham, S.F., Hardie, J., Dawson, G.W., Hick, A.J., Pickett, J.A., Wadhams, L.J. et.al (1991). Behavioral and electrophysiological responses of aphids to host and nonhost plant volatiles.*Journal of Chemical Ecology*, *17*(6), pp.1231-1242.

O'neill, C.M. and Bancroft, I. (2000). Comparative physical mapping of segments of the genome of Brassica oleracea var. alboglabra that are homoeologous to sequenced regions of chromosomes 4 and 5 of Arabidopsis thaliana. *The plant journal*, *23*(2), pp.233-243.

Palacios, I., Drucker, M., Blanc, S., Leite, S., Moreno, A. and Fereres, A. (2002).

Cauliflower mosaic virus is preferentially acquired from the phloem by its aphid vectors. *Journal of General Virology*, *83*(12), pp.3163-3171.

Palaniswamy, P. and Bodnaryk, R.P. (1994). A wild Brassica from Sicily provides trichome-based resistance against flea beetles, Phyllotreta cruciferae (Goeze)(Coleoptera: Chrysomelidae). *The Canadian Entomologist*,*126*(05), pp.1119-1130.

Park, J.Y., Koo, D.H., Hong, C.P., Lee, S.J., Jeon, J.W., Lee, S.H. et.al. (2005). Physical mapping and microsynteny of Brassica rapa ssp. pekinensis genome corresponding to a 222 kbp gene-rich region of Arabidopsis chromosome 4 and partially duplicated on chromosome 5. *Molecular Genetics and Genomics*, *274*(6), pp.579-588.

Partek. Inc. (2014). Partek® Genomics Suite™ version 6.6 Copyright©;2014. Partek Inc., St. Louis, MO, USA.

Pawełkowicz, M., Zieliński, K., Zielińska, D., Pląder, W., Yagi, K., Wojcieszek, M. et.al. (2016). Next generation sequencing and omics in cucumber (Cucumis sativus L.) breeding directed research. *Plant Science*, 242, pp.77-88.

Pegadaraju, V., Louis, J., Singh, V., Reese, J.C., Bautor, J., Feys, B.J. et.al. (2007). Phloem-based resistance to green peach aphid is controlled by Arabidopsis PHYTOALEXIN DEFICIENT4 without its signaling partner ENHANCED DISEASE SUSCEPTIBILITY1. The Plant Journal, 52(2), pp.332-341.

Pelgrom, K.T.B., Broekgaarden, C., Voorrips, R.E., Bas, N., Visser, R.G.F. and Vosman, B. (2015). Host plant resistance towards the cabbage whitefly in Brassica oleracea and its wild relatives. *Euphytica*, 202(2), pp.297-306.

Pérez-de-Castro, A.M.,Vilanova, S., Cañizares, J., Pascual, L., Blanca, J.M., Diez, M.J. et.al. (2012). Application of genomic tools in plant breeding. *Current genomics*, 13(3), pp.179-195.

Pimentel, D. (2005). Environmental and economic costs of the application of pesticides primarily in the United States. *Environment, development and sustainability*, 7(2), pp.229-252.

Poelman, E.H. (2015). From induced resistance to defence in plant-insect interactions. *Entomologia Experimentalis et Applicata*, 157(1), pp.11-17.

Poelman, E.H., Van Loon, J.J., Van Dam, N.M., Vet, L.E. and Dicke, M. (2010). Herbivore-induced plant responses in Brassica oleracea prevail over effects of constitutive resistance and result in enhanced herbivore attack. *Ecological Entomology*, 35(2), pp.240-247.

Poelman, E.H., Galiart, R.J., Raaijmakers, C.E., Van Loon, J.J. and Van Dam, N.M. (2008). Performance of specialist and generalist herbivores feeding on cabbage cultivars is not explained by glucosinolate profiles. *Entomologia Experimentalis et Applicata*, 127(3), pp.218-228.

Polge, C., Jossier, M., Crozet, P., Gissot, L. and Thomas, M. (2008). β-subunits of the SnRK1 complexes share a common ancestral function together with expression and function specificities; physical interaction with nitrate reductase specifically occurs via AKINβ1-subunit. *Plant physiology*, 148(3), pp.1570-1582.

Pollard, D.G. (1973). Plant penetration by feeding aphids (Hemiptera, Aphidoidea): a review. *Bulletin of Entomological Research*, 62(04), pp.631-714.

Portman, S.L., Kariyat, R.R., Johnston, M.A., Stephenson, A.G. and Marden, J.H. (2015). Inbreeding compromises host plant defense gene expression and improves herbivore survival. *Plant signaling & behavior*, 10(5), p.e998548.

Powell, G., Tosh, C.R. and Hardie, J. (2006). Host plant selection by aphids: behavioral, evolutionary, and applied perspectives. Annu. Rev. Entomol., 51, pp.309-330.

Powell, G. and Hardie, J. (2000). Host-selection behaviour by genetically identical aphids with different plant preferences. *Physiological Entomology*, 25(1), pp.54-62.

Prado, E. and Tjallingii, W.F. (1999). Effects of experimental stress factors on probing behaviour by aphids. *Entomologia experimentalis et applicata*, 90(3), pp.289-300.

Prado, E. and Tjallingii, W.F. (1997). Effects of previous plant infestation on sieve element acceptance by two aphids. *Entomologia Experimentalis et Applicata*, 82(2), pp.189-200.

Prado, E. and Tjallingii, W.F. (1994). Aphid activities during sieve element punctures. Entomologia experimentalis et applicata, 72(2), pp.157-165.

Pratt, C., Pope, T.W., Powell, G. and Rossiter, J.T. (2008). Accumulation of glucosinolates by the cabbage aphid Brevicoryne brassicae as a defense against two coccinellid species. *Journal of chemical ecology*, 34(3), pp.323-329.

Prescott-Allen, R. and Prescott-Allen, C. (1988). Genes from the wild: using wild genetic resources for food and raw materials. *Earthscan publications*, London.

Prokopy, R.J. and Owens, E.D. (1983). Visual Detection of Plants by Herbivorous Insects. *Annual Review of Entomology*, 28(1), pp.337–364.

Rakow, G. (2004). Species origin and economic importance of Brassica. In *Brassica* ;Pua, E.C., Douglas, C., (Eds) Volume 54, pp. 3–11. Springer Berlin Heidelberg.

Ray, S. and Satya, P. (2014). Next generation sequencing technologies for next generation plant breeding. *Frontiers in plant science*, 5, p.367.

Razaq, M.U., Mehmood, A., Aslam, M., Ismail, M., Afzal, M. and Shad, S.A. (2011). Losses in yield and yield components caused by aphids to late sown Brassica napus L., Brassica juncea L. and Brassica carrinata A. Braun at Multan, Punjab (Pakistan). *Pakistan Journal of Botany*, 43(1), pp.319–324.

