UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Aphid-Plant interactions and the possible role of an endosymbiont in aphid biotype development

by

Zacharias Hendrik Swanevelder

Submitted in partial fulfilment of the requirements for the degree

Philosophiae Doctor

In the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

August 2010

Supervisor: Prof A-M Botha-Oberholster

Co-supervisor: Dr E Venter

Believe is the gift of seeing His works all around you

Thank You for carrying me in those times I were unable to continue,

Thank You for the gifts of logic, science, and all the others You have so undeservingly bestowed upon me,

Thank You for supervisors, especially for their patience, during the completion of this study,

And

Thank You for a family and friends that You have given me in support while completing this task

For my family and friends,
Thank you for your support, patience and prayers

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Declaration

I, Zacharias Hendrik Swanevelder declare that the thesis, which I hereby submit for the degree, *Philosophiae Doctor* at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature: _____

Date: _____

ACKNOWLEDGEMENTS

My supervisors, professors and teachers help formed me into the scientist I am today. However, it all started with my parents. Thank you for answering all those silly questions when I was small, helping me with home work during my school education, supporting me during my university days and forming me into the free thinking individual I am today. It was through your example, guidance, love, support and a lot of tea and coffee during exams, which have led to the completion of this work and the fulfilment of a dream. Thanks for being great parents. I would also like to thank my sister. Thank you for your support during the completion of this study and for always being there as a friend, confidant and/or morel support. You and Carl's support helped me through in those bleak days before hand-in.

I would like to express my gratitude to my main supervisor, Prof A-M Botha-Oberholster, for giving me the opportunity to investigate the various hypotheses that constitute this study. Without her guidance, support and the various discussions we had on the different hypotheses, I would not have been able to successfully complete this research project. Under her guidance I've not only developed as a free thinking research scientist, but also as an individual. Though this is only a few words that express my gratitude, I would like her to know that the impact she had my career will continue throughout my life. Thank you, Anna-Maria.

I would also like to express my thanks to my co-supervisor, Dr. E. Venter for his support and guidance during the completion of this study. Eduard, though I'm sure I've caused you and Anna-Maria a lot of frustration during this research project, I just want to express my gratitude to you for your part in helping me to develop into the researcher that I'm today. Thank you for the various scientific discussions we had during my education.

And to all my friends: Thanks for the support, pep-talks, willing (and sometimes unwilling) ears, drinks and shoulders during this study, without those this study would probably have grinded to a halt. My thanks to a colleague and friend, Dr. A.K.J. Surridge, for friendship, help and facilities that were used during the denaturing gradient gel electrophoresis (DGGE) part of the study.

I would also like to thank my Bioinformatics professor, Prof F. Joubert at the Bioinformatics Unit of the University of Pretoria, for all his help during the execution of the scripts and the use of their servers during the analyses.

This research would not have been possible without various collaborators. I would like to express my thanks to: Prof N.L.V. Lapitan from Colorado State University, Dr. V.L. Tolmay of the

ARC-Small Grains Institute in Bethlehem, South Africa and Dr. G.J. Puterka of the USDA-ARS in Stillwater Oklahoma, for *Diuraphis* samples, and Mr. I. Millar of the South African National Collection of Insects (SANC, PPRI) situated in Pretoria (South Africa) for the other members of the Aphididae.

Finally, I would also like to thank the various funding bodies that allowed me to complete this study: the National Research Foundation of South Africa (NRF) for awarding me with a Prestigious PhD Bursary, the University of Pretoria for awarding me a bursary for my tuition, and the funding of the project by the NRF, THRIP (Technology and Human Resources for Industry Programme, South Africa) and Winter Cereal Trust.

ABBREVIATIONS

ΔG	Gibbs' free energy
°C	degrees Celsius
µg	microgram
µL	microliter
µM	micromolar
5' RACE	5' rapid amplification of cDNA ends
5' UTR	5' untranslated transcription region/leader sequence
AA	ascorbate
AA_a	apoplastic ascorbate
AO	ascorbate oxidase
APX	ascorbate peroxidase
AS	anthranilate synthase
avr	avirulence
BA	Baysian analysis
bp	base pair
CAM	calmodulin
CAT	catalase
CC	coiled-coil
CI	consistency index
COI	mitochondrial cytochrome oxidase subunit I
CPGR	Centre for Proteomic & Genomic Research
CW	cell walls
DGGE	denaturing gradient gel electrophoresis
DHA	dehydroascorbate
DHAR	DHA reductase
DKG	2,3-diketo-l-gulonic acid
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotides
FDR	false discovery rate
GCRMA	Robust Multi-array Analysis with probe GC-content
GPX	glutathione peroxidase
GR	glutathione reductase
GSH	glutathione
GSH	reduced glutathione
GSSH	oxidised glutathione
h	hour(s)
hpi	hours post infestation
HR	hypersensitive response
IWF	intracellular washing fluid

JA	jasmonic acid
JAs	jasmonates
kb	kilobases
kcal mol^{-1}	kilocalories per mole
kD	kiloDalton
LDF	linear discriminant function
limma	linear models for microarrays
LMM	linear mixed models
LRR	leucine rich repeat
LSU	Rubisco large subunit
LZ	leucine zipper
MAPK	mitogen-activated protein kinases
MAS5.0	Affymetrix microarray suite 5
MDA	monodehydroascorbate
MeJA	methyl jasmonate or jasmonic acid
min	minute or minutes
MM	mismatch probe pairs
Mol %	mole percentage
mRNA	messenger ribonucleic acid
NB-ARC	nucleotide binding apoptosis *R*-gene and *CED-4* like domain
NBD	nucleotide binding domain
NBS-LRR	nucleotide binding site leucine rich repeat
ng	nanogram
NIL	near isogenic lines
nL h^{-1} g^{-1}	nano liters per hour per gram
NO	nitric oxide
ORF	open reading frame
PAL	phenylalanine ammonia lyase
PCD	programmed cell death
PCR	polymerase chain reaction
PE	pectin esterase
Phe	phenylalanine
PK	protein kinases
pleuABCD	leucine plasmid
PLM	Probe Level Models
PM	perfect match probe pairs
PPO	polyphenol oxidases or catechol oxidase
PR-genes	pathogenesis related genes
PSII	photosystem II
pTrpEG	tryptophan biosynthesis plasmid
PX	peroxidase
qPCR	quantitative polymerase chain reaction

R	resistance
RBS	ribosomal binding sites
RI	retention index
RMA	Robust Multichip Average
RNA	ribonucleic acid
ROS	reactive oxygen species
RT-qPCR	quantitative real-time PCR
RWA	Russian wheat aphid
SA	salicylic acid
SA biotype	South African biotype
SAM	South African mutant biotype
SAR	systemic acquired resistance
SCOPE	Suite for Computational identification Of Promoter Elements
SE	sieve elements
sec	second(s)
SOD	superoxide dismutase
SSU	Rubisco small subunit
t/ha	ton per hectare
TBR	tree-bisection-reconnection
TD	Touchdown
TF	transcription factor
TFBs	transcription factor binding site
U	enzymatic unit
US$	United States dollar
USA	United States of America
v/v	volume per volume
VSN	Variance Stabilisation
σ^{32}	sigma factor *rpoH*
σ^{70}	sigma factor *rpoD*

TABLE OF CONTENTS

DECLARATION 3

ACKNOWLEDGEMENTS 4

ABBREVIATIONS 6

TABLE OF CONTENTS 9

SUMMARY 12

LIST OF FIGURES 14

LIST OF TABLES 29

PREFACE 32

CHAPTER 1 35

INTRODUCTION 35

CHAPTER 2 39

APHID-PLANT-ENDOSYMBIONT INTERACTION: THE RUSSIAN WHEAT APHID, ITS HOSTS AND ENDOSYMBIONT, *BUCHNERA APHIDICOLA* 39

DIURAPHIS NOXIA (APHIDIDAE: MACROSIPHINI) 40
Origin 40
Description 40
Biology 41
Genetics 41
Biotypes 41
Hosts 42
Symptoms 43
Economic losses 44
Control 44
APHIDS AND THEIR ENDOSYMBIONT, *BUCHNERA APHIDICOLA* 45
Buchnera aphidicola's role in aphid nutrition 45
Buchnera aphidicola-RWA interaction 47
TRITICUM AESTIVUM AS APHID HOST 48
Resistant wheat cultivars 48
PLANT RESISTANCE AND DEFENCE MECHANISMS 50
The origin and detection of aphid elicitors 52
Plant resistance genes 54
Redox signalling and plant defence 55
Transduction and defence pathways 56
APHID-PLANT INTERACTION 58
Aphid probing and feeding 58
Evading the host defence 59
Salivary enzymes: Neutralising potential signals and suppressing defences 61
Salivary enzymes of the RWA 64
RWA-WHEAT INTERACTION 66
CONCLUSION 75
REFERENCES 76

CHAPTER 3 99

LIMITED ENDOSYMBIONT VARIATION IN *DIURAPHIS NOXIA* (HEMIPTERA: APHIDIDAE) BIOTYPES FROM THE USA AND SOUTH AFRICA 99

INTRODUCTION 100
MATERIALS AND METHODS 102
Diuraphis species and biotypes 102
Biotypic endosymbiont investigation 102
Buchnera aphidicola sequence variation amongst biotypes 103
Structural analysis 103
Plasmid copy numbers 103
RESULTS 104
Biotypic endosymbiont investigation 104
Buchnera aphidicola sequence variation amongst biotypes 104
Primary and secondary structural analysis 105
Plasmid copy numbers 109
DISCUSSION 110
Biotypic endosymbiont investigation 110
Buchnera aphidicola sequence variation amongst biotypes 110
Primary and secondary structural analysis 111
Plasmid copy numbers 112
CONCLUSIONS 113
ACKNOWLEDGEMENTS 114
REFERENCES 114

CHAPTER 4 122

SYMBIOSIS: VARIATION IN *BUCHNERA APHIDICOLA*'S LEUCINE PLASMID MAY CONFER ADVANTAGE TO RUSSIAN WHEAT APHID BIOTYPE 122

INTRODUCTION 123
MATERIALS AND METHODS 124
Aphids 124
DNA and RNA extraction 125
Leader sequence determination 125
The inverted repeat region in the Aphididae 126
Software analysis 126
RT-qPCR 127
RESULTS 128
The leader sequence 128
The inverted repeat in the Aphididae 130
RT-qPCR 133
DISCUSSION 135
CONCLUSION 137
ACKNOWLEDGEMENTS 138
REFERENCES 138

CHAPTER 5 145

JUST HOW DO AFFYMETRIX NORMALIZATION METHODS COMPARE? STATISTICS CONTEMPLATE BIOLOGY 145

INTRODUCTION 146
MATERIALS AND METHODS 147
Experimental design 147
Aphids and plant material 148
Data analysis 149
RESULTS 152
Quality control of slides, background correction and normalization 152
Normalization and background correction methods 154
Identifying false positives 159

DISCUSSION 165
ACKNOWLEDGEMENTS 168
REFERENCES 169

CHAPTER 6 172

CONCLUSION 172

APPENDICES AND SUPPLEMENTARY DATA 176

APPENDIX – CHAPTER 3 177

Sequences submitted to Genbank 183

APPENDIX – CHAPTER 4 213

Sequences submitted to Genbank 213

APPENDIX – CHAPTER 5 215

Script 1 Data in, normalization & quality check 216
Script 2 Quality control analysis 220
Script 3 Getting differentially expressed genes 223
Target.txt 226
Convert 226
Data visualization before normalization and background correcting 227
Data visualization after background correcting and normalization 237

CURRICULUM VITAE 275

SUMMARY

Aphid-Plant interactions and the possible role of an endosymbiont in aphid biotype development

Zacharias Hendrik Swanevelder

Department of Genetics
University of Pretoria, Pretoria
Supervisor: Prof A-M Botha-Oberholster
Co-supervisor: Dr E Venter

Philosophiae Doctor

Diuraphis noxia Kurdjumov (Hemiptera: Aphididae: Macrosiphini) is a major agricultural pest that causes extensive economic losses to the wheat and barley industries. Resistant cultivars were relatively successful in controlling this pest until the recent development of new *D. noxia* (Russian wheat aphid, RWA) biotypes. The aim was to investigate the role of the aphid endosymbiont, *Buchnera aphidicola*, in the RWA-host interaction. It was hypothesized that variations in the endosymbiont's key essential amino acid biosynthetic pathway genes, their copy numbers, and/or expression levels, maybe a determining factor influence the RWA's success in the aphid-host interaction. Aphid symbiont species content, key essential amino acid biosynthetic gene variation, plasmid copy numbers and expression levels of ten different RWA biotypes were determined, using DGGE, RT-PCR, RT-qPCR, 5'-RACE and sequencing. The RWA biotypes were shown to be monosymbiotic, with plasmid copy numbers varying between biotypes. Only a single CCC-insert in a non-coding region of the leucine plasmid differed between the biotypes. Similar variations were identified in the family Aphididae, suggesting a regulatory function for this region. The presence of this CCC-insert in a plasmid led to an increase in the leader sequence length of the *leuA* gene. The insert may also have a functional role through gene regulation, since it increased the expression levels of subsequent genes (*leuA* and *leuB*). An endosymbiont that upgrade the host's diet with the required essential amino acids will be beneficial to RWAs when feeding on resistant wheat cultivars

as it will enhance aphid fitness. This suggests selective pressure of resistant wheat cultivars on the aphid, *i.e.* the incapability to change resistant cultivar essential amino acid content, could select for individuals with beneficial endosymbionts. *B. aphidicola* could therefore play a role in the development of RWA biotypes.

The influences that statistical normalization methods have on the final identification of differentially regulated Affymetrix probe sets in RWA-plant interactions were also investigated. The hypothesis was that a subset of the probe sets determined as differentially regulated would be consistent, regardless of the normalization and background method utilized, if all the other analyses are kept constant. This subset would be normalization-method-independent. The data of two Affymetrix RWA-plant interaction experiments were analyzed with five different normalization and background correcting methods and at three different confidence levels, with the results subjected to FDR and FWER correction algorithms. The results showed that on average a third of the regulated genes were only selected after normalization by a single method and that the total number of genes deemed regulated was highly normalization method dependent. Normalization-method-biases could also not be countered by increased confidence levels and these biases eventually determined the probe sets deemed differentially regulated, even after FDR and FWER corrections. Both these strategies actually increased normalization-method-biases and these could only be corrected by using multiple normalization methods to identify the normalization-method-biases-independent probe set subset.

LIST OF FIGURES

Figure 2.1 The South African (SA) biotype on the susceptible cultivar Scheepers, and the SA mutant (SAM) on the resistant wheat cultivar TugelaDN. No morphological differences are apparent. Photos taken under different magnifications.

Figure 2.2 (A) Heavily infested wheat plants showing longitudinal streaking and (B) tightly inward curling of the leaf edges (leaf rolling).

Figure 2.3 The interaction between aphids (italics and dashed lines) and the defence responses of plants (normal case and unbroken lines). Aphids activate similar signal transduction pathways as pathogens, which utilize the salicylic acid (SA), reactive oxygen species (ROS) and jasmonic acid (JA)/ethylene dependent pathways to activate gene expression in plants. Arrows indicate the interaction and red hexagons the inhibitions. Abbreviations: CW, cell wall; HR, hypersensitive response; NO, nitric oxide; PPO, polyphenol oxidase; PR-genes, pathogen related genes (Adapted from Miles 1999; Walling 2000).

Figure 2.4 Biotic elicitation in plant-aphid interaction is thought to be similar to that of microbial pathogens. In microbial-pathogen interaction an elicitor binds to plasma membrane/CW receptor and activates protein phosphorylation, protein kinases (PK) and mitogen-activated PK (MAPK), as well as G-proteins (intracellular proteins that interact/function with receptors to regulate various enzymes and ion channels). Ion transporters are activated resulting in changes of ion fluxes (calcium and hydrogen ion influx increases). Calcium ions bind to calmodulin (CAM, non-enzymatic intracellular Ca^{2+}-binding proteins) which then binds to other proteins for regulation, resulting in the expression of defence genes. Secondary messengers are also activated, resulting in calcium release and the activation of various pathways. The cytoplasm acidifies as the result of NADPH oxidase activation, decrease in membrane polarization and the inactivation of H^{+}-ATPase. ROS activation and PR-protein expression can cause HR cell death at infection site or systemic acquired resistance (SAR) (compiled from Radman *et al.* 2003).

Figure 2.5 An overview of aphid feeding (green) and the responses of the host. Antioxidants in the plant regulate the redox condition of phenolic compounds in the cells by keeping them in a reduced

form under normal cellular conditions, while regulating oxidation rates when wounded. Aphid salivary enzymes (green) alter the redox poise of the cell, thereby enhancing the oxidation of toxic phenolics/quinones into non-toxic phenol-protein conjugates and polymers. Refer to *Figures 2.6 & 2.7*. Compiled from Miles & Oertli 1993; Jarabak *et al.* 1997 and Ni *et al.* 2000.

Figure 2.6 Plant defence responses during unsuccessful aphid infestation. Increases in expression or substrate of phenylalanine ammonia lyase (PAL) and polyphenol oxidase (PPO) lead to increases in quinones. The antioxidant regeneration system (Antiox Regen) regulates the conversion from diphenols to quinone, thereby keeping the supply of toxic compounds steady. When antioxidants are used faster than they can be replenished (X), quinones reach the HR level. Phe, phenylalanine. Also see *Figure 2.7*. Compiled from Miles & Oertli 1993 and Miles 1999.

Figure 2.7 Known RWA salivary enzymes and their role in upsetting cellular and apoplastic redox poise in the host. In healthy tissue phenolics are kept in a reduced state by antioxidants and the continual reductase renewal. During wounding, membranes rupture and phenolics are released/produced. Quinone redox-cycling start (orange) and produce O_2^- and H_2O_2 resulting in ROS and HR. Quinone reductase transcription is up-regulated during this HR response, producing more ROS species. Toxic semi- and hydroquinones are also being recycled, thereby controlling the oxidation of quinones to non-toxic substances. Ingestion of un-oxidised phenolics/quinones, in the absence of the renewable antioxidants, autoxidise with proteins in the gut to form toxic substances. Salivary oxidoreductases shift the redox poise of the apoplast/cytosol and may even detoxify phloem sap *en route* to the gut. $AscH_2$, ascorbate; AO, ascorbate oxidase; APX, ascorbate peroxidase; CAT, catalase; DHA, dehydroascorbate; DHAR, DHA reductase; DKG, 2,3-diketo-l-gulonic acid; GR, glutathione reductase; GSH, reduced glutathione; GSSH, oxidised glutathione; GPX, glutathione peroxidase MDA, monodehydroascorbate; PPO, polyphenol oxidase (catechol oxidase); SOD, superoxide dismutase. (Miles & Oertli 1993; Jarabak *et al.* 1997; Miyake & Kurata 1998; Roginsky *et al.* 1999; Ni *et al.* 2000; Cape *et al.* 2006; Foyer & Noctor 2009); IUBMB Enzyme Nomenclature EC 1.6.5.5; Pfam PF01095; www.brenda-enzymes.org (EC3.1.1.11 – pectinesterase).

Figure 3.1 The inverted repeat region between *repA2* and *leuA* on leucine plasmids of the Aphididae. The end of *repA2* and the start of *leuA* are indicated in grey, with possible start codons indicated in red. Abbreviations: ‘pB’ indicates the leucine plasmid of *B. aphidicola* which is then

followed by the aphid species, 'Dn' and biotype 'SAM': pBDnSAM, *D. noxia* biotype SAM; pBDnUSA1, *D. noxia* biotype USA1; pBDn, *D. noxia* (Genbank); pBAp, *Acyrthosiphon pisum*; pBMd, *Metapolophium dirhodum*; pBMr, *Macrosiphum rosae*; pBAs, *Aulacorthum solani*; pBUs, *Uroleucon sonchi*; pBRc, *Rhopalosiphum cerasifoliae*; pBSg, *Schizaphis graminum*; pBRi, *R. insertum*; pBRp, *R. padi*; pBPp, *Ptercomma populeum*.

Figure 3.2 Minimum free energy diagrams for the inverted repeat region up to the start codon of *leuA* of *D. noxia* biotypes as predicted with Quickfold (DNA and temperature 25 °C, DINAMelt Server) (Markham & Zuker 2005), with (*Figure 3.2A*) and without (*Figure 3.2B*) CCC-insert. Similar structures and energy values were also obtained for predicted RNA structures of the same region. In both cases the CCC-insert resulted in an increase of ~5 kcal mol^{-1} free energy between the predicted structures. All free energy values (dG) are given as kcal mol^{-1}.

Figure 3.3 The region between *repA2* and *leuA* indicating predicted promoter and *Rho* independent terminator regions of RWA biotypes without (*Figure 3.3A*) and with (*Figure 3.3B*) the CCC-insert. The CCC-insert shifts the predicted *Rho* independent terminator from the coding to the non-coding strand (*B*). This increases the predicted free energy value of the terminator structure, suggesting a less stable terminator.

Figure 3.4 The ratio *leuB:trpB* was used to calculate the copy number of the different biotypes. The mean copy numbers for *leuB* (insert A, n=4) and of *trpB* (insert B, n=4) were determined for the different biotypes and then used to calculate the *leuB:trpB* ratio. South African biotypes had lower copy numbers than their USA counterparts which is supported by literature. The average across all the biotypes was 0.9 copies per chromosomal gene.

Figure 3.5 Phylogenetic relationships between tribal representatives of the subfamily Aphidinae. The first two phylogenies are an abbreviated representation of a combined nuclear and mitochondrial dataset of *tRNA, COII* and *TEF1α*, from von Dohlen *et al.* (2006). Bold lines indicate members of the tribe Macrosiphini. Species names are given where data are available for the phylogenies and genus names when different species of the same genus were used. An asterisk indicates absent genera.

Figure 4.1 *Buchnera aphidicola* of *D. noxia* with the leader regions (5' UTR) indicated by a blue line for accessions with an upstream CCC-insert and a red line for those without the CCC-insert. Predicted promoters, sigma factor binding sites, AU-rich region and ribosomal binding sites (RBS), including possible start sites, are indicated. Predicted secondary structures (Quickfold, DINAMelt Server, 25 °C) for the leader sequences are given in the inserts A and B. Six structures were predicted for the short 5' UTR (insert A) and three for the longer leader region of the CCC-insert containing *Buchnera* (insert B). The longer leader sequence produced predicted structures that are approximately -2.3 kcal mol^{-1} more stable than those of the shorter 5' UTR. These structures could increase the half life of the RNA molecules by preventing RNA degradation.

Figure 4.2 The region between the *repA2* and *leuA* on the leucine plasmid of *B. aphidicola* accessions originating from various aphid hosts. Variations within a species is indicated by bold red lettering, variations within a genus by bold blue lettering and the start site of the 5' UTR leader sequence by bold gold lettering. The genes *repA2* and *leuA* are shaded in grey with their end and start codons, respectively, in bold. The inverted repeat regions of the family Aphididae are underlined on the consensus sequence and the regions indicated by black arrows. The conserved core region (red double arrows) of the inverted repeat region (black) includes predicted promoters, transcription factor binding sites and start codons (see legend). The Gibbs' free energy (ΔG) values for the stemloop structures were calculated with Quickfold (DINAMelt Server, at 25 °C in kcal mol^{-1}). Predicted stemloop structures (Loop) is given as perfect (P) or imperfect (IP), with the number of additional loops indicated (in brackets). The number of suboptimal T-G pairs (TG) predicted is also listed. RSA-tools' Consensus and Convert-matrix programs were used to obtain the core regions of the stemloop structures that are conserved within the family.

Figure 4.3 The relative gene expression levels per plasmid copy of *leuA* and *leuB* after normalization with *rpoB.*

Figure 5.1 A 2 × 2 fractional design used in the analysis of the two experiments, with (A) representing the 18 slide experimental comparison and (B) the 12 slide comparison. Abbreviations (A): RWA1GR, *Dn7* resistant cultivar 94M370 infested with US RWA biotype 1; RWA1GS, the susceptible cultivar Gamtoos infested with US RWA biotype 1; RWA2GR, the *Dn7* resistant cultivar 94M370 infested with the US RWA biotype 2; RWA2GS, susceptible Gamtoos line infested

by US RWA biotype 2; GR, the resistant cultivar 94M370 control that was un-infested and GS, the susceptible Gamtoos control, un-infested. Abbreviations (B): Tug, the susceptible cultivar Tugela; TuD, the resistant NIL TugelaDN; Tu2, the resistant NIL Tugela *Dn2* and Tu5, the resistant NIL Tugela *Dn5* - all these lines were infested with the South African RWA biotype.

Figure 5.2 A layout of the analysis conducted on each experiment, including some of the output folders and filenames (with examples), as produced by the different R scripts. Files and folders produced during the running of script 1 are represented by a green block, with script 2 and 3 represented by a blue and red block respectively. Variable portions of filenames are given in italics. Details on the input files and the output files/folders are given within the header of each script (*Table Appx 5.1*).

Figure 5.3 The work flow of the data analysis done on the differentially regulated genes, before and after Bonferroni (Bon) and Benjamini-Hochberg (Ben or BenHoch) corrections for 95, 99 and 99.9 % confidences. The same procedures followed for Benjamini-Hochberg analysis were done for Bonferroni. These analyses were repeated for each of the different confidence levels. Files added after the analyses are indicated in red. The summary files are given in bold red.

Figure 5.4 Box plots of the $\log_2$ (PM) values, before and after background correction and normalization were done on the 12 slide dataset with the different statistical approaches. All the sample slides occur in the same order in the various graphs in their corresponding triplicate slides (Tugela, TugelaDN, Tugela *Dn2* and Tugela *Dn5*) but that the Y-axis differs in accordance to the method used.

Figure 5.5 *(previous page)* Normalization method contributions, according to the number of genes differentially up or down regulated, were grouped into the largest, second largest, *etc*. group for each specific slide comparison. The number of times a specific method, expressed as total percentage contribution across all slides, were then calculated and presented here, *i.e.* how many times does a method produce the most, second most, *etc*. number of genes for a specific confidence level, across all the slide comparisons. These values were calculated before FDR/FWER correction ((A), 12 slides and (D), 18 slides) and after FWER/FDR corrections (Bonferroni: (B), 12 slides and (E), 18 slides; Benjamini-Hochberg: (C), 12 slides and (F), 18 slides). The total percentage contribution (y-

axis) for that method is less than a 100 % if there are zero genes identified as differentially regulated for one or more slide comparisons. Slide comparisons sometimes delivered the same number of differentially regulated genes for two normalization methods. In these cases, both methods were placed in the same group. Following normalization method identification, "U" indicates up and "D" down regulated comparisons, followed by the confidence level (95, 99 and 99.9 %).

Figure 5.6 The mean number of genes/probe sets found to be differentially regulated as found by the LMM and BA after normalization with the five methods and at 95, 99 and 99.9 % confidence, before/after FDR/FWER corrections. The standard error over all the slide comparisons is indicated by the error bars. (A) and (B) represents the down and up regulated 12 slide experiment respectively, while (C) and (D) represent the down and up regulated genes of the 18 slide experiment, respectively.

Figure 5.7 The averages of the total percentages of geneIDs/probe sets predicted to be differentially regulated, after normalization at different confidences (95, 99, 99.9 %) and before/after FDR/FWER corrections. The standard errors were calculated on the percentage contribution of each slide comparison to a 1, 2, 3, 4, or 5 method occurrence that was initially used to determine the mean sum (average total) percentage for each specific confidence (95, 99, 99.9 %) and before/after FDR/FWER correction. (A)-(C) represent the 12 slide experiment, with (A) the total differentially regulated genes, *i.e.* the combined down and up regulated percentages, while (B) and (C) represent the down and up regulated genes respectively. Similarly (D)-(F) represent the 18 slide experiment, with (D) the combined, (E) the down and (F) the up regulated differentially expressed gene.

Figure Appx 3.1 Phylogeny of the biotypes to their *16S rDNA* DGGE BLAST results (Genbank). Parsimony topology and branch lengths are shown. Bootstrap percentages from parsimony (70 % and greater, 1000 bootstrap replications) are shown above and Bayesian posterior probability values below branches. The third heuristic tree (out of 23) is presented here. The phylogram had a tree length of 93 and was based on 74 parsimony informative characters with a consistency index of 0.3647 and a retention index of 0.7164. *Escherichia coli* was used as outgroup to root the tree. Branch labels show the Genbank accession, followed by the aphid species name except for the outgroup. Tribes and families for the hosts obtained from Genbank accessions were indicated for the Aphididae. All phylogenetic analyses were done with PAUP 4.0b10 (Swofford 2002) and Bayesian

analysis with MrBayes 3.1.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Multiple sequence alignments were done with ClustalW version 2 (Larkin *et al.* 2007) and manually evaluated before further analysis. Ambiguous characters and uninformative nucleotides were excluded from data prior to analysis and all characters were re-weighted to the consistency index. The best statistical model of DNA evolution for each dataset was determined using modeltest 3.7 (Posada & Crandall 1998) and used in subsequent analysis. Heuristic searches using random sequence additions were performed with the tree-bisection-reconnection (TBR) branch-swapping algorithm and MaxTrees set to auto increase. Phylogenetic signal, *i.e.* consistency index (CI) and retention index (RI), was assessed by evaluating the tree length distributions in each dataset after 100 random generated trees. Only groups with a 70 % or more support were retained in bootstrap analyses over a 1000 replicates. MrBayes utilizes a Metropolis-coupled Markov Chain Monte Carlo search algorithm (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). The general time reversal (GTR) evolutionary model was selected with codon site-specific rate variation. One million generations were run for each dataset, using one cold and 3 heated chains. Bayesian inference posterior probabilities were calculated after the appropriate burn in was determined.

Figure Appx 3.2 Differences between the biotypes (FJ705299-FJ705308) and the Genbank sequences (AF041837, NC001911) for the *pleuABCD* plasmid. The CCC-insert only occurred in the biotypes SA, SAM, USA3 and USA7. A single nucleotide insert, a T-insert, between *leuA* and *leuB* genes, changed *leuB* into the same ORF as the *leuA* gene. The new *leuA-leuB* ORF is indicated in blue. The mutations notation used here indicate the nucleotide of the original Genbank sequences, a slash, followed by the changed nucleotide from sequencing results.

Figure Appx 5.1 Images of the scanned Affymetrix slides after hybridization with the different treatments. The images were produced from the *CEL* files. The 12 slides images were the originally produced *CEL* images, while the 18 slides images are the $\log_2$ derivatives of the original 18 slides images. Both types were produced by script 1 and saved in */R results/Quality control raw data/Slide Images raw/*. The filenames for the 12 slides are: (A) Tug1_raw.jpg, (B) Tug2_raw.jpg, (C) Tug3_-raw.jpg, (D) TugDn_1_raw.jpg, (E) TugDn_2_raw.jpg, (F) TugDn_3_raw.jpg, (G) TugDn2_1_-raw.jpg, (H) TugDn2_2_raw.jpg, (I) TugDn2_3_raw.jpg, (J) TugDn5_1_raw.jpg, (K) TugDn5_2_-raw.jpg, and (L) TugDn5_3_raw.jpg.

The $\log_2$ derivatives filenames for the 18 slides are: (A) RWA1_GS_3_log2_raw.jpg, (B) RWA1_GS_2_log2_raw.jpg, (C) RWA1_GS_1_log2_raw.jpg, (D) RWA2_GS_1_log2_raw.jpg, (E) RWA2_GS_2_log2_raw.jpg, (F) RWA2_GS_3_log2_raw.jpg, (G) RWA1_GR_1_log2_raw.jpg, (H) RWA1_GR_2_log2_raw.jpg, (I) RWA1_GR_3_log2_raw.jpg, (J) RWA2_GR_1_log2_raw.jpg, (K) RWA2_GR_2_log2_raw.jpg, (L) RWA2_GR_3_log2_raw.jpg, (M) Gam_S_1_log2_raw.jpg, (N) Gam_S_2_log2_raw.jpg, (O) Gam_S_3_log2_raw.jpg, (P) Gam_R_1_log2_raw.jpg, (Q) Gam_R_2_log2_raw.jpg, (R) Gam_R_3_log2_raw.jpg.