Redden, R. (2013). New Approaches for Crop Genetic Adaptation to the Abiotic

Stresses Predicted with Climate Change. *Agronomy*, 3(2), pp.419–432.

Rehman, A. and Powell, W. 92010). Host selection behaviour of aphid parasitoids (Aphidiidae: Hymenoptera). *Journal of Plant Breeding and Crop Science*, 2(10), pp.299-311.

Reis-Filho, J.S. and Pusztai, L. (2011). Gene expression profiling in breast cancer: classification, prognostication, and prediction. *The Lancet*, 378(9805), pp.1812-1823.

Reymond, P., Bodenhausen, N., Van Poecke, R.M., Krishnamurthy, V., Dicke, M. and Farmer, E.E. (2004). A conserved transcript pattern in response to a specialist and a generalist herbivore. *The Plant Cell*, 16(11), pp.3132-3147.

Reymond, P., Weber, H., Damond, M. and Farmer, E.E. (2000). Differential gene expression in response to mechanical wounding and insect feeding in Arabidopsis. The Plant Cell, 12(5), pp.707-719.

Rich, T.C.G. (1991). *Crucifers of Great Britain and Ireland* ISBN 10: 0., BSBI Publication.

Ronald, P. (2011). Plant genetics, sustainable agriculture and global food security. *Genetics*, 188(1), pp.11-20.

Rosenthal, J.P. and Dirzo, R.O.D.O.L.F.O. (1997). Effects of life history, domestication and agronomic selection on plant defence against insects: evidence from maizes and wild relatives. *Evolutionary Ecology*, 11(3), pp.337-355.

Sadowski, J., Gaubier, P., Delseny, M. and Quiros, C.F. (1996). Genetic and physical mapping inBrassica diploid species of a gene cluster defined in Arabidopsis thaliana. *Molecular and General Genetics*, 251(3), pp.298-306.

Saint Pierre, C. (2012). Phenotyping in controlled environments vs. field conditions. In Physiological Breeding I: Interdisciplinary Approaches to Improve Crop Adaptation (pp. 124-130). CIMMYT Mexico, DF.

Salvador-Recatalà, V. and Tjallingii, W.F. (2015). A New Application of the Electrical Penetration Graph (EPG) for Acquiring and Measuring Electrical Signals in Phloem Sieve Elements. *JoVE (Journal of Visualized Experiments)*, (101), pp.e52826-e52826.

Sander, J.D. and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology*, 32(4), pp.347–55.

Sangha, J.S., Chen, Y.H., Kaur, J., Khan, W., Abduljaleel, Z., Alanazi, M.S. et.al. (2013). Proteome analysis of rice (Oryza sativa L.) mutants reveals differentially induced proteins during brown planthopper (Nilaparvata lugens) infestation. *International journal of molecular sciences*, 14(2), pp.3921-3945.

Santamaria, M.E., Martínez, M., Cambra, I., Grbic, V. and Diaz, I. (2013). Understanding plant defence responses against herbivore attacks: an essential first step towards the development of sustainable resistance against pests. *Transgenic*

research, 22(4), pp.697-708.

Sauge, M.H. (1998). Analysis of the mechanisms of resistance to the green peach aphid in several Prunus genotypes. *Acta horticulturae*. 456, pp.737–739.

Sauge, M.H., Kervella, J. and Pascal, T. (1998). Settling behaviour and reproductive potential of the green peach aphid Myzus persicae on peach varieties and a related wild Prunus. *Entomologia experimentalis et applicata*, 89(3), pp.233-242.

Schenk, P.M., kazan, K., Wilson, I., Anderson, J.P., Richmond, T., Somerville, S.C. et al. (2000). Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 97, pp.11655–11660.

Schilling, C.H., Edwards, J.S. and Palsson, B.O. (1999). Toward metabolic phenomics: analysis of genomic data using flux balances. *Biotechnology Progress*, 15(3), pp.288-295.

Schoonhoven, L.M., Van Loon, J.J. and Dicke, M. (1998) *Insect-Plant biology* (No. Ed.1). Chapman & Hall.

Schoonhoven, L.M., Van Loon, J.J. and Dicke, M. (2005). *Insect-plant biology* (No. Ed. 2). Oxford University Press.

Schwachtje, J., Minchin, P.E.H., Jahnke, S., van Dongen, J.T., Schittko, U. and Baldwin I.T. (2006). SNF1-related kinases allow plants to tolerate herbivory by allocating carbon to roots. *Proceedings of the National Academy of Sciences of the United States of America*, 103(34), pp.12935–12940.

Schwachtje, J. and Baldwin, I.T. (2008). Why does herbivore attack reconfigure primary metabolism? *Plant physiology*, 146(3), pp.845–851.

Schwartzberg, E.G. and Tumlinson, J.H. (2014). Aphid honeydew alters plant defence responses. *Functional ecology*, 28(2), pp.386-394.

Schwarzkopf, A., Rosenberger, D., Niebergall, M., Gershenzon, J. and Kunert, G. (2013). To feed or not to feed: plant factors located in the epidermis, mesophyll, and sieve elements influence pea aphid's ability to feed on legume species. *PloS one*, 8(9), p.e75298.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z. et.al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature biotechnology*, 31(8), pp.686-688.

Shi, J., Huang, S., Zhan, J., Yu, J., Wang, X., Hua, W. et.al. (2014). Genome-wide microsatellite characterization and marker development in the sequenced Brassica crop species. *DNA research*, 21(1), pp.53-68.

Silva, A.X., Jander, G., Samaniego, H., Ramsey, J.S. and Figueroa, C.C. (2012). Insecticide resistance mechanisms in the green peach aphid Myzus persicae (Hemiptera: Aphididae) I: a transcriptomic survey. *PloS one*, 7(6), p.e36366.

Simmons, A.T., McGrath, D. and Gurr, G.M. (2005). Trichome characteristics of F1 Lycopersicon esculentum× L. cheesmanii f. minor and L. esculentum× L. pennellii hybrids and effects on Myzus persicae. *Euphytica*, 144(3), pp.313-320.

Simpson, C., Thomas, C., Findlay, K., Bayer, E. and Maule, A.J. (2009). An Arabidopsis GPI-anchor plasmodesmal neck protein with callose binding activity and potential to regulate cell-to-cell trafficking. *The Plant Cell*, 21(2), pp.581-594.

Singh, R., Ellis, P.R., Pink, D.A.C. and Phelps, K. (1994). An investigation of the resistance to cabbage aphid in brassica species. *Annals of Applied Biology*, 125(3), pp.457-465.