Figure Appx 5.2 Summary figures, including histograms and boxplots, of all the slide data before normalization for the 12 slide and 18 slide experiment. These figures were produced with script 1 and saved as *Summary of all data.jpg.*

Figure Appx 5.3 The $\log_2$ expression comparisons of the raw, pre-normalized data within treatments. The 12 slides files are: (A) Tug1_vs_Tug2_log2.jpg, (B) Tug1_vs_Tug3_log2.jpg, (C) Tug2_vs_Tug3_log2.jpg, (D) TugDn_1_vs_TugDn_2_log2.jpg, (E) TugDn_1_vs_TugDn_3_log2.-jpg, (F) TugDn_2_vs_TugDn_3_log2.jpg, (G) TugDn2_1_vs_TugDn2_2_log2.jpg, (H) TugDn2_-1_vs_TugDn2_3_log2.jpg, (I) TugDn2_2_vs_TugDn2_3_log2.jpg, (J) TugDn5_1_vs_TugDn5_2_-log2.jpg, (K) TugDn5_1_vs_TugDn5_3_log2.jpg, (L) TugDn5_2_vs_TugDn5_3_log2.jpg.

The 18 slides files are: (A) RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (B) RWA2_GS_1_-vs_RWA2_GS_3_log2.jpg, (C) RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (D) RWA2_GR_1_vs_-RWA2_GR_2_log2.jpg, (E) RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (F) RWA2_GR_2_-vs_RWA2_GR_3_log2.jpg, (G) RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (H) RWA1_GS_3_-vs_RWA1_GS_1_log2.jpg, (I) RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (J) RWA1_GR_1_-vs_RWA1_GR_2_log2.jpg, (K) RWA1_GR_1_vs_RWA1_GR_3_log2.jpg, (L) RWA1_GR_2_-vs_RWA1_GR_3_log2.jpg, (M) Gam_S_1_vs_Gam_S_2_log2.jpg, (N) Gam_S_1_vs_Gam_S_3_-log2.jpg, (O) Gam_S_2_vs_Gam_S_3_log2.jpg, (P) Gam_R_1_vs_Gam_R_2_log2.jpg, (Q) Gam_-R_1_vs_Gam_R_3_log2.jpg, (R) Gam_R_2_vs_Gam_R_3_log2.jpg.

Figure Appx 5.4 MA-plots of the raw, pre-normalized slide comparisons within treatments. The 12 slides files are: (A) Tug1_vs_Tug2_raw.jpg, (B) Tug1_vs_Tug3_raw.jpg, (C) Tug2_vs_Tug3_-raw.jpg, (D) TugDn_1_vs_TugDn_2_raw.jpg, (E) TugDn_1_vs_TugDn_3_raw.jpg, (F) TugDn_2_-vs_TugDn_3_raw.jpg, (G) TugDn2_1_vs_TugDn2_2_raw.jpg, (H) TugDn2_1_vs_TugDn2_3_-

raw.jpg, (I) TugDn2_2_vs_TugDn2_3_raw.jpg, (J) TugDn5_1_vs_TugDn5_2_raw.jpg, (K) TugDn5_1_vs_TugDn5_3_raw.jpg, (L) TugDn5_2_vs_TugDn5_ 3_raw.jpg.

The 18 slides files are: (A) RWA2_GS_1_vs_RWA2_GS_2_raw.jpg, (B) RWA2_GS_1_vs_-RWA2_GS_3_raw.jpg, (C) RWA2_GS_2_vs_RWA2_GS_3_raw.jpg, (D) RWA2_GR_1_vs_-RWA2_GR_2_raw.jpg, (E) RWA2_GR_1_vs_RWA2_GR_3_raw.jpg, (F) RWA2_GR_2_vs_-RWA2_GR_3_raw.jpg, (G) RWA1_GR_1_vs_RWA1_GR_2_raw.jpg, (H) RWA1_GR_1_vs_-RWA1_GR_3_raw.jpg, (I) RWA1_GR_2_vs_RWA1_GR_3_raw.jpg, (J) RWA1_GS_2_vs_-RWA1_GS_1_raw.jpg, (K) RWA1_GS_3_vs_RWA1_GS_1_raw.jpg, (L) RWA1_GS_3_vs_-RWA1_GS_2_raw.jpg, (M) Gam_S_1_vs_Gam_S_2_raw.jpg, (N) Gam_S_1_vs_Gam_S_3_-raw.jpg, (O) Gam_S_2_vs_Gam_S_3_raw.jpg, (P) Gam_R_1_vs_Gam_R_2_raw.jpg, (Q) Gam_R_-1_vs_Gam_R_3_raw.jpg, (R) Gam_R_2_vs_Gam_ R_3_raw.jpg.

Figure Appx 5.5 Histograms of the slides before ((A) Histogram_rawdata.jpg) and after normalization ((B) Histogram_esetVSN.jpg, (C) Histogram_esetRMA.jpg, (D) Histogram_esetMAS.jpg, (E) Histogram_eset GCRMA.jpg) for the 12 Slide and 18 Slide experiments.

Figure Appx 5.6 Box plots of the $\log_2$(PM) values, before ((A) Boxplot_rawdata.jpg) and after normalization ((B) Boxplot_esetVSN.jpg, (C) Boxplot_esetRMA.jpg, (D) Boxplot_esetPLM.jpg, (E) Boxplot_esetMAS.jpg, (F) Boxplot_esetGCRMA.jpg) for the 12 Slide (see *Figure 5.4*) and the 18 Slide experiment using different statistical approaches. All the sample slides occur in the same order in the various graphs in their corresponding triplicate slides, and are: RWA1_GS_3, RWA1_GS_2, RWA1_GS_1, RWA2_GS_1, RWA2_GS_2, RWA2_GS_3, RWA1_GR_1, RWA1_GR_2, RWA1_GR_3, RWA2_GR_1, RWA2_GR_2, RWA2_GR_3, Gam_S_1, Gam_S_2, Gam_S_3, Gam_R_1, Gam_R_2, Gam_R_3. The Y-axis differs in accordance to the method used.

Figure Appx 5.7 The $\log_2$ expression slide comparison plots for all the normalized datasets. The raw/pre-normalized dataset is also produced and included for comparison. The files are written to the folder */R results/Graphs of raw and normalized data/Log2 plots/*. The 12 Slides files are: (A) rawdata_Tug1_vs_Tug2_log2.jpg, (B) rawdata_Tug1_vs_Tug3_log2.jpg, (C) rawdata_Tug2_vs_-Tug3_log2.jpg, (D) rawdata_TugDn_1_vs_TugDn_2_log2.jpg, (E) rawdata_TugDn_1_vs_TugDn_-3_log2.jpg, (F) rawdata_TugDn_2_vs_TugDn_3_log2.jpg, (G) rawdata_TugDn2_1_vs_TugDn2_-

2_log2.jpg, (H) rawdata_TugDn2_1_vs_TugDn2_3_log2.jpg, (I) rawdata_TugDn2_2_vs_TugDn2_-3_log2.jpg, (J) rawdata_TugDn5_1_vs_TugDn5_2_log2.jpg, (K) rawdata_TugDn5_1_vs_TugDn5_-3_log2.jpg, (L) rawdata_TugDn5_2_vs_TugDn5_3_log2.jpg, (M) esetVSN_Tug1_vs_Tug2_log2.-jpg, (N) esetVSN_Tug1_vs_Tug3_log2.jpg, (O) esetVSN_Tug2_vs_Tug3_log2.jpg, (P) esetVSN_-TugDn_1_vs_TugDn_2_log2.jpg, (Q) esetVSN_TugDn_1_vs_TugDn_3_log2.jpg, (R) esetVSN_-TugDn_2_vs_TugDn_3_log2.jpg, (S) esetVSN_TugDn2_1_vs_TugDn2_2_log2.jpg, (T) esetVSN_-TugDn2_1_vs_TugDn2_3_log2.jpg, (U) esetVSN_TugDn2_2_vs_TugDn2_3_log2.jpg, (V) eset-VSN_TugDn5_1_vs_TugDn5_2_log2.jpg, (W) esetVSN_TugDn5_1_vs_TugDn5_3_log2.jpg, (X) esetVSN_TugDn5_2_vs_TugDn5_3_log2.jpg, (Y) esetRMA_Tug1_vs_Tug2_log2.jpg, (Z) eset-RMA_Tug1_vs_Tug3_log2.jpg, (AA) esetRMA_Tug2_vs_Tug3_log2.jpg, (AB) esetRMA_TugDn-_1_vs_TugDn_2_log2.jpg, (AC) esetRMA_TugDn_1_vs_TugDn_3_log2.jpg, (AD) esetRMA_-TugDn_2_vs_TugDn_3_log2.jpg, (AE) esetRMA_TugDn2_1_vs_TugDn2_2_log2.jpg, (AF) eset-RMA_TugDn2_1_vs_TugDn2_3_log2.jpg, (AG) esetRMA_TugDn2_2_vs_TugDn2_3_log2.jpg, (AH) esetRMA_TugDn5_1_vs_TugDn5_2_log2.jpg, (AI) esetRMA_TugDn5_1_vs_TugDn5_3_-log2.jpg, (AJ) esetRMA_TugDn5_2_vs_TugDn5_3_log2.jpg, (AK) esetMAS_Tug1_vs_Tug2_log2-.jpg, (AL) esetMAS_Tug1_vs_Tug3_log2.jpg, (AM) esetMAS_Tug2_vs_Tug3_log2.jpg, (AN) eset-MAS_TugDn_1_ vs_TugDn_2_log2.jpg, (AO) esetMAS_TugDn_1_vs_TugDn_3_log2.jpg, (AP) esetMAS_TugDn_2_vs_TugDn_3_log2.jpg, (AQ) esetMAS_TugDn2_1_vs_TugDn2_2_log2.jpg, (AR) esetMAS_TugDn2_1_vs_TugDn2_3_log2.jpg, (AS) esetMAS_TugDn2_2_vs_TugDn2_3_-log2.jpg, (AT) esetMAS_TugDn5_1_vs_TugDn5_2_log2.jpg, (AU) esetMAS_TugDn5_1_vs_Tug-Dn5_3_log2.jpg, (AV) esetMAS_TugDn5_2_vs_TugDn5_3_log2.jpg, (AW) esetGCRMA_Tug1_-vs_Tug2_log2.jpg, (AX) esetGCRMA_Tug1_vs_Tug3_log2.jpg, (AY) esetGCRMA_Tug2_vs_-Tug3_log2.jpg, (AZ) esetGCRMA_TugDn_1_vs_TugDn_2_log2.jpg, (BA) esetGCRMA_TugDn_-1_vs_TugDn_3_log2.jpg, (BB) esetGCRMA_TugDn_2_vs_TugDn_3_log2.jpg, (BC) eset-GCRMA_TugDn2_1_vs_TugDn2_2_log2.jpg, (BD) esetGCRMA_TugDn2_1_vs_TugDn2_3_-log2.jpg, (BE) esetGCRMA_TugDn2_2_vs_TugDn2_3_log2.jpg, (BF) esetGCRMA_TugDn5_1_-vs_TugDn5_2_log2.jpg, (BG) esetGCRMA_TugDn5_1_vs_TugDn5_3_log2.jpg, (BH) eset-GCRMA_TugDn5_2_vs_TugDn5_3_log2.jpg.

The 18 Slides files are: (A) rawdata_RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (B) raw-data_RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (C) rawdata_RWA2_GS_2_vs_RWA2_GS_3_-log2.jpg, (D) rawdata_RWA2_GR_1_vs_RWA2_GR_2_log2.jpg, (E) rawdata_RWA2_GR_1_-vs_RWA2_GR_3_log2.jpg, (F) rawdata_RWA2_GR_2_vs_RWA2_GR_3_log2.jpg, (G) rawdata_-

RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (H) rawdata_RWA1_GR_1_vs_RWA1_GR_3_log2.jpg, (I) rawdata_RWA1_GR_2_vs_RWA1_GR_3_log2.jpg, (J) rawdata_RWA1_GS_2_vs_RWA1_-GS_1_log2.jpg, (K) rawdata_RWA1_GS_3_vs_RWA1_GS_1_log2.jpg, (L) rawdata_RWA1_GS_-3_vs_RWA1_GS_2_log2.jpg, (M) rawdata_Gam_S_1_vs_Gam_S_2_log2.jpg, (N) rawdata_Gam_-S_1_vs_Gam_S_3_log2.jpg, (O) rawdata_Gam_S_2_vs_Gam_S_3_log2.jpg, (P) rawdata_Gam_R_-1_vs_Gam_R_2_log2.jpg, (Q) rawdata_Gam_R_1_vs_Gam_R_3_log2.jpg, (R) rawdata_Gam_R_-2_vs_Gam_R_3_log2.jpg, (S) esetVSN_RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (T) esetVSN_-RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (U) esetVSN_RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (V) esetVSN_RWA2_GR_1_vs_RWA2_GR_2_log2.jpg, (W) esetVSN_RWA2_GR_1_vs_-RWA2_GR_3_log2.jpg, (X) esetVSN_RWA2_GR_2_vs_RWA2_GR_3_log2.jpg, (Y) esetVSN_-RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (Z) esetVSN_RWA1_GR_1_vs_RWA1_GR_3_log2.-jpg, (AA) esetVSN_RWA1_GR_2_vs_RWA1_GR_3_log2.jpg, (AB) esetVSN_RWA1_GS_2_-vs_RWA1_GS_1_log2.jpg, (AC) esetVSN_RWA1_GS_3_vs_RWA1_GS_1_log2.jpg, (AD) eset-VSN_RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (AE) esetVSN_Gam_S_1_vs_Gam_S_2_log2.jpg, (AF) esetVSN_Gam_S_1_vs_Gam_S_3_log2.jpg, (AG) esetVSN_Gam_S_2_vs_Gam_S_3_log2.-jpg, (AH) esetVSN_Gam_R_1_vs_Gam_R_2_log2.jpg, (AI) esetVSN_Gam_R_1_vs_Gam_R_3_-log2.jpg, (AJ) esetVSN_Gam_R_2_vs_Gam_R_3_log2.jpg, (AK) esetRMA_RWA2_GS_1_vs_-RWA2_GS_2_log2.jpg, (AL) esetRMA_RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (AM) eset-RMA_RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (AN) esetRMA_RWA2_GR_1_vs_RWA2_GR_-2_log2.jpg, (AO) esetRMA_RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (AP) esetRMA_RWA2_-GR_2_vs_RWA2_GR_3_log2.jpg, (AQ) esetRMA_RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (AR) esetRMA_RWA1_GR_1_vs_RWA1_GR_3_log2.jpg, (AS) esetRMA_RWA1_GR_2_vs_-RWA1_GR_3_log2.jpg, (AT) esetRMA_RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (AU) eset-RMA_RWA1_GS_3_vs_RWA1_GS_1_log2.jpg, (AV) esetRMA_RWA1_GS_3_vs_RWA1_GS_-2_log2.jpg, (AW) esetRMA_Gam_S_1_vs_Gam_S_2_log2.jpg, (AX) esetRMA_Gam_S_1_vs_-Gam_S_3_log2.jpg, (AY) esetRMA_Gam_S_2_vs_Gam_S_3_log2.jpg, (AZ) esetRMA_Gam_R_-1_vs_Gam_R_2_log2.jpg, (BA) esetRMA_Gam_R_1_vs_Gam_R_3_log2.jpg, (BB) esetRMA_-Gam_R_2_vs_Gam_R_3_log2.jpg, (BC) esetMAS_RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (BD) esetMAS_RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (BE) esetMAS_RWA2_GS_2_vs_RWA2_-GS_3_log2.jpg, (BF) esetMAS_RWA2_GR_1_vs_RWA2_GR_2_log2.jpg, (BG) esetMAS_-RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (BH) esetMAS_RWA2_GR_2_vs_RWA2_GR_3_log2.-jpg, (BI) esetMAS_RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (BJ) esetMAS_RWA1_GR_1_vs_-

RWA1_GR_3_log2.jpg, (BK) esetMAS_RWA1_GR_2_vs_RWA1_GR_3_log2.jpg, (BL) esetMAS_RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (BM) esetMAS_RWA1_GS_3_vs_RWA1_GS_1_log2.jpg, (BN) esetMAS_RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (BO) esetMAS_Gam_S_1_vs_Gam_S_2_log2.jpg, (BP) esetMAS_Gam_S_1_vs_Gam_S_3_log2.jpg, (BQ) esetMAS_Gam_S_2_vs_Gam_S_3_log2.jpg, (BR) esetMAS_Gam_R_1_vs_Gam_R_2_log2.jpg, (BS) esetMAS_Gam_R_1_vs_Gam_R_3_log2.jpg, (BT) esetMAS_Gam_R_2_vs_Gam_R_3_log2.jpg, (BU) esetGCRMA_RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (BV) esetGCRMA_RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (BW) esetGCRMA_RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (BX) esetGCRMA_RWA2_GR_1_vs_RWA2_GR_2_log2.jpg, (BY) esetGCRMA_RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (BZ) esetGCRMA_RWA2_GR_2_vs_RWA2_GR_3_log2.jpg, (CA) esetGCRMA_RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (CB) esetGCRMA_RWA1_GR_1_vs_RWA1_GR_3_log2.jpg, (CC) esetGCRMA_RWA1_GR_2_vs_RWA1_GR_3_log2.jpg, (CD) esetGCRMA_RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (CE) esetGCRMA_RWA1_GS_3_vs_RWA1_GS_1_log2.jpg, (CF) esetGCRMA_RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (CG) esetGCRMA_Gam_S_1_vs_Gam_S_2_log2.jpg, (CH) esetGCRMA_Gam_S_1_vs_Gam_S_3_log2.jpg, (CI) esetGCRMA_Gam_S_2_vs_Gam_S_3_log2.jpg, (CJ) esetGCRMA_Gam_R_1_vs_Gam_R_2_log2.jpg, (CK) esetGCRMA_Gam_R_1_vs_Gam_R_3_log2.jpg, (CL) esetGCRMA_Gam_R_2_vs_Gam_R_3_log2.jpg.

Figure Appx 5.8 The MA-plot comparisons of the different slides after normalization by the different methods. The raw/pre-normalized plots are also included for comparison. The files are written to the subfolder */R results/Graphs of raw and normalized data/MA plots/*. The 12 slides files are: (A) rawdata_Tug1_vs_Tug2.jpg, (B) rawdata_Tug1_vs_Tug3.jpg, (C) rawdata_Tug2_vs_Tug3.jpg, (D) rawdata_TugDn_1_vs_TugDn_2.jpg, (E) rawdata_TugDn_1_vs_TugDn_3.jpg, (F) rawdata_TugDn_2_vs_TugDn_3.jpg, (G) rawdata_TugDn2_1_vs_TugDn2_2.jpg, (H) rawdata_TugDn2_1_vs_TugDn2_3.jpg, (I) rawdata_TugDn2_2_vs_TugDn2_3.jpg, (J) rawdata_TugDn5_1_vs_TugDn5_2.jpg, (K) rawdata_TugDn5_1_vs_TugDn5_3.jpg, (L) rawdata_TugDn5_2_vs_TugDn5_3.jpg, (M) esetVSN_Tug1_vs_Tug2.jpg, (N) esetVSN_Tug1_vs_Tug3.jpg, (O) esetVSN_Tug2_vs_Tug3.jpg, (P) esetVSN_TugDn_1_vs_TugDn_2.jpg, (Q) esetVSN_TugDn_1_vs_TugDn_3.jpg, (R) esetVSN_TugDn_2_vs_TugDn_3.jpg, (S) esetVSN_TugDn2_1_vs_TugDn2_2.jpg, (T) esetVSN_TugDn2_1_vs_TugDn2_3.jpg, (U) esetVSN_TugDn2_2_vs_TugDn2_3.jpg, (V) esetVSN_TugDn5_1_vs_TugDn5_2.jpg, (W) esetVSN_TugDn5_1_vs_TugDn5_3.jpg, (X) esetVSN_TugDn5_2_vs_TugDn5_3.jpg, (Y) esetRMA_Tug1_vs_Tug2.jpg, (Z) esetRMA_-

Tug1_vs_Tug3.jpg, (AA) esetRMA_Tug2_vs_Tug3.jpg, (AB) esetRMA_TugDn_1_vs_TugDn_2.jpg, (AC) esetRMA_TugDn_1_vs_TugDn_3.jpg, (AD) esetRMA_TugDn_2_vs_TugDn_3.jpg, (AE) esetRMA_TugDn2_1_vs_TugDn2_2.jpg, (AF) esetRMA_TugDn2_1_vs_TugDn2_3.jpg, (AG) esetRMA_TugDn2_2_vs_TugDn2_3.jpg, (AH) esetRMA_TugDn5_1_vs_TugDn5_2.jpg, (AI) esetRMA_TugDn5_1_vs_TugDn5_3.jpg, (AJ) esetRMA_TugDn5_2_vs_TugDn5_3.jpg, (AK) esetPLM_Tug1_vs_Tug2.jpg, (AL) esetPLM_Tug1_vs_Tug3.jpg, (AM) esetPLM_Tug2_vs_Tug3.jpg, (AN) esetPLM_TugDn_1_ vs_TugDn_2.jpg, (AO) esetPLM_TugDn_1_vs_TugDn_3.jpg, (AP) esetPLM_TugDn_2_vs_TugDn_3.jpg, (AQ) esetPLM_TugDn2_1_vs_TugDn2_2.jpg, (AR) esetPLM_TugDn2_1_vs_TugDn2_3.jpg, (AS) esetPLM_TugDn2_2_vs_TugDn2_3.jpg, (AT) esetPLM_TugDn5_1_vs_TugDn5_2.jpg, (AU) esetPLM_TugDn5_1_vs_TugDn5_3.jpg, (AV) esetPLM_TugDn5_2_vs_TugDn5_3.jpg, (AW) esetMAS_Tug1_vs_Tug2.jpg, (AX) esetMAS_Tug1_vs_Tug3.jpg, (AY) esetMAS_Tug2_vs_Tug3.jpg, (AZ) esetMAS_TugDn_1_vs_TugDn_2.jpg, (BA) esetMAS_TugDn_1_vs_TugDn_3.jpg, (BB) esetMAS_TugDn_2_vs_TugDn_3.jpg, (BC) esetMAS_TugDn2_1_vs_TugDn2_2.jpg, (BD) esetMAS_TugDn2_1_vs_TugDn2_3.jpg, (BE) esetMAS_TugDn2_2_vs_TugDn2_3.jpg, (BF) esetMAS_TugDn5_1_vs_TugDn5_2.jpg, (BG) esetMAS_TugDn5_1_vs_TugDn5_3.jpg, (BH) esetMAS_TugDn5_2_vs_TugDn5_3.jpg, (BI) esetGCRMA_Tug1_vs_Tug2.jpg, (BJ) esetGCRMA_Tug1_vs_Tug3.jpg, (BK) esetGCRMA_Tug2_vs_Tug3.jpg, (BL) esetGCRMA_TugDn_1_vs_TugDn_2.jpg, (BM) esetGCRMA_TugDn_1_vs_TugDn_3.jpg, (BN) esetGCRMA_TugDn_2_vs_TugDn_3.jpg, (BO) esetGCRMA_TugDn2_1_vs_TugDn2_2.jpg, (BP) esetGCRMA_TugDn2_1_vs_TugDn2_3.jpg, (BQ) esetGCRMA_TugDn2_2_vs_TugDn2_3.jpg, (BR) esetGCRMA_TugDn5_1_vs_TugDn5_2.jpg, (BS) esetGCRMA_TugDn5_1_vs_TugDn5_3.jpg, (BT) esetGCRMA_TugDn5_2_vs_TugDn5_3.jpg.

The 18 Slides files are: (A) rawdata_RWA2_GS_1_vs_RWA2_GS_2.jpg, (B) rawdata_RWA2_GS_1_vs_RWA2_GS_3.jpg, (C) rawdata_RWA2_GS_2_vs_RWA2_GS_3.jpg, (D) rawdata_RWA2_GR_1_vs_RWA2_GR_2.jpg, (E) rawdata_RWA2_GR_1_vs_RWA2_GR_3.jpg, (F) rawdata_RWA2_GR_2_vs_RWA2_GR_3.jpg, (G) rawdata_RWA1_GR_1_vs_RWA1_GR_2.jpg, (H) rawdata_RWA1_GR_1_vs_RWA1_GR_3.jpg, (I) rawdata_RWA1_GR_2_vs_RWA1_GR_3.jpg, (J) rawdata_RWA1_GS_2_vs_RWA1_GS_1.jpg, (K) rawdata_RWA1_GS_3_vs_RWA1_GS_1.jpg, (L) rawdata_RWA1_GS_3_vs_RWA1_GS_2.jpg, (M) rawdata_Gam_S_1_vs_Gam_S_2.jpg, (N) rawdata_Gam_S_1_vs_Gam_S_3.jpg, (O) rawdata_Gam_S_2_vs_Gam_S_3.jpg, (P) rawdata_Gam_R_1_vs_Gam_R_2.jpg, (Q) rawdata_Gam_R_1_vs_Gam_R_3.jpg, (R) rawdata_Gam_R_2_vs_Gam_R_3.jpg, (S) esetVSN_RWA2_GS_1_vs_RWA2_GS_2.jpg, (T) esetVSN_-

RWA2_GS_1_vs_RWA2_GS_3.jpg, (U) esetVSN_RWA2_GS_2_vs_RWA2_GS_3.jpg, (V) esetVSN_RWA2_GR_1_vs_RWA2_GR_2.jpg, (W) esetVSN_RWA2_GR_1_vs_RWA2_GR_3.jpg, (X) esetVSN_RWA2_GR_2_vs_RWA2_GR_3.jpg, (Y) esetVSN_RWA1_GR_1_vs_RWA1_GR_2.jpg, (Z) esetVSN_RWA1_GR_1_vs_RWA1_GR_3.jpg, (AA) esetVSN_RWA1_GR_2_vs_RWA1_GR_3.jpg, (AB) esetVSN_RWA1_GS_2_vs_RWA1_GS_1.jpg, (AC) esetVSN_RWA1_GS_3_vs_RWA1_GS_1.jpg, (AD) esetVSN_RWA1_GS_3_vs_RWA1_GS_2.jpg, (AE) esetVSN_Gam_S_1_vs_Gam_S_2.jpg, (AF) esetVSN_Gam_S_1_vs_Gam_S_3.jpg, (AG) esetVSN_Gam_S_2_vs_Gam_S_3.jpg, (AH) esetVSN_Gam_R_1_vs_Gam_R_2.jpg, (AI) esetVSN_Gam_R_1_vs_Gam_R_3.jpg, (AJ) esetVSN_Gam_R_2_vs_Gam_R_3.jpg, (AK) esetRMA_RWA2_GS_1_vs_RWA2_GS_2.jpg, (AL) esetRMA_RWA2_GS_1_vs_RWA2_GS_3.jpg, (AM) esetRMA_RWA2_GS_2_vs_RWA2_GS_3.jpg, (AN) esetRMA_RWA2_GR_1_vs_RWA2_GR_2.jpg, (AO) esetRMA_RWA2_GR_1_vs_RWA2_GR_3.jpg, (AP) esetRMA_RWA2_GR_2_vs_RWA2_GR_3.jpg, (AQ) esetRMA_RWA1_GR_1_vs_RWA1_GR_2.jpg, (AR) esetRMA_RWA1_GR_1_vs_RWA1_GR_3.jpg, (AS) esetRMA_RWA1_GR_2_vs_RWA1_GR_3.jpg, (AT) esetRMA_RWA1_GS_2_vs_RWA1_GS_1.jpg, (AU) esetRMA_RWA1_GS_3_vs_RWA1_GS_1.jpg, (AV) esetRMA_RWA1_GS_3_vs_RWA1_GS_2.jpg, (AW) esetRMA_Gam_S_1_vs_Gam_S_2.jpg, (AX) esetRMA_Gam_S_1_vs_Gam_S_3.jpg, (AY) esetRMA_Gam_S_2_vs_Gam_S_3.jpg, (AZ) esetRMA_Gam_R_1_vs_Gam_R_2.jpg, (BA) esetRMA_Gam_R_1_vs_Gam_R_3.jpg, (BB) esetRMA_Gam_R_2_vs_Gam_R_3.jpg, (BC) esetPLM_RWA2_GS_1_vs_RWA2_GS_2.jpg, (BD) esetPLM_RWA2_GS_1_vs_RWA2_GS_3.jpg, (BE) esetPLM_RWA2_GS_2_vs_RWA2_GS_3.jpg, (BF) esetPLM_RWA2_GR_1_vs_RWA2_GR_2.jpg, (BG) esetPLM_RWA2_GR_1_vs_RWA2_GR_3.jpg, (BH) esetPLM_RWA2_GR_2_vs_RWA2_GR_3.jpg, (BI) esetPLM_RWA1_GR_1_vs_RWA1_GR_2.jpg, (BJ) esetPLM_RWA1_GR_1_vs_RWA1_GR_3.jpg, (BK) esetPLM_RWA1_GR_2_vs_RWA1_GR_3.jpg, (BL) esetPLM_RWA1_GS_2_vs_RWA1_GS_1.jpg, (BM) esetPLM_RWA1_GS_3_vs_RWA1_GS_1.jpg, (BN) esetPLM_RWA1_GS_3_vs_RWA1_GS_2.jpg, (BO) esetPLM_Gam_S_1_vs_Gam_S_2.jpg, (BP) esetPLM_Gam_S_1_vs_Gam_S_3.jpg, (BQ) esetPLM_Gam_S_2_vs_Gam_S_3.jpg, (BR) esetPLM_Gam_R_1_vs_Gam_R_2.jpg, (BS) esetPLM_Gam_R_1_vs_Gam_R_3.jpg, (BT) esetPLM_Gam_R_2_vs_Gam_R_3.jpg, (BU) esetMAS_RWA2_GS_1_vs_RWA2_GS_2.jpg, (BV) esetMAS_RWA2_GS_1_vs_RWA2_GS_3.jpg, (BW) esetMAS_RWA2_GS_2_vs_RWA2_GS_3.jpg, (BX) esetMAS_RWA2_GR_1_vs_RWA2_GR_2.jpg, (BY) esetMAS_RWA2_GR_1_vs_RWA2_GR_3.jpg, (BZ) esetMAS_RWA2_GR_2_vs_RWA2_GR_3.jpg, (CA) esetMAS_RWA1_GR_1_vs_RWA1_GR_2.jpg, (CB) esetMAS_RWA1_-

GR_1_vs_RWA1_GR_3.jpg, (CC) esetMAS_RWA1_GR_2_vs_RWA1_GR_3.jpg, (CD) esetMAS_RWA1_GS_2_vs_RWA1_GS_1.jpg, (CE) esetMAS_RWA1_GS_3_vs_RWA1_GS_1.jpg, (CF) esetMAS_RWA1_GS_3_vs_RWA1_GS_2.jpg, (CG) esetMAS_Gam_S_1_vs_Gam_S_2.jpg, (CH) esetMAS_Gam_S_1_vs_Gam_S_3.jpg, (CI) esetMAS_Gam_S_2_vs_Gam_S_3.jpg, (CJ) esetMAS_Gam_R_1_vs_Gam_R_2.jpg, (CK) esetMAS_Gam_R_1_vs_Gam_R_3.jpg, (CL) esetMAS_Gam_R_2_vs_Gam_R_3.jpg, (CM) esetGCRMA_RWA2_GS_1_vs_RWA2_GS_2.jpg, (CN) esetGCRMA_RWA2_GS_1_vs_RWA2_GS_3.jpg, (CO) esetGCRMA_RWA2_GS_2_vs_RWA2_GS_3.jpg, (CP) esetGCRMA_RWA2_GR_1_vs_RWA2_GR_2.jpg, (CQ) esetGCRMA_RWA2_GR_1_vs_RWA2_GR_3.jpg, (CR) esetGCRMA_RWA2_GR_2_vs_RWA2_GR_3.jpg, (CS) esetGCRMA_RWA1_GR_1_vs_RWA1_GR_2.jpg, (CT) esetGCRMA_RWA1_GR_1_vs_RWA1_GR_3.jpg, (CU) esetGCRMA_RWA1_GR_2_vs_RWA1_GR_3.jpg, (CV) esetGCRMA_RWA1_GS_2_vs_RWA1_GS_1.jpg, (CW) esetGCRMA_RWA1_GS_3_vs_RWA1_GS_1.jpg, (CX) esetGCRMA_RWA1_GS_3_vs_RWA1_GS_2.jpg, (CY) esetGCRMA_Gam_S_1_vs_Gam_S_2.jpg, (CZ) esetGCRMA_Gam_S_1_vs_Gam_S_3.jpg, (DA) esetGCRMA_Gam_S_2_vs_Gam_S_3.jpg, (DB) esetGCRMA_Gam_R_1_vs_Gam_R_2.jpg, (DC) esetGCRMA_Gam_R_1_vs_Gam_R_3.jpg, (DD) esetGCRMA_Gam_R_2_vs_Gam_R_3.jpg.