Singh, S.P., Kumar, S., Singh, Y.P. and Singh, R. (2014). Reliable screening technique for evaluation of wild crucifers against mustard aphid Lipaphis erysimi (Kalt.). *Indian journal of experimental biology*, 52, pp.1201-1205.

Smith, C.M. and Boyko, E.V. (2007). The molecular bases of plant resistance and defense responses to aphid feeding: current status. *Entomologia Experimentalis et Applicata*, 122(1), pp.1–16.

Smith, C.M. and Clement, S.L. (2012). Molecular Bases of Plant Resistance to Arthropods. *Annual Review of Entomology*, 57(1), pp.309–328.

Soffan, A. and Aldawood, A.S. (2015). Electrical Penetration Graph monitored feeding behavior of cowpea aphid, Aphis craccivora Koch.(Hemiptera: Aphididae), on faba bean, Vicia faba L.(Fabaceae), cultivars. *Turkish Journal of Entomology*, 39(4).pp.401–411.

Southern, E. M., Maskos, U. and Elder, J. K. (1992). Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. *Genomics*, 13(4), 1008-1017.

Spiller, N.J., Koenders, L. and Tjallingii, W.F. (1990). Xylem ingestion by aphids–a strategy for maintaining water balance. *Entomologia Experimentalis et Applicata*, 55(2), pp.101-104.

Starzycki, M., Starzycka, E., Pszczoa, J. and Pakos, T. (1996). Comparison of the efficiency of breeding methods of Brassica napus L. resistant to Phoma lingam (Tode ex Fr.) Desm. *Plant Breeding and Seed Science* 40(1/2): 25-30.

Suehiro, N., Natsuaki, T., Watanabe, T. and Okuda, S. (2004). An important determinant of the ability of Turnip mosaic virus to infect Brassica spp. and/or Raphanus sativus is in its P3 protein. *Journal of general virology*, 85(7), pp.2087-2098.

Sun, P., García, J., Comach, G., Vahey, M.T., Wang, Z., Forshey, B.M., Morrison, A.C. et.al. (2013). Sequential waves of gene expression in patients with clinically defined dengue illnesses reveal subtle disease phases and predict disease severity. *PLoS Negl Trop Dis*, 7(7), p.e2298.

Swanson, J.M., Wood, G.C., Xu, L., Tang, L.E., Meibohm, B., Homayouni, R. et.al.

(2012). Developing a gene expression model for predicting ventilator-associated pneumonia in trauma patients: a pilot study. *PloS one*, 7(8), p.e42065.

Tan, Q., Thomassen, M., Jochumsen, K.M., Mogensen, O., Christensen, K. and Kruse, T.A. (2008). Gene selection for predicting survival outcomes of cancer patients in microarray studies. *In Advances in Computer and Information Sciences and Engineering* (pp. 405-409). Springer Netherlands.

Tang, J.Y., Zielinski, R.E., Zangerl, A.R., Crofts, A.R., Berenbaum, M.R. and DeLucia, E.H. (2006). The differential effects of herbivory by first and fourth instars of Trichoplusia ni (Lepidoptera: Noctuidae) on photosynthesis in Arabidopsis thaliana. *Journal of Experimental Botany*, 57(3), pp.527-536.

Tanksley, S.D. and McCouch, S.R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. Science, 277(5329), pp.1063-1066.

Tarca, A.L., Romero, R. and Draghici, S. (2006). Analysis of microarray experiments of gene expression profiling. *American journal of obstetrics and gynecology*, 195(2), pp.373-388.

Teetes, G.L. (2014). Plant Resistance to Insects: A Fundamental Component of IPM. *In: E. B. Radcliffe,W. D. Hutchison & R. E. Cancelado [eds.], Radcliffe's IPM World Textbook*.

Tew, M. and Bekeko, Z. (2014). Diversity and conservation of wild crop relatives for source of resistance to major biotic stress : Experiences in Ethiopia. , 2(3), pp.331–348.

Thompson, G.A. and Goggin, F.L. (2006). Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. *Journal of Experimental Botany*, 57(4), pp.755-766.

Tjallingii, W.F. and Mayoral, A. (1992). Criteria for host-plant acceptance by aphids. *In Proceedings of the 8th International Symposium on Insect-Plant Relationships* (pp. 280-282). Springer Netherlands.

Tjallingii, W.F. (1995). Aphid-plant interactions: what goes on in the depth of the tissues. In *Proc. Exp. Appl. Entomol.*, 6, pp.163–169.

Tjallingii, W.F. (1985). Electrical nature of recorded signals during stylet penetration by aphids. *Entomologia Experimentalis et Applicata*, 38(2), pp.177-186.

Tjallingii, W.F. (2006). Salivary secretions by aphids interacting with proteins of phloem wound responses. *Journal of Experimental Botany*, 57(4), pp.739-745.

Tjallingii, W.F. and Esch, T.H. (1993). Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. *Physiological entomology*, 18(3), pp.317-328.

Tjallingii, W.F. and Gabryś, B. (1999). Anomalous stylet punctures of phloem sieve elements by aphids. *In Proceedings of the 10th International Symposium on Insect-Plant Relationships* (pp. 97-103). Springer Netherlands.

Tobias Züst & Agrawal, A.A. (2016). Mechanisms and evolution of plant resistance to aphids. *Nature Plants*, 2, pp.1–9.

Torres-Avilés, F., Romeo, J.S. and López-Kleine, L. (2014). Data mining and influential analysis of gene expression data for plant resistance gene identification in tomato (solanum lycopersicum). *Electronic Journal of Biotechnology*, 17(2), pp.79–82.

Townend, J. and Dickinson, A.L. (1995). A comparison of rooting environments in containers of different sizes. *Plant and Soil*, 175(1), pp.139–146.

Traka, M.H., Melchini, A. and Mithen, R.F. (2014). Sulforaphane and prostate cancer interception. *Drug discovery today*, 19(9), pp.1488–92.

Trevino, V., Tadesse, M.G., Vannucci, M., Al-Shahrour, F., Antczak, P., Durant, S. et.al. (2011). Analysis of normal-tumour tissue interaction in tumours: prediction of prostate cancer features from the molecular profile of adjacent normal cells. *PloS one*, 6(3), p.e16492.

Turcotte, M.M., Turley, N.E. and Johnson, M.T.J. (2014). The impact of domestication on resistance to two generalist herbivores across 29 independent domestication events. *The New phytologist*, 204(3), pp.671–81.

van Poppel, G., Verhoeven, D.T., Verhagen, H. and Goldbohm, R.A. (1999). Brassica vegetables and cancer prevention. Epidemiology and mechanisms. Advances in experimental medicine and biology, 472, pp.159–168.

Varambally, S., Yu, J., Laxman, B., Rhodes, D.R., Mehra, R., Tomlins, S.A. et.al. (2005). Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer cell*, 8(5), pp.393-406.

Varshney, R.K., Nayak, S.N., May, G.D. and Jackson, S.A. (2009). Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in biotechnology*, 27(9), pp.522-530.

Varshney, R.K., Hoisington, D. A. and Tyagi, A.K. (2006). Advances in cereal genomics and applications in crop breeding. *Trends in Biotechnology*, 24(11), pp.490–499.

Varshney, R.K., Terauchi, R. and McCouch, S.R. (2014). Harvesting the Promising Fruits of Genomics: Applying Genome Sequencing Technologies to Crop Breeding J. L. Dangl, ed. *PLoS Biology*, 12(6), p.e1001883.

Verdugo, J.A., Méndez, T., Ortiz–Martínez, S.A., Cumsille, R. and Ramírez, C.C. (2012). Variation in Resistance Mechanisms to the Green Peach Aphid Among Different Prunus persica Commercial Cultivars. *Journal of economic entomology*, 105(5), pp.1844-1855.

Visser, J.H. and Piron, P.G.M. (1997). Odour response profiles in aphids differentiating for species, clone, form and food. *Proceedings of the Section Experimental and Applied Entomology of the Netherlands Entomological Society*, 8, pp.115–120.

Voelckel, C. and Baldwin, I.T. (2004). Generalist and specialist lepidopteran larvae elicit different transcriptional responses in Nicotiana attenuata, which correlate with larval FAC profiles. *Ecology Letters*, 7(9), pp.770–775.

Wagner, G.J., Wang, E. and Shepherd, R.W. (2004). New approaches for studying and exploiting an old protuberance, the plant trichome. *Annals of Botany*, 93(1), pp.3-11.

Walia, H., Wilson, C., Condamine, P., Liu, X., Ismail, A.M., Zeng, L. et.al. (2005). Comparative transcriptional profiling of two contrasting rice genotypes under salinity stress during the vegetative growth stage. *Plant Physiology*, 139(2), pp.822-835.

Walling, L. (2000). The Myriad Plant Responses to Herbivores. *Journal of plant growth regulation*, 19(2), pp.195–216.

Walling, L.L. (2008). Avoiding effective defenses: strategies employed by phloem-feeding insects. *Plant physiology*, 146(3), pp.859–66.

Walsh, J.A., Sharpe, A.G., Jenner, C.E. and Lydiate, D.J. (1999). Characterisation of resistance to turnip mosaic virus in oilseed rape (Brassica napus) and genetic mapping of TuRB01. *Theoretical and Applied Genetics*, 99(7-8), pp.1149-1154.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S. et.al. (2011). The genome of the mesopolyploid crop species Brassica rapa. *Nature genetics*, 43(10), pp.1035-1039.

Wang, Y., Wang, X., Yuan, H., Chen, R., Zhu, L., He, R. and He, G. (2008). Responses of two contrasting genotypes of rice to brown planthopper. *Molecular plant-microbe interactions*, 21(1), pp.122-132.

War, A.R., Paulraj, M.G., Ahmad, T., Buhroo, A.A., Hussain, B., Ignacimuthu, S. and Sharma, H.C. (2012). Mechanisms of plant defense against insect herbivores. *Plant signaling & behavior*, 7(10), pp.1306-1320.

Wei, L., Xiao, M., Hayward, A. and Fu, D. (2013). Applications and challenges of next-generation sequencing in Brassica species. *Planta*, 238(6), pp.1005-1024.

Westman, A.L. and Kresovich, S. (1999). Simple sequence repeat (SSR)-based marker variation in Brassica nigra genebank accessions and weed populations. *Euphytica*, 109(2), pp.85–92.

Wilkinson, T., Ashford, D., Pritchard, J. and Douglas, A. (1997). Honeydew sugars and osmoregulation in the pea aphid Acyrthosiphon pisum. *Journal of Experimental Biology*, 200(15), pp.2137-2143.

Will, T., Steckbauer, K., Hardt, M., and van Bel, A. J. (2012). Aphid gel saliva: sheath structure, protein composition and secretory dependence on stylet-tip milieu. *PloS one*, 7(10), p.e46903.

Will, T., Tjallingi, W.F., Thönnessen, A. and van Bel, A.J.E. (2007). Molecular sabotage of plant defense by aphid saliva. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25), pp.10536–10541.

Will, T. and van Bel, A.J.E. (2006). Physical and chemical interactions between aphids and plants. *Journal of experimental botany*, 57(4), pp.729–37.

Will, T., Furch, A.C.U. and Zimmermann, M.R. (2013). How phloem-feeding insects face the challenge of phloem-located defenses. *Frontiers in plant science*, 4(August), p.336.

Will, T. and Vilcinskas, A. (2013). Aphid-proof plants: biotechnology-based approaches for aphid control. *Advances in biochemical engineering/biotechnology*, 136, pp.179–203.

Willcox, J.K., Catignani, G.L. and Lazarus, S. (2003). Tomatoes and cardiovascular health. *Critical reviews in food science and nutrition*, 43(1), pp.1–18.

Wink, M. (1988). Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. *Theoretical and Applied Genetics*, 75(2), pp.225–233.

Wittstock, U. and Gershenzon, J. (2002). Constitutive plant toxins and their role in defense against herbivores and pathogens. *Current Opinion in Plant Biology*, 5(4), pp.300–307.

Witzel, K., Neugart, S., Ruppel, S., Schreiner, M., Wiesner, M. and Baldermann, S. (2015). Recent progress in the use of 'omics technologies in brassicaceous vegetables. *Frontiers in plant science*, 6, p.244.

Wu, J. and Baldwin, I.T. (2010). New insights into plant responses to the attack from insect herbivores. *Annual review of genetics*, 44, pp.1–24.

Xia, X., Shao, Y., Jiang, J., Ren, L., Chen, F., Fang W. et al. (2014). Gene expression profiles responses to aphid feeding in chrysanthemum (Chrysanthemum morifolium). *BMC genomics*, 15(1), p.1050.

Xu, D., Li, S., Lin, L., Qi, F., Hang, X. and Sun, Y. (2016). Gene expression profiling to identify the toxicities and potentially relevant disease outcomes due to endosulfan exposure. *Toxicol. Res.*, 5(2), pp.621–632.

Yauk, C.L. and Berndt, M.L. (2007). Review of the literature examining the correlation among DNA microarray technologies. *Environmental and molecular mutagenesis*, 48(5), pp.380–94.

Yumurtaci, A., (2015). Utilization of wild relatives of wheat, barley, maize and oat in developing abiotic and biotic stress tolerant new varieties. *Emir. J. Food Agric*, 27(1), pp.1–23.

Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. *Nature reviews. Genetics*, 2(12), pp.983–9.

Zhang, C., Shi, H., Chen, L., Wang, X., Lü, B., Zhang, S. et.al. (2011). Harpin-induced expression and transgenic overexpression of the phloem protein gene AtPP2-A1 in Arabidopsis repress phloem feeding of the green peach aphid Myzus persicae.

BMC plant biology, 11(1), p.11.

Zhang, F., Yao, J., Ke, J., Zhang, L., Lam, V.Q., et al. (2015). Structural basis of JAZ repression of MYC transcription factors in jasmonate signalling. *Nature*, 525(7568), pp.269–73.

Zhang, L., Li, Z., Quan, R., Li, G., Wang, R., Huang, R. (2011). An AP2 domain-containing gene, *ESE1*, targeted by the ethylene signaling component EIN3 is important for the salt response in Arabidopsis. *Plant Physiol*, **157, pp** 854–865.

Zhang, W., Yu, Y., Hertwig, F., Thierry-Mieg, J., Zhang, W., Thierry-Mieg, D. et.al (2015). Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. *Genome biology*, 16(1), p.133.

Zhang, Z., Ober, J.A. and Kliebenstein, D.J. (2006). The gene controlling the quantitative trait locus EPITHIOSPECIFIER MODIFIER1 alters glucosinolate hydrolysis and insect resistance in Arabidopsis. *The Plant cell*, 18(6), pp.1524–36.

Zhao, H., Lin, J., Grossman, H.B., Hernandez, L.M., Dinney, C.P., Wu, X. (2007). Dietary isothiocyanates, GSTM1, GSTT1, NAT2 polymorphisms and bladder cancer risk. *International journal of cancer. Journal international du cancer*, 120(10), pp.2208–13.

Zhu, J. and Struss, D., (1991). Transfer of Phoma lingam resistance from B. nigra into B. napus. *Cruciferae Newsletter*, (17) pp14-15.

Appendix I: List of accessions used in the field experiment to explore the level of resistance in the brassica germplasm

S.no	Seed nr	cwr/ lr	**INSTCODE**	**ACCENUMB**	**GENUS**	**SPECIES**	**SUBTAXA**	**SUBTAUTHOR**	**CROPNAME**	**ORIGCTY**
1	1	cwr	DEU146	BRA 2850	Brassica	balearica				
2	2	cwr	DEU146	K 7557	Brassica	balearica				
3	3	cwr	DEU146	BRA 2848	Brassica	bourgeaui				ESP
4	4	cwr	DEU146	K 9825	Brassica	bourgeaui				ESP
5	5	cwr	DEU146	K 6501	Brassica	cretica				
6	6	cwr	GRC005	639	Brassica	cretica	aegaea			GRC
7	7	cwr	GRC005	1964	Brassica	cretica	aegaea			GRC
8	8	cwr	DEU146	K 10120	Brassica	cretica				TUR
9	9	cwr	DEU146	K 6631	Brassica	cretica				GRC
10	10	cwr	GBR004	3252	Brassica	cretica				
11	11	cwr	GRC005	657	Brassica	cretica				GRC
12	12	cwr	JPN059	Cr-5	Brassica	cretica				
13	13	cwr	JPN059	Cr-6	Brassica	cretica				TUR
14	14	cwr	JPN059	Cr-7	Brassica	cretica				TUR
15	15	cwr	JPN059	Cr-8	Brassica	cretica				TUR
16	16	cwr	JPN059	Cr-9	Brassica	cretica				TUR
17	17	cwr	DEU146	BRA 1810	Brassica	fruticulosa	fruticulosa			ESP
18	18	cwr	JPN059	Fr-401	Brassica	fruticulosa	mauritanica	(Coss.) Maire		DZA
19	19	cwr	JPN059	Fr-301	Brassica	fruticulosa	radicata			DZA
20	20	cwr	JPN059	Fr-503	Brassica	fruticulosa				MAR
21	21	cwr	GBR006	HRIGRU 12483	Brassica	hilarionis			st hilarion cabbage	CYP
22	22	cwr	DEU146	BRA 1166	Brassica	incana				
23	23	cwr	DEU146	BRA 1262	Brassica	incana				

24	24	cwr	DEU146	BRA 2856	Brassica	incana				ITA
25	25	cwr	DEU146	BRA 2918	Brassica	incana				ITA
26	26	cwr	DEU146	K 10373	Brassica	incana				UKR
27	27	cwr	DEU146	K 9238	Brassica	incana				
28	28	cwr	GBR006	HRIGRU 6691	Brassica	incana			wild species	ITA
29	29	cwr	JPN059	Inc-1	Brassica	incana				
30	30	cwr	DEU146	K 5997	Brassica	insularis				
31	31	cwr	JPN059	Ins-1	Brassica	insularis				TUN
32	32	cwr	DEU146	K 7635	Brassica	macrocarpa				
33	33	cwr	JPN059	O-502	Brassica	macrocarpa				ITA
34	34	cwr	JPN059	O-503	Brassica	macrocarpa				ITA
35	35	cwr	DEU146	K 8824	Brassica	maurorum				
36	36	cwr	DEU146	BRA 2920	Brassica	maurorum				
37	37	cwr	DEU146	K 9404	Brassica	montana				ITA
38	38	cwr	DEU146	BRA 1644	Brassica	montana				
39	39	cwr	DEU146	K 6675	Brassica	montana				ITA
40	40	cwr	DEU146	K 7220	Brassica	montana				
41	41	cwr	DEU146	K 7223	Brassica	montana				
42	42	cwr	DEU146	K 8380	Brassica	montana				
43	189	cwr	DEU146	BRA 2411	Brassica	oleracea	botrytis italica		broccoli	DEU
44	199	cwr	DEU146	BRA 2401	Brassica	oleracea	capitata		cabbage	TUR
45	200	cwr	DEU146	BRA 2476	Brassica	oleracea	capitata		cabbage	TUR
46	275	cwr	DEU146	K 8793	Brassica	oleracea	costata			ESP
47	313	cwr	DEU146	K 9249	Brassica	oleracea				ESP
48	320	cwr	GBR004	22785	Brassica	oleracea				GBR
49	321	cwr	GBR004	57071	Brassica	oleracea				GBR
50	322	cwr	GBR004	70373	Brassica	oleracea				GBR