LIST OF TABLES

Table 2.1 *Triticum aestivum* RWA-resistant genes, their origins and mode of resistance.

Table 2.2 Morphological and physiological changes in cereals over time in response to RWA infestation.

Table 2.3 Proteins present in the apoplast and vascular system with special focus on wheat.

Table 3.1 Sequences analyzed and the differences observed against homologues on Genbank and between the different biotypes.

Table 5.1 Selected geneIDs/probe sets shown to be up regulated (12 slide experiment), either generally or for specific slide comparisons (TuD-Tug, Tu2-Tug, *etc*.). Values under the different normalization methods (VSN, RMA, *etc*.) indicate the number of times the specific probe set was obtained across the six slide comparisons within that specific normalization method, *i.e.* 6 indicates that the probe set was obtained in all six slide comparisons. The slide comparisons indicate the number of times the probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences. *Table Appx 5.2* shows examples of down regulated probe sets.

Table 5.2 GeneIDs/probe sets, selected in *Table 5.1*, that were still shown to be differentially regulated after FDR (Benjamini-Hochberg method) or FWER (Bonferroni) correction. Values under the different normalization methods (VSN, RMA, *etc*.) indicate the number of times the specific probe set was obtained across the six slide comparisons (TuD-Tug, Tu2-Tug, *etc*.) within that specific normalization method, *i.e.* 6 indicates that the probe set was obtained in all six slide comparisons. The slide comparisons indicate the number of times the probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the

"occurrence" (Occ.) value at specific confidences. *Table Appx 5.3* shows examples of differentially up regulated probe sets after FDR/FWER correction.

Table 5.3 The influences of an increased confidence on the same three geneIDs/probe sets, before and after FDR/FWER correction. Values under the different normalization methods (VSN, RMA, *etc.*) indicate the number of times the specific probe set was obtained across the six slide comparisons (TuD-Tug, Tu2-Tug, *etc.*) within that specific normalization method, *i.e.* 6 indicates that the geneID/probe set was obtained in all six slide comparisons. The slide comparisons indicate the number of times the probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences.

Table Appx 3.1 Primer pairs, targets and PCR conditions used in the different analyses.

Table Appx 5.1 Scripts, codes and files used and/or produced during the data analyses of the different Affymetrix experiments.

Table Appx 5.2 Selected GeneIDs/probe sets shown to be up regulated (12 slide experiment), either generally or for specific slide comparisons (TuD-Tug, Tu2-Tug, *etc.*). Values under the different normalization methods (VSN, RMA, *etc.*) indicate the number of times the specific probe set was obtained across the six slide comparisons within that specific normalization method, *i.e.* 6 indicates that the geneID was obtained in all six slide comparisons. The slide comparisons indicate the number of times the geneID/probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences. *Table 5.1* shows examples of up regulated GeneIDs/probe sets.

Table Appx 5.3 GeneIDs/probe sets, selected in *Table 5.1*, that were still shown to be differentially regulated after FDR (Benjamini-Hochberg method) or FWER (Bonferroni) correction. Values under the different normalization methods (VSN, RMA, *etc.*) indicate the number of times the specific probe set was obtained across the six slide comparisons (TuD-Tug, Tu2-Tug, *etc.*) within that

specific normalization method, *i.e.* 6 indicates that the geneID was obtained in all six slide comparisons. The slide comparisons indicate the number of times the geneID/probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences. *Table 5.2* shows examples of differentially up regulated GeneIDs/probe sets after FDR/FWER correction.

PREFACE

Plant-insect interactions are complex with plants having various constitutive and inducible defences that protect against an insect attack, while insects continually attempt to circumvent these defensive obstructions. Many studies have looked at wheat cultivars, their resistance to *Diuraphis noxia* Kurdjumov (Russian wheat aphid, RWA), RWA biotypes and the interaction between the RWA and various resistant and susceptible wheat cultivars. However insects, as in the case with the family Aphididae, are often in symbiotic interaction(s) with bacteria. These bacteria usually allow insects, in this case aphids, to feed on plant components that are not nutritious enough to sustain life. Although the bacterium usually associated with aphids, *Buchnera aphidicola*, has been the subject of numerous studies, few have investigated the role of this endosymbiont in *D. noxia* biotypes' ability to overcome wheat resistance. Previous research did suggest that the relationship between die RWA and *B. aphidicola* was degrading in regards to nutrition production by the bacterium, but this was before the new RWA biotypes appeared. This poses some interesting questions: (i) Are the *B. aphidicola* the same in all the RWA biotypes, or (ii) Can an improved bacterial contribution to the aphid host play a role in the establishment of new RWA biotypes? This study investigates the plant-aphid interaction, but with special focus on the role of the endosymbiont(s) on this interaction.

This project has resulted in various publications, presentations and posters.

Published peer-reviewed papers

Swanevelder ZH, Surridge AKJ, Venter E and Botha A-M (2010). Limited endosymbiont variation in *Diuraphis noxia* (Hemiptera: Aphididae) biotypes from the USA and South Africa. *Journal of Economic Entomology* **103:** 887-897.

Botha A-M, **Swanevelder ZH** and Lapitan NLV (2010). Transcript profiling of wheat genes expressed during feeding by two different biotypes of *Diuraphis noxia*. *Journal of Environmental Entomology* **39(4):** 1206-1231.

Published conference papers

Botha A-M, **Swanevelder ZH**, Schultz T, Van Eck L and Lapitan NLV (2008). Deciphering defense strategies that are elucidated in wheat containing different *Dn* resistance genes. *Proceedings of the 11th International Wheat Genetics Symposium, Brisbane, Australia*: pp. O29.1-O29.3

Conference presentations

Swanevelder ZH, Venter E and Botha-Oberholster A-M (2006). Does *Buchnera aphidicola* hold the key to the development of "new" *Diuraphis noxia* biotypes? *South African Genetics Society meeting*.

Swanevelder ZH, Venter E and Oberholster A-M (2007). The role of the endosymbiont *Buchnera aphidicola* in the development of new *Diuraphis noxia* biotypes. *University of Johannesburg Symposium*.

Botha A-M, **Swanevelder ZH**, Schultz T, Van Eck L and Lapitan NLV (2008). Several specific defense strategies are elucidated in wheat containing different *Dn* genes. *International Plant Resistance to Insects Meeting*: Abstracts p.55.

Swanevelder ZH, Venter E and Botha A-M (2008). The effect of normalization methods on the identification of differentially regulated genes after Affymetrix analysis. Poster. *South African Genetics Society meeting*.

Awards

Swanevelder ZH, Venter E and Botha-Oberholster A-M (2008). Endosymbiont involvement in the development of new *Diuraphis noxia* biotypes. WERA66 (*Integrated Management of Russian Wheat Aphid and other Cereal Arthropod Pests*) and IPRI (*International Plant Resistance to Insects Workshop*). This paper presented at the Joint Meeting of the WERA66 and IPRI received a 3rd place presentation award.

Genbank submissions/accessions

The following sequences have been deposited in Genbank: FJ705277-FJ705318 (Appendix Chapter 3, Sequences submitted to Genbank) and GU145279-GU145289 (Appendix Chapter 4, Sequences submitted to Genbank).

Papers in preparation/submitted

Swanevelder ZH and Botha A-M. The interaction between *Diuraphis noxia* (Kurdjumov) and its cereal host, *Triticum aestivum* L.

Swanevelder ZH, Venter E and Botha A-M. Variation in *Buchnera aphidicola*'s leucine plasmid confers advantage to Russian Wheat Aphid biotypes.

Swanevelder ZH and Botha A-M. The influence of normalization methods on the identification of differentially regulated transcripts during Affymetrix analysis: Just how often do you get the same transcript?

CHAPTER 1

INTRODUCTION

Diuraphis noxia Kurdjumov (Russian wheat aphid, RWA) is considered a major agricultural pest to the wheat and barley industry, with losses in the USA exceeding US$1 billion attributed to this pest. The introduction of various resistant wheat cultivars in the latter part of the 1980s significantly reduced the impact this pest had on the industry. However, the recent development of new RWA biotypes in the USA and South Africa nullify the resistance of many existing wheat cultivars. Again the RWA became a serious threat to wheat production in these regions. The new RWA biotypes, however, present an opportunity to investigate biotype development and the mechanisms aphids utilize to overcome cultivar resistance. Wheat transcriptome studies, where different cultivars are infested by one or more RWA biotypes, could explain the mechanisms of the different modes of plant resistance against the RWA. Furthermore, these studies could also highlight which resistant pressures influences aphid biotype development.

Literature is reviewed in Chapter 2. Here the literature on the members and the bacterium-aphid and aphid-plant interactions is reviewed. The chapter begins with an introduction to the aphid, *D. noxia*, which includes details like its origin, taxonomy, morphology, biology, preferred hosts, economic impact, symptoms and control. This is followed by a brief introduction on the symbiotic relationship between bacterial endosymbionts and their role in aphid success. The plant host, *Triticum aestivum*, is reviewed before the biotic interaction, *i.e.* aphid-plant interaction, is introduced. The chapter concludes with the aphid-plant interaction. Here the physical interaction between the aphid and plant, including plant defence evading mechanisms with special focus on the aphid salivary enzymes, are discussed.

Studies on plant-pest interactions usually deal with the specific pest or with the host plant (resistance). In the last couple of years, molecular techniques have increasingly driven the research in these fields and more studies now deal with both organisms. However, few plant-aphid investigations include the insects' endosymbiont. Endo- or symbiont-aphid interaction investigations highlighted the inter-dependencies of both on each other and the major influences that the endosymbiont(s) have on aphid biology and fitness as pest. Aphid success as plantsap feeders are directly contributed to an endosymbiont-aphid symbiosis. *Diuraphis noxia* have the bacterium *Buchnera aphidicola* as the endosymbiont. *Buchnera aphidicola* contains the biosynthetic pathways needed to produce leucine and tryptophan, two essential amino acids that are usually in low concentrations within the phloem. Aphids do not have the ability to produce these essential amino acids and are therefore dependent on their endosymbiont for amino acid production. The bacterium has moved rate limiting enzymes to multi-copy plasmids, thereby facilitating higher expression of

these essential amino acids. However, in the case of RWA-*B. aphidicola* relationship, the RWA seems less dependent on its endosymbiont since lower plasmid copy numbers and the presence of pseudogenes, have been reported for the endosymbiont of this aphid. Furthermore, *D. noxia* was shown to up-regulate leucine and tryptophan levels in phloem of susceptible wheat cultivars. Together, these finding suggests degradation in the aphid-endosymbiont relationship regarding essential amino acid production. The most cost effective regulation of RWA infestations are the employment of resistant wheat cultivars.

However, *D. noxia* is unable to up-regulate leucine and tryptophan in resistant wheat cultivars, therefore one could argue that the aphid was under severe selection pressure regarding essential amino acid production since the 1980s. This led to the first hypothesis of this study: There are no nucleotide differences in *B. aphidicola*'s leucine and selected tryptophan biosynthetic genes from the different RWA biotypes found in South Africa and the USA. Furthermore, the second hypothesis states the leucine plasmid copy numbers were the same for *B. aphidicola* of all the RWA biotypes. However, other aphid symbionts could also affect host fitness. In the first part of Chapter 3 the presence of other symbionts in the RWA biotypes was investigated using denaturing gradient gel electrophoresis (DGGE) of the bacterial 16S rRNA gene. With the confirmation of the RWA's monosymbiotic status, the chapter continues to investigate the first hypothesis, using *B. aphidicola* accessions from ten different RWA biotypes. Except for a single CCC-insert upstream of *leuA* gene in four of the ten RWA biotypes, no other differences were observed. Could this CCC-insert play a role in the regulation of the *leuA*-*leuB* operon? This led to a third hypothesis: The CCC-insert, found upstream of the *leuA* gene in *B. aphidicola* of some RWA biotypes, has a functional effect. This hypothesis was investigated in Chapter 4 by comparing accessions with and without the CCC-insert regarding the 5' UTR leader sequences, gene expression levels, *Rho*-independent terminator sites and predicted promoters. The chapter also investigated this region within other aphids as a possible regulatory mechanism within the family.

In the second part of the study (Chapter 5) the influences that statistical mechanisms have on the identification of differentially regulated genes within the RWA-host interaction, were investigated. The hypothesis of this chapter states that, though different background correction and normalization methods for Affymetrix datasets depend on different assumptions, they would eventually identify the same subset genes/probe sets as differentially regulated, especially under increased stringencies. In this chapter the influences of 5 different normalization and background correction methods, under three different confidence levels, with/without false discovery rate (FDR)

and family-wise type I error rate (FWER) correction, were investigated for two different experiments. The genes identified after these analyses as differentially regulated, were subsequently compared to identify how often a gene/probe set was deemed differentially regulated by all 5 smethods. The chapter's analyses were done using scripts written in the statistical program R.

The thesis's major results and conclusions are briefly discussed in the last chapter (Chapter 6). All the supplementary data for the different chapters are given in the appendices.

CHAPTER 2

APHID-PLANT-ENDOSYMBIONT INTERACTION: THE RUSSIAN WHEAT APHID, ITS HOSTS AND ENDOSYMBIONT, *BUCHNERA APHIDICOLA*

The symbiotic relationship between the bacterial endosymbiont *Buchnera aphidicola* and aphids enables highly specialised phloem feeding on host plants, thereby enabling aphids to feed almost undetected (Srivastava 1987; Douglas 1998; Moran *et al.* 2002; Voelckel *et al.* 2004). Plants try to counter this exploitation by employing constitutive and induced defences (Walling 2000). However, the influence that this aphid-plant interaction has on the bacterium, and the influence that the aphid-bacterium interaction has on overcoming plant resistance, have not been investigate very often (Walling 2000). This chapter aims to present an overview on the members, *i.e.* the Russian wheat aphid (*Diuraphis noxia* Kurdjumov), the endosymbiont (*Buchnera aphidicola* Munson *et al.*) and wheat (*Triticum aestivum* L.), and the biotic interactions between them.

Diuraphis noxia (Aphididae: Macrosiphini)

Diuraphis noxia (Russian wheat aphid, RWA) resides within the tribe Macrosiphini of the subfamily Aphidinae (Heie 1992). It forms part of the phytophagous suborder Sternorrhyncha and can therefore be either in the Homoptera (together with the Auchenorrhyncha) or the Hemiptera (the Homoptera and Heteroptera) (Miles 1999).

Origin

The RWA is palaearctic in origin, *i.e.* central Asia to the Middle East. However, various introductions have resulted in a worldwide distribution to all arid and semi-arid cereal producing regions, with the only exception being Australasia (Hewitt *et al.* 1984; Du Toit 1986; 1987; Zemetra *et al.* 1990; Souza *et al.* 1991; Gonzalez *et al.* 1992; Blackman & Eastop 2000; Stray 2000; Stary & Lukasova 2002; Baker *et al.* 2003; Haley *et al.* 2004). In the Republic of South Africa, it was first detected in 1978 with major yield losses resulting in subsequent years (Du Toit & Walters 1984; Du Toit 1987).