51	323	cwr	GBR004	70410	Brassica	oleracea				GBR
52	324	cwr	GBR004	70421	Brassica	oleracea				GBR
53	325	cwr	GBR004	70432	Brassica	oleracea				GBR
54	326	cwr	GBR004	70465	Brassica	oleracea				GBR
55	327	cwr	GBR004	70476	Brassica	oleracea				GBR
56	328	cwr	GBR004	70498	Brassica	oleracea				GBR
57	329	cwr	GBR004	70502	Brassica	oleracea				GBR
58	330	cwr	GBR004	70513	Brassica	oleracea				GBR
59	331	cwr	GBR004	75194	Brassica	oleracea				GBR
60	332	cwr	GBR004	75208	Brassica	oleracea				GBR
61	333	cwr	GBR004	75219	Brassica	oleracea				GBR
62	334	cwr	GBR004	75220	Brassica	oleracea				GBR
63	335	cwr	GBR004	75231	Brassica	oleracea				GBR
64	336	cwr	GBR004	75242	Brassica	oleracea				GBR
65	337	cwr	GBR004	92483	Brassica	oleracea				GBR
66	342	cwr	GBR006	HRIGRU 2075	Brassica	oleracea			wild species	GBR
67	343	cwr	GBR006	HRIGRU 7234	Brassica	oleracea			white flowered kale	ZIM
68	344	cwr	GBR006	HRIGRU 7338	Brassica	oleracea			wild cabbage	NZL
69	345	cwr	GBR006	HRIGRU 7343	Brassica	oleracea			wild species	
70	346	cwr	GBR006	HRIGRU 7795	Brassica	oleracea			wild cabbage	GBR
71	347	cwr	GBR006	HRIGRU 7796	Brassica	oleracea			wild cabbage	GBR
72	348	cwr	GBR006	HRIGRU 7797	Brassica	oleracea			wild cabbage	GBR
73	349	cwr	GBR006	HRIGRU	Brassica	oleracea			wild cabbage	GBR

				8694						
74	350	cwr	GBR006	HRIGRU 8705	Brassica	oleracea			wild cabbage	GBR
75	351	cwr	GBR006	HRIGRU 8707	Brassica	oleracea			wild cabbage	GBR
76	352	cwr	GBR006	HRIGRU 8714	Brassica	oleracea			wild cabbage	GBR
77	353	cwr	GBR006	HRIGRU 8724	Brassica	oleracea			wild cabbage	GBR
78	354	cwr	GBR006	HRIGRU 9156	Brassica	oleracea			wild species	GBR
79	363	cwr	DEU146	Bra 2922	Brassica	bivoniana				ITA
80	386	cwr	SWE002	NGB16241	Brassica	oleracea			cabbage	DEU
81	387	cwr	SWE002	NGB21657	Brassica	oleracea				DNK
82	391	cwr	JPN059	Ro-101	Brassica	robertiana				ESP
83	392	cwr	DEU146	K 7690	Brassica	rupestris				ITA
84	393	cwr	DEU146	K 6877	Brassica	rupestris				ITA
85	394	cwr	DEU146	BRA 2851	Brassica	rupestris				ITA
86	395	cwr	DEU146	K 8823	Brassica	spinescens				
87	396	cwr	DEU146	K 9402	Brassica	villosa		drepanensis		ITA
88	397	cwr	DEU146	BRA 2923	Brassica	villosa		drepanensis		ITA
89	398	cwr	DEU146	K 6926	Brassica	villosa	drepanensis			ITA
90	399	cwr	GBR006	HRIGRU 6848	Brassica	villosa			wild species	ITA
91	400	cwr	JPN059	Vill-1	Brassica	villosa				ITA
92	401	cwr	DEU146	K 10259	Brassica					ITA
93	402	cwr	DEU146	K 10260	Brassica					ITA
94	404	cwr	DEU146	K 8025	Brassica					GEO
95	466	cwr	NLD037	CGN06903	Brassica	oleracea				FRA

96	467	cwr	NLD037	CGN18947	Brassica	oleracea				DEU
97	468	cwr	NLD037	CGN07149	Brassica					
98	469	cwr	NLD037	CGN14116	Brassica	villosa				
99	43	lr	BIH039	GB00039	Brassica	oleracea	acephala		kale	
100	47	lr	ESP009	BRS-0158	Brassica	oleracea	acephala		kale	ESP
101	53	lr	ESP009	BRS-0166	Brassica	oleracea	acephala		kale	ESP
102	54	lr	ESP009	BRS-0103	Brassica	oleracea	acephala		kale	ESP
103	56	lr	ESP009	BRS0410	Brassica	oleracea	acephala		kale	
104	59	lr	ESP009	BRS-0064	Brassica	oleracea	acephala		kale	ESP
105	61	lr	ESP009	BRS-0161	Brassica	oleracea	acephala		kale	ESP
106	63	lr	ESP009	BRS-0160	Brassica	oleracea	acephala		kale	ESP
107	66	lr	ESP009	BRS-0201	Brassica	oleracea	acephala		kale	ESP
108	67	lr	ESP009	BRS0555	Brassica	oleracea	acephala		kale	
109	69	lr	ESP009	BRS-0292	Brassica	oleracea	acephala		kale	ESP
110	74	lr	FRA010	10	Brassica	oleracea	acephala		fodder kale	FRA
111	75	lr	FRA010	364	Brassica	oleracea	acephala		fodder kale	FRA
112	77	lr	FRA010	17	Brassica	oleracea	acephala		fodder kale	FRA
113	82	lr	FRA010	320	Brassica	oleracea	acephala		fodder kale	FRA
114	87	lr	GBR006	HRIGRU 12261	Brassica	oleracea	acephala		portuguese kale	PRT
115	96	lr	GBR006	HRIGRU 10000	Brassica	oleracea	acephala		kale	ESP
116	97	lr	GBR006	HRIGRU 9994	Brassica	oleracea	acephala		kale	ESP
117	100	lr	GBR006	HRIGRU 4502	Brassica	oleracea	acephala		fodder kale	IRL
118	101	lr	GBR006	HRIGRU 12092	Brassica	oleracea	acephala		portuguese kale	PRT
119	105	lr	GBR006	HRIGRU	Brassica	oleracea	acephala		portuguese	PRT