Description

Diuraphis noxia is a small, pale yellow-green or grey-green, spindle-shaped apterae (1.4-2.3 mm) that is often covered with a white powdery wax. Under adverse conditions it occurs as an alatae (1.5-2.0 mm) with pale-green abdomen (Hewitt *et al.* 1984; Walters *et al.* 1984; Gonzalez *et al.* 1992; Blackman & Eastop 2000). Overcrowding or a decline in host quality due to seasonal or

host morphological changes, trigger an ontogenetic switch from apterous (wingless) to alatae (winged) aphids that enables relocation and distribution (Walters *et al.* 1984; Dixon 1998). Characteristically to the RWA is the short antennae, a 'forked tail', *i.e.* a projection above the caudal, and an apparently absent siphunculi (Walters *et al.* 1984).

Biology

Females are viviparous and parthenogenetic, *i.e.* eggs commence development directly after ovulation. Nymphs also have the ability to produce embryos themselves, allowing for fast population expansions. Females can produce up to 4 nymphs per day that mature after approximately two weeks (Walters *et al.* 1984; Dixon 1998). Temperature affect the RWA, with higher temperatures having a negative effect on its the lifecycle (Girma *et al.* 1990).

Genetics

Very little is known about the genetics of the RWA. Aphids possesses holocentric chromosomes, *i.e.* chromosome with centromeric activity along the whole axis (Bizzaro *et al.* 2000 & references there in), that allow for karyotype rearrangements. This is thought to support the high occurrence of chromosomal polymorphism found in many aphid species (Manicardi *et al.* 2002).

Biotypes

Different RWA biotypes are distinguished from each other based on their ability to overcome host resistance, their fecundity and the amount of damage they cause to plants in a differential (Puterka *et al.* 1992; Jyoti & Michaud 2005; Burd *et al.* 2006; Jyoti *et al.* 2006; Weiland *et al.* 2008). *Diuraphis noxia* has a high worldwide biotypic diversity which is geographically limited (Puterka *et al.* 1992). Resistant lines is therefore only effective to the specific regions for which they've been bred, *e.g.* a Hungarian biotype was shown to be more virulent to South African resistant lines than the local biotype (Basky 2003). Similarly, Hungarian RWA populations were shown to be more virulent to resistant lines with *Dn1*, *Dn2*, *Dn4* and *Dn5* resistance genes (Basky 2003; Smith *et al.* 2004), while Russian, Syrian (Puterka *et al.* 1992), Chilean, Czech Republic and Ethiopian (Smith *et al.* 2004) populations showed resistance to *Dn4*. Czech populations showed resistance to *Dnx* and the Ethiopian population to *Dny* containing lines (Smith *et al.* 2004).

A second biotype has recently been added to the single biotype initially present in South African fields (Du Toit 1989; Tolmay *et al.* 2007), with a third biotype in the laboratory (*Figure 2.1*) (Van Zyl 2007). Today, eight biotypes are known to occur in the USA (Haley *et al.* 2004; Jyoti & Michaud 2005; Burd *et al.* 2006; Weiland *et al.* 2008). Molecular analysis of the USA biotypes showed little nuclear and mitochondrial variation (Lapitan *et al.* 2007b; Shufran *et al.* 2007), therefore suggesting adaptation or diversification of the existing USA biotype(s) rather than reintroductions. However, the biotype designation is assigned based on plant phenotypic responses to aphid feeding and no genetic, taxonomic or other differences can be presumed (Smith *et al.* 2005).

Figure 2.1 The South African (SA) biotype on the susceptible cultivar Scheepers, and the SA mutant (SAM) on the resistant wheat cultivar TugelaDN. No morphological differences are apparent. Photos taken under different magnifications.

Hosts

Host plants of the RWA include mainly barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*), but they have been found on other cereals and grasses. These include *H. murinum, T. cylindricum, Elymus arenarius* and *H. pusillum* (Butts & Pakendorf 1984b; Hewitt *et al.* 1984; alternative hosts reviewed in Kindler & Springer 1989; Kindler *et al.* 1992; Belefant-Miller *et al.* 1994; Blackman & Eastop 2000; Stray 2000). Post-harvest survival of RWA has been contributed to cultivation practices and volunteer plants (Hewitt *et al.* 1984; Kriel *et al.* 1986; Stray 2001). *Bromus*

species, *E. trachycaulum* and *E. agrotricum* may also play a role in the seasonal cycle of this aphid (Fouche *et al.* 1984; Hewitt *et al.* 1984; Kindler *et al.* 1992).

Symptoms

Several symptoms are characteristic of RWA infestation. These symptoms in susceptible cultivars include severe longitudinal streaking (*Figure 2.2*) or spotting that is white, yellow or purple in colour (chlorosis), and tightly inward curling leaf edges (leaf rolling) (Du Toit 1986). Growth can be retarded and heavily infested plants have a flattened appearance (Elsidaig & Zwer 1993). Flag leaf infestation leads to white bended ears and in severe cases susceptible cultivars die. Aphids are mainly found on the newest growth and axils of leaves. Infestations occur in a patchy distribution under field conditions and a 20 % infestation of the crop can escalate to 80 % in just 2 weeks if left unchecked. Resistant cultivars are identified by necrotic spots on leaves with no leaf rolling. Seed number, thousand-kernel mass, mass of ear per plant and total seed mass per plant for both susceptible and resistant lines is also reduced. Early infested plants usually have fewer productive tillers (up to 50 % less) and in many cases are dwarfed and of uneven height.

Figure 2.2 (A) Heavily infested wheat plants showing longitudinal streaking and (B) tightly inward curling of the leaf edges (leaf rolling).

Infested flag leaves can result in contorted grain heads that interfere with pollination and head extension (Butts & Pakendorf 1984c; Walters *et al.* 1984; Du Toit 1986; Kriel *et al.* 1986; Du Toit 1989; Quick *et al.* 1991; Smith *et al.* 1991; Dong & Quick 1995). Various screening keys have been developed to help quantify cultivar symptoms and resistance (Butts & Pakendorf 1984b; Du Toit 1987; Quick *et al.* 1991; Burd & Burton 1992). The longitudinal streaking/spotting and chlorosis observed during RWA feeding (Fouche *et al.* 1984; Hewitt *et al.* 1984; Kriel *et al.* 1984; Kriel *et al.* 1986) are sometimes mistakenly identified as virus transmission (Blackman & Eastop 2000; Kazemi *et al.* 2001).

Economic losses

Major yield losses over the years have been credited to the RWA. In the USA it is estimated to have caused losses exceeding US$ 1 billion (Quick *et al.* 1991; Gonzalez *et al.* 1992; Porter *et al.* 1998). Earlier research suggested yield losses of almost 35 % for plants (at growth stage 6) when infested with 9.1 aphids per plant, this converts to a reduction of 1 t/ha in yield (Du Toit & Walters 1984). Later reports calculated that RWA infested or damaged plants have yield reductions of 60-80 % for the 1987/88 season (Archer & Bynum 1992). Initial research suggested that major yield losses occurred when RWA feeding took place before growth stage 30 (stem elongation) (Tottman *et al.* 1979), but later studies suggest that wheat compensate for infestations stopped before this growth stage (Du Toit 1986; Kriel *et al.* 1986).

Control

Expensive systemic insecticides have proven relatively successful in controlling the RWA (Botha 1984; Walters *et al.* 1984; Du Toit 1989; Zemetra *et al.* 1990). Soil systemic insecticides were successful against the RWA but have a negative impact on yield (Du Toit 1984). Similarly, systemic insecticide treated seed controlled the RWA but were deemed economically unfeasible (Butts & Pakendorf 1984a). Leaf rolling as a result of RWA feeding, forms a protective enclosure for the aphids, making it difficult to reach them with contact insecticides and biological agents (Gonzalez *et al.* 1992; Bergeson & Messina 1998) — even though various natural enemies have been identify that may act as possible biological controls (Aalbersberg *et al.* 1984; Archer & Bynum 1992; Gonzalez *et al.* 1992; Reed *et al.* 1992; Clark & Messina 1998; Stray 2000; Bosque-Perez *et al.* 2002; Nowierski & Fitzgerald 2002; Prinsloo *et al.* 2002; Baker *et al.* 2003; Brewer *et al.* 2005).

The only viable options that remain are culture practices, *i.e.* delayed plantings, the eradication of volunteer wheat (Du Toit 1989), and resistant cultivars for the control of this pest. In recent years, the release of resistant cultivars in conjunction with biological controls, have reduced the effects of this pest on wheat fields (Bosque-Perez *et al.* 2002; Jyoti & Michaud 2005). Resistant cultivars, with normal root and shoot development, produced higher grain yields (Zwer *et al.* 1994). However, this scenario is changing with the new RWA biotypes discovered in the USA and South Africa (Haley *et al.* 2004; Jyoti & Michaud 2005; Tolmay *et al.* 2007) and with more virulent populations found across the globe (Puterka *et al.* 1992; Basky 2003; Smith *et al.* 2004).

Aphids and their endosymbiont, *Buchnera aphidicola*

Aphids primarily target the phloem sap of plants for all their nutritional needs (Srivastava 1987). Phloem sap, studied *via* stylet seepage (Bornman & Botha 1973), can be regarded as a excellent food source. It is high in carbohydrates, with nitrogen predominantly occurring as free amino acids. Generally, phloem is also toxin and feeding deterrent free, as these are usually localised within the vacuole and apoplast (Douglas 2006). Though high in sugars, phloem is low in nitrogen (~20 mol %), especially essential amino acids (Lai *et al.* 1996; Dixon 1998; Sandstrom & Moran 1999; Douglas 2006). This presents two problems to aphids: firstly, the ratio of non-essential to essential amino acids in phloem (1:4-1:20) is much lower than the 1:1 ratio found in animal proteins. Secondly, the highly concentrated sugars need to be regulated to prevent disturbing the osmotic pressure within the insect. This is done by the excretion of honeydew or by drinking from xylem (Douglas 2006; Will & van Bel 2006).

Aphids cannot produce essential amino acids themselves. This suggests that aphid growth is related to amino acid composition, rather than sucrose:amino acid ratio (Karley *et al.* 2002). However, essential amino acid concentrations within phloem, compared to that present in the aphid, cannot explain the high growth rates observed (Douglas 2006). It is a symbiotic relationship with an endosymbiont, *B. aphidicola*, that allows aphids to exploit phloem as food source (Douglas 1998).

Buchnera aphidicola's role in aphid nutrition

The primary endosymbiont of aphids, *B. aphidicola*, compensates for the deficiency in the diet by synthesising and recycling the necessary essential amino acids (Munson *et al.* 1991; Munson & Baumann 1993; Dixon 1998; Douglas 1998; Wilkinson 1998; Febvay *et al.* 1999; Sandstrom &

Moran 1999; Birkle *et al.* 2002; Moran *et al.* 2005). *Buchnera aphidicola* is transovarially and maternally inherited between generations, where it's maintained in specialised, aphid produced cells, called mycetocytes (or bacteriocytes) (Lai *et al.* 1994; Lai *et al.* 1996; Rouhbakhsh *et al.* 1996; Dixon 1998; Douglas 1998; Gil *et al.* 2004). Genome analysis highlighted the dependency of the endosymbiont on the host. Many genes from key pathways are absent in *B. aphidicola*, implying that the host provide many important substrates (Shigenobu *et al.* 2000; Tamas *et al.* 2002; Van Ham *et al.* 2003; Zientz *et al.* 2004). This supports the obligatory relationship known to exist between these organisms (Douglas 1998; Wernegreen & Moran 2000).

Genomics and transcriptomics of *B. aphidicola* demonstrated that most of the essential amino acid biosynthetic pathways, though not always complete, were retained and functional (Shigenobu *et al.* 2000; Tamas *et al.* 2002; Nakabachi *et al.* 2005). The type of essential amino acid biosynthetic pathways retained seems to be dependent on the aphid's diet (Tamas *et al.* 2002; Zientz *et al.* 2004). Depending on the need for a specific essential amino acid, *B. aphidicola* has duplicated genes or moved rate limiting enzymes, or even whole pathways, to single or multiple copy plasmids (Baumann *et al.* 1999). These rate limiting pathway enzymes are encoded on vertically, long-term transmission plasmids (Lai *et al.* 1996; Silva *et al.* 1998; Thao *et al.* 1998; Baumann *et al.* 1999; Soler *et al.* 2000; Wernegreen & Moran 2001).

Anthranilate synthase (AS), the first enzyme and limiting factor in tryptophan biosynthesis, is encoded as two subunits on the plasmid *TrpEG* (Lai *et al.* 1996). In most members of the Aphididae and even distantly related families like Pemphigidae, AS occurs as highly conserved repetitive units, inter-dispersed with less conserved intergenic regions on plasmids (Baumann *et al.* 1995; Rouhbakhsh *et al.* 1996; Van Ham *et al.* 1999). Tryptophan production is mainly controlled by feedback inhibition of the product, tryptophan (Lai *et al.* 1994; Lai *et al.* 1996). The overproduction of AS is an attempt to compensate for the inhibitory effect of tryptophan feedback inhibition (Lai *et al.* 1994). The rest of the genes in the pathway are located within the *B. aphidicola* genome (Baumann *et al.* 1998). The whole leucine gene pathway is also located on a plasmid, with gene organisation conserved in the Aphididae (Van Ham *et al.* 1997; Silva *et al.* 1998; Soler *et al.* 2000).

***Buchnera aphidicola*-RWA interaction**

Conversely, RWA seems less dependent on its endosymbiont for essential amino acids (Telang *et al.* 1999; Porter & Webster 2000; Sandstrom *et al.* 2000; Ni *et al.* 2001). This hypothesis is supported by various observations. Firstly, the tryptophan plasmid of *B. aphidicola* occurs in lower copy numbers than in other aphids that feed on similar hosts (Lai *et al.* 1996; Thao *et al.* 1998). Secondly, this plasmid also contains seven pseudogenes, with amino acid measurements indicating a reduction in tryptophan availability (Lai *et al.* 1994; Lai *et al.* 1996; Wernegreen & Moran 2000). Retention of a single working *trpEG* gene copy by *B. aphidicola* is thought to be enough to allow sufficient tryptophan production when the RWA feeds on a less suitable host of preference. Thirdly, only a single copy of the leucine plasmid is retained (Thao *et al.* 1998; Baumann *et al.* 1999). This is in contrast to the 24 copies present in *B. aphidicola* of *Schizaphis graminum*, an aphid feeding on similar hosts (Thao *et al.* 1998). These reductions in copy numbers of genes involved in essential amino acid production is thought to have occurred in response to the ability of the RWA to alter the phloem sap composition of susceptible wheat cultivars (Telang *et al.* 1999; Porter & Webster 2000; Sandstrom *et al.* 2000; Ni *et al.* 2001).

Initially, like *S. graminum*, phloem sap of RWA susceptible cultivars are low in tryptophan and leucine (Thao *et al.* 1998). However, RWA feeding increases the essential amino acid component, especially leucine and tryptophan, in phloem of susceptible cultivars (Telang *et al.* 1999). The elevated levels of essential amino acids found in susceptible accessions infested by RWA could be the result of the elevated expression of amino acid transporter(s) (Van Niekerk 2003). Amino acid transporters, for specific amino acids, are known to be expressed under different environmental conditions (for review see Fischer *et al.* 1998; Delrot *et al.* 2000; Ortiz-Lopez *et al.* 2000). This observation suggests that a RWA diet of susceptible cultivars contain sufficient tryptophan and leucine, thereby removing the evolutionary pressure on the endosymbiont to retain all its gene copies (Lai *et al.* 1996; Telang *et al.* 1999). However, RWA cannot up-regulate these essential amino acids in resistant cultivars (Telang *et al.* 1999). This suggests that the selective pressure of lower leucine and tryptophan levels are back on *B. aphidicola* when the RWA is forced to feed on resistance wheat cultivars.

Triticum aestivum as aphid host

Triticum aestivum L. em. Thell. is believed to have originated ~10 000 years ago from the hybridisation of *T. tauschii* (Coss.) Schmal and *T. turgidum* L. This hybridisation resulted in today's allopolyploid bread wheat (2*n* = 6x = 42, AABBDD), where *T. tauschii* is considered to be the DD genome donor (2*n* = 14) and the allotetraploid *T. turgidum*, the AABB genome donor (2*n* = 28) (Gill *et al.* 1991 & references; Feldman & Levy 2005). *Triticum urartu* (AA) and an *Aegilops* species of section Sitopsis (BB) are considered the progenitor species for *T. turgidum* (Feldman & Levy 2005). This allopolyploidy of wheat, caused an evolutionary bottleneck with a narrow genetic base as shown by limited polymorphism found in its cultivars (Gill *et al.* 1991; Kubalakova *et al.* 2002).