				9591					kale	
120	111	lr	HRV044	A	Brassica	oleracea	acephala		kale	HRV
121	113	lr	NLD037	CGN14111	Brassica	oleracea	acephala		kale	ESP
122	115	lr	NLD037	CGN18466	Brassica	oleracea	acephala		kale	TUR
123	116	lr	NLD037	CGN18468	Brassica	oleracea	acephala		kale	TUR
124	119	lr	POL003	PL173503	Brassica	oleracea	acephala		kale	POL
125	121	lr	DEU271	CR 2361	Brassica	oleracea	acephala viridis			EGY
126	122	lr	DEU271	CR 2364	Brassica	oleracea	acephala viridis			GEO
127	125	lr	DEU271	CR 2384	Brassica	oleracea	acephala viridis			GEO
128	127	lr	DEU271	CR 2378	Brassica	oleracea	acephala viridis			ITA
129	128	lr	DEU271	CR 2370	Brassica	oleracea	acephala viridis			GEO
130	131	lr	DEU271	CR 2390	Brassica	oleracea	acephala viridis			YUG
131	132	lr	DEU271	CR 2590	Brassica	oleracea	acephala viridis			ESP
132	133	lr	DEU271	CR 2363	Brassica	oleracea	acephala viridis			GEO
133	134	lr	DEU271	CR 2388	Brassica	oleracea	acephala viridis			GEO
134	135	lr	DEU271	CR 2371	Brassica	oleracea	acephala viridis			YUG
135	136	lr	DEU146	K 4699	Brassica	oleracea	acephala medullosa		marrow stem kale	POL
136	137	lr	DEU271	CR 2181	Brassica	oleracea	acephala medullosa		marrow stem kale	POL
137	139	lr	DEU271	CR 1288	Brassica	oleracea	acephala medullosa		marrow stem kale	FRA
138	144	lr	DEU146	BRA 1493	Brassica	oleracea	acephala sabellica		borecole	
139	145	lr	DEU146	BRA 1840	Brassica	oleracea	acephala sabellica		borecole	DEU
140	146	lr	NLD037	CGN15120	Brassica	oleracea	acephala sabellica		borecole	NLD
141	150	lr	NLD037	CGN14037	Brassica	oleracea	alboglabra		chinese kale	THA
142	152	lr	GBR006	HRIGRU 4495	Brassica	oleracea	botrytis		winter cauliflower	IRL
143	156	lr	ESP026	A-B-14	Brassica	oleracea	botrytis botrytis		cauliflower	ESP

144	162	lr	FRA010	638	Brassica	oleracea	botrytis botrytis		cauliflower	FRA
145	166	lr	FRA010	634	Brassica	oleracea	botrytis botrytis		cauliflower	FRA
146	167	lr	FRA010	633	Brassica	oleracea	botrytis botrytis		cauliflower	FRA
147	169	lr	NLD037	CGN11982	Brassica	oleracea	botrytis botrytis		cauliflower	IDN
148	174	lr	NLD037	CGN14026	Brassica	oleracea	botrytis botrytis		cauliflower	ITA
149	179	lr	DEU146	BRA 114	Brassica	oleracea	botrytis italica		broccoli	ITA
150	188	lr	DEU146	K 6599	Brassica	oleracea	botrytis italica		broccoli	ITA
151	190	lr	DEU146	BRA 1559	Brassica	oleracea	capitata		cabbage	GEO
152	198	lr	DEU146	BRA 379	Brassica	oleracea	capitata		cabbage	CHE
153	201	lr	ESP009	BRS-0083	Brassica	oleracea	capitata		cabbage	ESP
154	203	lr	ESP009	BRS-0152	Brassica	oleracea	capitata		cabbage	ESP
155	208	lr	ESP026	CL-B-5	Brassica	oleracea	capitata		cabbage	ESP
156	211	lr	FRA010	383	Brassica	oleracea	capitata		cabbage	FRA
157	212	lr	FRA010	396	Brassica	oleracea	capitata		cabbage	FRA
158	217	lr	GBR006	HRIGRU 7824	Brassica	oleracea	capitata		jersey cabbage	GBR
159	218	lr	GBR006	HRIGRU 12478	Brassica	oleracea	capitata		cabbage	TUR
160	219	lr	GBR006	HRIGRU 12477	Brassica	oleracea	capitata		cabbage	TUR
161	220	lr	GBR006	HRIGRU 11615	Brassica	oleracea	capitata		savoy cabbage	PRT
162	223	lr	GBR006	HRIGRU 7826	Brassica	oleracea	capitata		pickling cabbage	YUG
163	226	lr	GBR006	HRIGRU 9978	Brassica	oleracea	capitata		fodder cabbage	ESP
164	227	lr	GBR006	HRIGRU 12429	Brassica	oleracea	capitata		drumhead x pickling	GBR
165	229	lr	GBR006	HRIGRU 6568	Brassica	oleracea	capitata		white cabbage	EGY

166	231	lr	POL003	PL174885	Brassica	oleracea	capitata		cabbage	UKR
167	232	lr	POL003	PL174868	Brassica	oleracea	capitata		cabbage	POL
168	234	lr	YUG002	NS-K-1	Brassica	oleracea	capitata		white cabbage	SRB
169	235	lr	ESP026	AN-B-14	Brassica	oleracea	capitata alba		white cabbage	ESP
170	237	lr	ESP026	V-B-46	Brassica	oleracea	capitata alba		white cabbage	ESP
171	245	lr	NLD037	CGN07007	Brassica	oleracea	capitata alba		white cabbage	HUN
172	250	lr	NLD037	CGN11060	Brassica	oleracea	capitata alba		white cabbage	EGY
173	254	lr	NLD037	CGN14114	Brassica	oleracea	capitata alba		white cabbage	RUS
174	257	lr	NLD037	CGN23712	Brassica	oleracea	capitata alba		white cabbage	KGZ
175	259	lr	DEU146	BRA 1919	Brassica	oleracea	capitata rubra		red cabbage	ITA
176	260	lr	DEU146	BRA 915	Brassica	oleracea	capitata rubra		red cabbage	GBR
177	261	lr	NLD037	CGN18436	Brassica	oleracea	capitata rubra		red cabbage	DEU
178	262	lr	NLD037	CGN18438	Brassica	oleracea	capitata rubra		red cabbage	ROU
179	265	lr	DEU146	BRA 2722	Brassica	oleracea	capitata sabauda		savoy cabbage	ITA
180	266	lr	DEU146	K 9015	Brassica	oleracea	capitata sabauda		savoy cabbage	ITA
181	269	lr	NLD037	CGN07121	Brassica	oleracea	capitata sabauda		savoy cabbage	NLD
182	272	lr	NLD037	CGN18456	Brassica	oleracea	capitata sabauda		savoy cabbage	SUN
183	276	lr	DEU271	CR 2558	Brassica	oleracea	costata			ESP
184	282	lr	DEU146	BRA 1465	Brassica	oleracea	gongylodes		kohl rabi	GEO
185	285	lr	DEU146	BRA 1460	Brassica	oleracea	gongylodes		kohl rabi	GEO
186	289	lr	DEU146	BRA 1452	Brassica	oleracea	gongylodes		kohl rabi	DEU
187	293	lr	PRT001	06033-BPGV	Brassica	oleracea	oleracea			PRT
188	297	lr	GBR006	HRIGRU 9490	Brassica	oleracea	tronchuda		tronchuda cabbage	PRT
189	307	lr	GBR006	HRIGRU 11604	Brassica	oleracea	tronchuda		tronchuda cabbage	PRT
190	308	lr	GBR006	HRIGRU 11612	Brassica	oleracea	tronchuda		tronchuda cabbage	PRT

191	310	lr	DEU146	K 9403	Brassica	oleracea				FRA
192	311	lr	DEU146	K 8788	Brassica	oleracea				ESP
193	316	lr	ESP026	A-B-15	Brassica	oleracea				ESP
194	317	lr	ESP026	AN-B-11	Brassica	oleracea				ESP
195	340	lr	GBR006	HRIGRU 4520	Brassica	oleracea			cabbage x cauli	GRC
196	355	lr	GBR165	SASASC1726	Brassica	oleracea			shetland cabbage	GBR
197	356	lr	GBR165	SASASC1348	Brassica	oleracea			shetland cabbage	GBR
198	357	lr	GBR165	SASASC1365	Brassica	oleracea			shetland cabbage	GBR
199	361	lr	GRC005	2680	Brassica	oleracea				GRC
200	362	lr	GRC005	2687	Brassica	oleracea				GRC
201	364	lr	GRC005	2697	Brassica	oleracea				GRC
202	368	lr	POL003	PL174872	Brassica	oleracea				SVK
203	371	lr	PRT001	03631-BPGV	Brassica	oleracea			portuguese kale	PRT
204	373	lr	PRT001	03653-BPGV	Brassica	oleracea			portuguese kale	PRT
205	379	lr	PRT001	07435-BPGV	Brassica	oleracea			portuguese kale	PRT
206	382	lr	PRT001	02756-BPGV	Brassica	oleracea			portuguese kale	PRT
207	384	lr	PRT001	04031-BPGV	Brassica	oleracea			portuguese kale	PRT
208	408	lr	RUS001	502201734	Brassica	oleracea		capitata alba	white cabbage	RUS
209	409	lr	RUS001	502201826	Brassica	oleracea		capitata alba	white cabbage	RUS
210	410	lr	RUS001	502201894	Brassica	oleracea		capitata alba	white cabbage	RUS
211	413	lr	RUS001	502202043	Brassica	oleracea		capitata alba	white cabbage	RUS

212	415	lr	RUS001	502202047	Brassica	oleracea		capitata alba	white cabbage	RUS
213	416	lr	RUS001	502202048	Brassica	oleracea		capitata alba	white cabbage	RUS
214	417	lr	RUS001	502202053	Brassica	oleracea		capitata alba	white cabbage	RUS
215	418	lr	RUS001	502202086	Brassica	oleracea		capitata alba	white cabbage	RUS
216	421	lr	RUS001	502202113	Brassica	oleracea		capitata alba	white cabbage	RUS
217	422	lr	RUS001	502202120	Brassica	oleracea		capitata alba	white cabbage	RUS
218	427	lr	RUS001	502202357	Brassica	oleracea		capitata alba	white cabbage	RUS
219	429	lr	RUS001	502202452	Brassica	oleracea		capitata alba	white cabbage	RUS
220	430	lr	RUS001	502202504	Brassica	oleracea		capitata alba	white cabbage	RUS
221	434	lr	ROM007	ROM007-13708	Brassica	oleracea	capitata		cabbage	ROU
222	436	lr	ROM007	ROM007-14938	Brassica	oleracea				ROU
223	438	lr	ROM007	ROM007-4656	Brassica	oleracea			Wild cabbage	ROU
224	449	cwr	Plantbreeding	Bol2009-0076	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
225	450	cwr	Plantbreeding	Bol2009-0077	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
226	451	cwr	Plantbreeding	Bol2009-0078	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	Italy
227	452	cwr	Plantbreeding	Bol2009-0079	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	Spain
228	453	cwr	Plantbreeding	Bol2009-0080	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
229	454	cwr	Plantbreeding	Bol2009-0081	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
230	455	cwr	Plantbreeding	Bol2009-0082	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
231	456	cwr	Plantbreeding	Bol2009-0083	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	

232	457	cwr	Plantbreeding	Bol2009-0084	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
233	458	cwr	Plantbreeding	Bol2009-0085	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
234	459	cwr	Plantbreeding	Bol2009-0086	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
235	460	cwr	Plantbreeding	Bol2009-0087	Brassica	fruticulosa	Cirillo subsp fruticulosa		wild	
236	461	cwr	Plantbreeding	Bol2009-0088	Brassica	villosa			wild	B villosa CGN14116
237	462	cwr	Plantbreeding	Bol2009-0089	Brassica	incana			wild	B incana CGN18470
238	463	cwr	Plantbreeding	Bol2009-0090	Brassica	incana			wild	B incana CGN18471
239	464	cwr	Plantbreeding	Bol2009-0091	Brassica	montana			wild	B montana CGN18472
240	465	cwr	Plantbreeding	BOL2010-0437	Brassica			IPK BRA1810		

Appendix II MIAME/ Plant data

MIAME/PLANT frame work	**Experiment information**
Array Design Description	
Manufacturer	Affymetrix
Chip type	Arabidopsis Gene 1.0 ST array
No. of Probes	28,501 gene level probesets
Genome build	TAIR 10
Experimental design	
1. Plant Experiment Design	
A) Pooling of samples	
Number of plants in each pool	4 leaf discs
When pooled	12 weeks age
Genotype pooled	individual
Planted on the same day	Yes
B) Experimental design	
Number of blocks	4
Randomised between blocks	Yes
2. Plant sample used, extract preparation and labelling	
1) Biosource properties	
Germplasm Accession	15
Starting material	Seed
Development stage	12 weeks old (vegetative stage)
Organism part	Leaf tissue
Growth substrate	Multipurpose compost (HUMAX)
2) Biomaterial manipulations	
Growth environment	Control growth room
Light duration	16L: 8D
3) Environmental conditions	
Light intensity	250-280 μmmol.m-^{2}s-^{1}
Light source	Fluorescent lamps
Humidity	60 ± 10%
Watering conditions	Manually
Temperature	24 ± 3^{0}C
Pots	11 cm diameter
Growth/Control agents	None
Harvesting conditions	Growth room temperature
Treatment type	Aphid induction
Degree of stress	15 adult aphid, per clip cage
Stress duration	24 hrs
Isolation techniques	Leaf tissue under clip cage removed by scalpel and flash frozen in liquid nitrogen
Extraction Method	
Quantity extracted	100mg
Extraction source	Fresh sample
Extraction method	Invitrogen TRIzol and Qiagen RNeasy Plant Mini kit (see chapter 4)
Labelling	As per manufacturer's instructions (see chapter 4)