Resistant wheat cultivars

Plant resistance to RWA have been broadly categorized into antixenosis, antibiosis and tolerance (Painter 1958; Kogan & Ortman 1978; Du Toit 1987). Antixenosis is the non-preference of a plant to provide an insect with shelter, food or oviposition, *i.e.* the inability to sustain and serve as host plant (Painter 1936; 1958; Kogan & Ortman 1978; Du Toit 1987; Bahlmann 2002). Antixenosis is attributed to multiple genes located on chromosomes 2B, 6A and 7D (Castro *et al.* 2001). Antibiosis of a resistant accession is the ability to adversely affect the insect's biology and is measured using aphid fecundity (Painter 1958; Bahlmann 2002). It is regarded as the most common form of wheat resistance (Du Toit 1987; 1989) and causes a decrease in aphid longevity and fecundity, a delay in development and increased restlessness (Kindler *et al.* 1992; Smith *et al.* 1992). Antibiosis to RWA is attributed to the involvement of different and independent genes on different chromosomes (Castro *et al.* 1999). Genes responsible for antibiosis and antixenosis resistance were shown to be different (Castro *et al.* 1999). Tolerant plants have the ability to survive infestations classified as fatal or severe injury causing to a susceptible cultivar (Painter 1958). Research suggests that adjustments to photosynthesis may be the underlining process used by tolerant cultivars to overcome infestations (Haile *et al.* 1999; Botha *et al.* 2006). Tolerance is also thought to be under multi-gene control with genes located on chromosomes 1A, 1D and 6D (Castro *et al.* 2001).

Initial, hexaploid wheat germplasm showed little resistance against the RWA (Du Toit 1987). Therefore, RWA resistance breeding programmes utilized related species as sources for resistance genes. These included *T. monococcum*, *T. timopheevii*, *T. ventricosum*, *T. tauschii*, *A. squarrosa*, *T. dicoccoides* and *Secale cereale* (rye) (Butts & Pakendorf 1984b; Du Toit 1987; Nkongolo *et al.*

1991a; Potgieter *et al.* 1991; Marais *et al.* 1994). Breeding endeavours helped in the identification of various RWA resistance genes (*Table 2.1*). Most of these resistance genes formed part of mapping studies (Ma *et al.* 1998; Myburg *et al.* 1998; Fritz *et al.* 1999; Venter & Botha 2000; Liu *et al.* 2001; Miller *et al.* 2001; Liu *et al.* 2002).

Table 2.1 *Triticum aestivum* RWA-resistant genes, their origins and mode of resistance.

Gene	***Accession***	***General information***
Dn1	PI 137739 /SA 1684	Hard white spring wheat, Iran, single dominant gene, antibiosis and antixenosis. One major gene on chromosome 7D (another, unidentified resistant gene on chromosome 7B)[1,3,5,6]
Dn2	PI 262660 /SA 2199	Hard white winter wheat, Bulgaria, single dominant gene, antibiosis and antixenosis[1,3,5]
Dn3	*T. tauschii* /SQ24	Recessive gene[2]
Dn4	PI 372129 /CORWA1	Single, dominant, nuclear gene, on chromosome 1DS, tolerance most significant resistance method, with least RWA damage to plants. Lower levels of chlorosis and fewer days to death than PI 137739 (*Dn1*) and PI 262660 (*Dn2*). High resistance with low RWA damage[3,8,10]
Dn5	PI 294994 /SA 463	Hard red winter wheat, Bulgaria, chromosome 7D. Possibly linked with *Dn1*. Resistance and inheritance unclear[3,4,5,8,11]
Dn6	PI 243781	Dominant gene[8,9]
Dn7	Rye	Unknown[7]
Dn8	PI 294994	Unknown[12]
Dn9	PI 294994	Unknown[12]
Dnx	PI 220127	Single, dominant gene[12,14]
Dny	'Stanton'	Unknown[13]

[1]Du Toit (1987; 1989), [2]Nkongolo *et al.* (1991a), [3]Nkongolo *et al.* (1991b), [4]Elsidaig & Zwer (1993), [5]Marais & Du Toit (1993), [6]Schroeder-Teeter *et al.* (1993), [7]Marais *et al.* (1994), [8]Dong & Quick (1995), [9]Saidi & Quick (1996), [10]Dong *et al.* (1997), [11]Zhang *et al.* (1998), [12]Liu *et al.* (2001), [13]Smith *et al.* (2004), [14]Boyko *et al.* (2006).

Plant resistance and defence mechanisms

Plants are at any moment bombarded with various biotic and abiotic stimuli. In the case of biotic interactions, plants employ an integrated defence strategy where stimuli may activate various signal transduction pathways that interlink (or cross talk) with each other to ensure the right defence response at the right time (Genoud & Metraux 1999; Pieterse & Van Loon 1999). This ensures a flexible defence with better energy and resource management (Baldwin & Preston 1999; Walling 2000).

Aphids, as phloem-feeders, produce smaller injuries during their long-lasting feeding interactions with plants, unlike chewing and cell-content feeders. Defence signals activated by aphid feeding are thought to be similar to those activated by bacteria, viral or fungal pathogens (*Figure 2.3*) (Walling 2000). The interactions between aphid/pathogen and plant could be compatible (slow reaction with hypersensitive response (HR) usually absent) or incompatible (fast reaction usually accompanied by HR) (Walling 2000). Incompatible interactions could again be divided into non-host resistance and host resistance (Odjakova & Hadjiivanova 2001). Non-host resistance, present in all plants, is the resistance of a species against a particular pathogen, *e.g.* general resistance for the majority of potential pathogens (Heath 2000b; Odjakova & Hadjiivanova 2001). Non-host resistance can include both preformed/constitutive and induced defences. Constitutive plant defences are various structural and chemical traits that physically limit/prevent insects accessing tissue, and/or deters colonization (antixenosis), growth, development, reproduction or survival (antibiosis) (Walling 2000; Peeters 2002). These include leaf morphological features, like waxy cuticula, cell walls (CW), silica depositions in CWs, callose, peptides, suberin and surface features, like trichomes and waxes, and stored allelochemicals (Hammond-Kosack & Jones 1996; Heath 2000b; Walling 2000; Dangl & Jones 2001; Peeters 2002; Walling 2008). Generally, sessile phloem feeders prefer thin leaf lamina and cuticles (Peeters 2002). In grasses, cutin and silica in epidermal cell walls and cuticles, are known to provide additional barriers to insects (Brett & Waldron 1990).

Plants have an active approach for non-host induced resistance. Elicitors, usually originating from the plant CW or the specific pathogen are recognized and activate constitutive and inducible defence responses *via* a complex signalling network, encompassing Ca^{2+} and H^{+}-ion fluxes, reactive oxygen species (ROS) intermediates, ethylene, salicylic acid (SA), jasmonic acid (JA), nitric oxide (NO) and membrane depolarisations. Inducible defence responses include

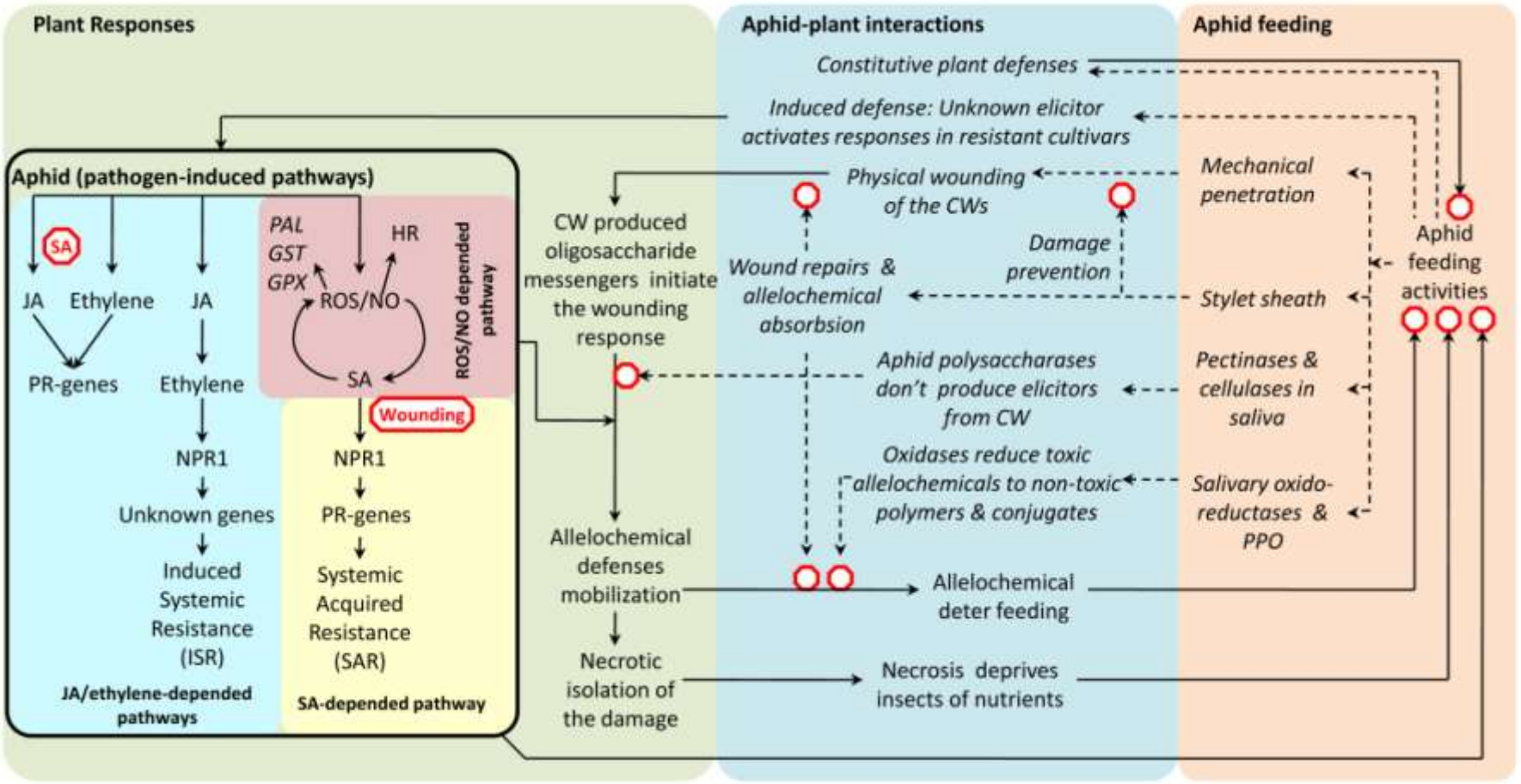


Figure 2.3 The interaction between aphids (italics and dashed lines) and the defence responses of plants (normal case and unbroken lines). Aphids activate similar signal transduction pathways as pathogens, which utilize the salicylic acid (SA), reactive oxygen species (ROS) and jasmonic acid (JA)/ethylene dependent pathways to activate gene expression in plants. Arrows indicate the interaction and red hexagons the inhibitions. Abbreviations: CW, cell wall; HR, hypersensitive response; NO, nitric oxide; PPO, polyphenol oxidase; PR-genes, pathogen related genes (Adapted from Miles 1999; Walling 2000).

the expression of pathogen-related proteins like protein kinases (PK), elements of the mitogen-activated protein kinases (MAPK) pathway, ROS, enzymes of the phytoalexin biosynthesis pathway and ultimately, programmed cell death (PCD) (Mittler *et al.* 1999; Heath 2000b; Walling 2000; Odjakova & Hadjiivanova 2001). These "non-specific" elicitors also plays a role in non-host HR generally associated with pathogen containment, elimination and the activation of defence genes (Heath 2000b). These collective pressures from various biotic and abiotic stresses have produced non-specific, broad-spectrum defences in response — an idea supported by the discovery of a cysteine protease with anti-fungal and feeding deterrent reactive sites (Joshi *et al.* 1999). The durability of non-host resistance, *i.e.* pathogens rarely alter host species range, makes it an ideal target for commercial breeding endeavours (Heath 2000b).

Host resistance (race/cultivar resistance), in contrast, is limited to certain expressed genotypes within an otherwise susceptible species and usually pathogen-genotype-specific or parasite-specific (Hammond-Kosack & Jones 1996; Heath 2000b; Odjakova & Hadjiivanova 2001). Host resistance is where a pathogen avirulence (*avr*) protein interacts/binds specifically with plant *R* protein(s), resulting in the recognition of the invader and the launch of the appropriate defence pathways (Hammond-Kosack & Jones 1996; Odjakova & Hadjiivanova 2001). Recognition by *R* genes can either be through this "gene-for-gene" ("receptor-ligand") interaction (van der Biezen & Jones 1998; Dangl & McDowell 2006), or through the recognition of *avr* mediated plant changes that promotes pathogen virulence ("guard hypothesis") (Dangl & Jones 2001; Chisholm *et al.* 2006). Both non-host and host resistance responses are proposed in aphid-plant interactions. A host resistant response, with a gene-for-gene interaction, is proposed for aphid-specific responses in resistant plants (Kaloshian 2004; Botha *et al.* 2006; Smith & Boyko 2007), while a non-host resistant response, activated as part of the general stress response after aphid tissue damage, is proposed for both susceptible and resistant plants (Van der Westhuizen *et al.* 1998b; Smith & Boyko 2007).

The origin and detection of aphid elicitors

Aphid probing can damage various tissues, many of which may produce defence responses as eliciting agents. Substances originating from the aphid's endosymbiont(s) are also possible elicitor candidates (Walling 2000; Smith & Boyko 2007 and references). Microbial elicitors (*Figure 2.4*), for example, can be peptides, proteins, lipids and oligosaccharides to name but a few (Montesano *et al.* 2003). Furthermore, saliva, as a foreign substance, may introduce elicitors into the plant. Digested cuticle products are well characterized in fungal recognition and defence response elicitors

in plants (Lequeu *et al.* 2003; Chassot & Metraux 2005). Exposure of a variety of plants and/or cell cultures to cutin monomers led to the effective protection of susceptible individuals, medium alkalinisation, ethylene production, activation of defence related genes and even H_2O_2 production (Chassot & Metraux 2005 & references therein). The complexity of the cuticle, consisting of surface wax, cutin and pectin (Brett & Waldron 1990), have many potential elicitors

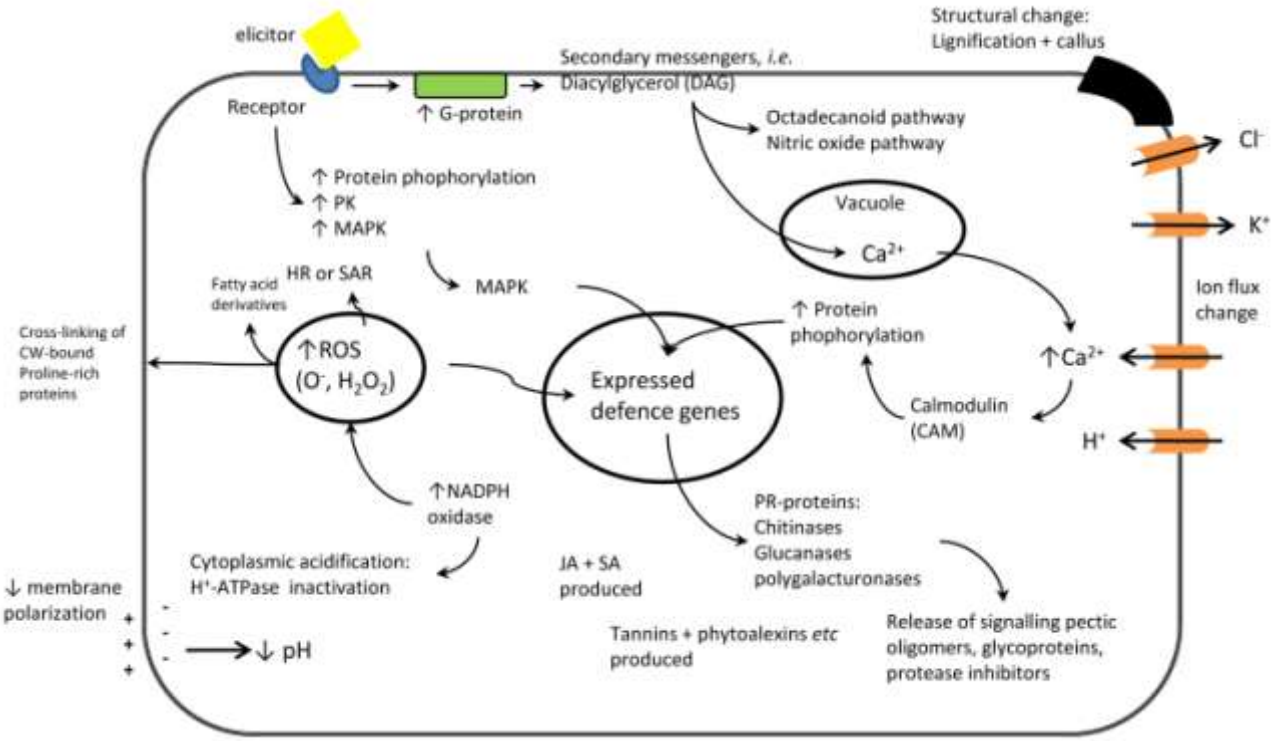


Figure 2.4 Biotic elicitation in plant-aphid interaction is thought to be similar to that of microbial pathogens. In microbial-pathogen interaction an elicitor binds to plasma membrane/CW receptor and activates protein phosphorylation, protein kinases (PK) and mitogen-activated PK (MAPK), as well as G-proteins (intracellular proteins that interact/function with receptors to regulate various enzymes and ion channels). Ion transporters are activated resulting in changes of ion fluxes (calcium and hydrogen ion influx increases). Calcium ions bind to calmodulin (CAM, non-enzymatic intracellular Ca^{2+}-binding proteins) which then binds to other proteins for regulation, resulting in the expression of defence genes. Secondary messengers are also activated, resulting in calcium release and the activation of various pathways. The cytoplasm acidifies as the result of NADPH oxidase activation, decrease in membrane polarization and the inactivation of H^+-ATPase. ROS activation and PR-protein expression can cause HR cell death at infection site or systemic acquired resistance (SAR) (compiled from Radman *et al.* 2003).

and may be the first step in RWA recognition. During a pathogen attack the CW degrading products, polysaccharides, diffuse into neighbouring cells triggering a defence response (Brett & Waldron 1990). Furthermore, phytoalexins, *i.e.* non-specific toxins produced by higher plants in response to pathogen attack, can also be induced from abiotic elicitors (*e.g.* from mechanical wounding), resulting in necrosis of host tissues (Brett & Waldron 1990). Proteinase inhibitors can also be produced in response to mechanical wounding (Brett & Waldron 1990).

Successful defence depends on the rapid recognition of a foreign invader which then triggers the appropriate response. The apoplast, with its interaction with the environment, is thought to play a key role in stress and pathogen detection (Foyer & Noctor 2005). Apoplastic transport moves various substances through the CW matrix (*i.e.* the apoplast) and across the plasma membrane located between neighbouring or widely separated cells (Brett & Waldron 1990). Many pathogenesis-related (PR) proteins accumulate in the apoplast and is therefore generally considered as the site where many eliciting signals of the defence responses originate and where defence products accumulates (Bowles 1990). Elicitors from fungal pathogens, can for example be detected by receptors in the plasma lemma (Nürnberger 1999) or CW (Heath 2000b), while bacterial and viral elicitors are perceived intracellular (Nürnberger 1999).

Plant resistance genes

In aphid-plant interactions, only a single plant resistance (*R*) gene has been cloned, *.e.g. Mi*, from the NBS-LRR (nucleotide binding site and leucine rich repeat) class, isolated from *Lycopersicon peruvianum* that confers resistance to the potato aphid, *Macrosiphum euphorbiae* (Kaloshian *et al.* 1995; Kanazin *et al.* 1996; Rossi *et al.* 1998; Smith & Boyko 2007). Other *R*-gene candidates in aphid-plant interactions include serine/theronine kinase, LRR-glycoproteins and leucine zipper (LZ)-NBS-LRR (reviewed by Smith & Boyko 2007). The various *Dn*-genes, characterized by the strength and type(s) of resistance they convey to cultivars, origin and mode of inheritance (*Table 2.1*), have yet to be successfully isolated. The current hypothesis is that the initial recognition and signalling of a co-ordinated defence response of resistant cultivars to the RWA, is due to specific plant *R*-genes, probably from the NBS-LRR protein class (Botha *et al.* 2006; Smith & Boyko 2007).

Resistance proteins are generally considered as mediators in recognising elicitors and activating downstream signalling responses, like the HR (Heath 2000a; Shirasu & Schulze-Lefert 2000; Moffett *et al.* 2002). Plant *R*-genes, based on amino acid homologies and characteristic

domain organisation, can be divided into five classes (Hammond-Kosack & Jones 1997; Young 2000; Dangl & Jones 2001). Intracellular or cytoplasmic serine or threonine PK forms the first distinct class (Hammond-Kosack & Jones 1997; Dangl & Jones 2001; Cannon & Young 2003), with sequence homology suggesting involvement in the signal transduction pathway, *e.g. Pto* of *Solanum lycopersicum* (syn. *Lycopersicon esculentum*) (Martin *et al.* 1993; Dangl & Jones 2001; Mysore *et al.* 2003). The second class contains a trans-membrane region connecting an extracellular leucine rich repeat (LRR) motif to a short cytoplasmic region, *e.g. Cf-X* genes of tomato (Jones *et al.* 1994; Cai *et al.* 1997; Thomas *et al.* 1997; Durrant *et al.* 2000; Dangl & Jones 2001; Cannon & Young 2003). The *Arabidopsis RPW8* is an example of the third class that consist of a trans-membrane or signal peptide and a coiled-coil (CC) domain (Dangl & Jones 2001; Xiao *et al.* 2001). The fourth class contains a large trans-membrane receptor, a large extracellular LRR domain and a cytoplasmic protein kinase domain, *e.g. Xa21* in *Oryza sativa* (Song *et al.* 1995; Pan *et al.* 2000a; Dangl & Jones 2001). The largest *R*-gene class contains both a nucleotide binding site (NBS) and a leucine-rich repeat (LRR) domain (Lagudah *et al.* 1997; Michelmore 2000; Dangl & Jones 2001; Halterman *et al.* 2001; Cannon & Young 2003).

NBS-LRR proteins are structurally characterized by a conserved nucleotide binding site (NBS), a variable N-terminal region and the varying number of LRRs at the carboxyl-terminal, (Hammond-Kosack & Jones 1997; Dangl & Jones 2001). The NBS region, a.k.a. nucleotide binding apoptosis *R*-gene & *CED-4* like domain (NB-ARC) or nucleotide binding domain (NBD), is further subdivided into small, highly conserved amino acid motifs (Bourne *et al.* 1991; Traut 1994; Cannon *et al.* 2002; Cannon & Young 2003). The NBS domain contains sequence motifs suggesting ATP binding or ATPases activity (Traut 1994; Pan *et al.* 2000b; Dangl & Jones 2001; Deslandes *et al.* 2002). LRR motifs provide the structural framework necessary in signalling transduction pathways for protein-protein, peptide-ligand and protein-carbohydrate interactions (Kobe & Deisenhofer 1994; 1995; Dangl & Jones 2001; Kobe & Kajava 2001). LRRs were also influential in determining the specificity of pathogen-specific gene-for-gene interactions, including the downstream signalling events (Ellis *et al.* 1999; Fluhr 2001; Kobe & Kajava 2001; Deslandes *et al.* 2002).

Redox signalling and plant defence

Redox signalling plays an important part in regulating defence gene expression (Baier & Dietz 2005 and references). In this signalling system, MAPKs play an important role and are again regulated by redox changes through phosphatase activities, with MAPK cascades resembling redox

signalling transduction pathways found in animals (Baier & Dietz 2005). Under normal conditions in the cytosol, antioxidants and antioxidant enzymes will maintain a reducing state, thus preventing ROS signals and the triggering of redox signalling pathways (Baier & Dietz 2005). ROS signalling of defences are determined by antioxidants, while the redox status of antioxidants, like ascorbate (AA), determines general plant defences to wounding and biotic stresses (Foyer & Noctor 2005). Indeed, changes in redox balance could induce defence-related genes, including PR-proteins. This can be seen when decreasing AA (or low levels of AA) leads to increase of PR-gene expression, this is similar for changes in glutathione (GSH), while dehydroascorbate (DHA) and oxidised glutathione (GSSH) accumulates during stress (Foyer & Noctor 2005; Noctor 2006). This 'redox sensing' by plants occurs when stresses change or adjust the redox state, thus leading to plant adjustments through other signalling systems like PK, phyto-hormones, protein phosphatases, ROS and calcium (Noctor 2006).

Transduction and defence pathways

Aphid feeding activates similar defence transduction pathways as bacteria, viruses or fungal pathogens, which include the JA/ethylene- and salicylic acid (SA)-dependent signalling pathways (*Figure 2.3*) (Walling 2000; Kaloshian 2004). These pathways act in a complex way to regulate each other and thus defence responses (Kunkel & Brooks 2002). The various signalling cascades results in local and systemic accumulation of defence RNAs and proteins, which includes enzymes capable of hydrolysing pathogen CWs, modifying plant CWs, synthesising secondary metabolites, producing signals for defence-signal modulation, protein turnover, *etc*. (Walling 2000). However, the aphid-plant interaction doesn't usually involve the wounding response. During the wounding response, breakdown products of plant tissues themselves act as elicitors to produce enzymatic cell breakdown (Wheeler 2001). Endogenous phenolics now leaking through broken membranes, are oxidized by oxidative enzymes and the production of these phenolics are stimulated in adjacent cells, thereby initiating repair processes like lignifications (Lamikanra 2002).

The HR is characterized by the rapid and localised death of plant cells, initiated by the plant (host) to prevent, restrict or confine the pathogen's growth and/or spreading (Heath 2000a; b; Lam *et al.* 2001; Odjakova & Hadjiivanova 2001). Typically, HR is recognised by one too many brown, dead cells at the infection site, usually as a result of ROS. Surrounding cells, not directly in contact or that is not physically invaded, may also die in HR, with sufficient numbers giving rise to necrotic lesions (Heath 2000a; Lam *et al.* 2001; Odjakova & Hadjiivanova 2001). Furthermore, cells

adjacent to these necrotic lesions become totally resistant to subsequent infections, *i.e.* show localised acquired resistance (Odjakova & Hadjiivanova 2001). HRs may or may not include cell death, but also includes the expression of *R*-genes (Heath 2000a; Lam *et al.* 2001). Host resistance that induces HR is generally controlled by gene-for-gene pathogen/parasite-specific *R*-genes, *i.e.* a specific *avr*-gene matching a specific *R*-gene (usually for fungal pathogens) or one *R*-gene capable of recognising multiple *avr*-genes (bacterial pathogens) (Heath 2000a).

Programmed cell death (PCD) is the activation of genetic programs inductive of cellular suicide and forms part of the HR, developmental programs, senescence, differentiation, development, seed germination and can even be induced by abiotic stresses (Beers & McDowell 2001; Lam *et al.* 2001). The systemic acquired resistance (SAR) pathway is activated after the HR response. The onset of SAR follows the accumulation of SA and is based on the expression of specific PR-genes that allows plants to maintain resistance against a pathogen, and includes PR-proteins like β-1,3-glucanase, chitinase, *etc*. (SAR reviewed by Ryals *et al.* 1996; Pieterse *et al.* 2001). SAR can be seen as a state of heightened defence and helps plants to protect themselves against a wide array of subsequent pathogen attacks (Kunkel & Brooks 2002; La Camera *et al.* 2004 reviewed microbial-plant interactions). Jasmonic acid (JA) or methyl JA (MeJA), known as jasmonates (JAs), are usually associated with chewing insects or the wounding responses (McConn *et al.* 1997; Titarenko *et al.* 1997; Halitschke & Baldwin 2005). However, JAs also plays a role in other cellular processes, like seed germination, development, senescence and leaf abscission (Seo *et al.* 2001). JA is produced by the octadecanoid pathway and forms part of the oxylipin signalling pathway (Halitschke & Baldwin 2005). Cross-talk, overlap or interaction of these various mechanisms exist (Beers & McDowell 2001), *e.g.* during HR the NO and ROS pathways interacts (Zaninotto *et al.* 2006) and SA, JA and ethylene pathways interact and control each other (Pieterse *et al.* 2001; Kunkel & Brooks 2002). Thus, though various expression studies have been done on insect-plant interactions (Fidantsef *et al.* 1999; Moran & Thompson 2001; Lacock *et al.* 2003; Mochida *et al.* 2003; Ogihara *et al.* 2003; Van Niekerk 2003) and on wheat-RWA interaction (Lacock *et al.* 2003; Van Niekerk 2003), the resistance mechanism involved in RWA host interaction is still unclear.

Aphid-plant interaction

Aphid probing and feeding

Aphids use a variety of chemical and physical stimuli in recognising their host of preference (Dixon 1987). A surface scan with the tip of the proboscis allows the detection of the vein contour, the preferred feeding site (Dixon 1998). This is confirmed with a drop of saliva, dissolving the cuticle which is then sensed by chemoreceptors on the labium tip (Srivastava 1987). The saliva also forms a flange on the surface through which the aphid will start probing (Miles 1968). Mandibular and maxillary stylets are arranged to form a needle-like structure with two deep grooves: one for pumping saliva out and the other for ingesting. This is used to probe into the plant while a continuous proteinaceous stylet sheath is secreted (Miles 1968; Dixon 1998; Brennan *et al.* 2001). The stylet sheath is thought to consist of mainly proteins, lipoproteins and phospholipids (Miles 1965; 1967; 1968). Sheath material, originating from the lateral lobe of the salivary gland, starts gelling immediately after secretion - probably due to hydrogen bonding and enzymatic oxidation of sulphydryl groups (Miles 1967; Dixon 1998). The product is a relatively impermeable, though flexible, salivary stylet sheath that provides some rigidity and directional control for the flexible stylet apex (Dixon 1998; Miles 1999).

Probing is the forward-backwards movement of the stylet, with a drop of saliva secreted before the next forward thrust. This movement drives the stylet through the gelling sheath material (Miles 1968; 1999), while brief stops allow sampling to determine the stylet position (Dixon 1998). These different actions produce two distinct EPG (electrical penetration graphs) wave patterns: one associated with salivation, while the other is associated with both salivation and feeding/testing (Dixon 1998). EPG, with transmission electron microscopy, showed that stylet paths follow an intercellular route, *i.e.* through the middle lamella, intercellular air spaces, secondary cell wall or amid the cell wall and plasmalemma, before going intracellular when entering the phloem (Dixon 1998; Miles 1999). However, when entering phloem cells, only the tip of the stylet penetrates and withdraw before sheath material is secreted (Miles 1968). RWA probing also follows an intercellular route, though the stylet tracks may have a branched appearance due to redirection (Fouche *et al.* 1984), that causes massive damage to the surrounding cells/tissue (Botha & Matsiliza 2004; Saheed *et al.* 2007b).

Following sieve element penetration, the viscous sheath forming saliva change into watery saliva (Miles 1965; Dixon 1998). Small volumes of watery saliva are also discharged intermittently

with sheath forming saliva when puncturing parenchyma cells and during intracellular probing. This is ingested with the surrounding substrate, thereby enabling the aphid to "taste" the current tissue being probed (Martin *et al.* 1997; Miles 1999). The functions of the watery saliva is unknown but could range from lubrication and food digestion, to food maintenance (reviewed by Miles 1999). Salivation is thought to take place when the aphid is not feeding or when the stylet is inserted/removed, but is continuous when parenchyma cells are used as food (Miles 1968; 1999). Aphids feeding on resistant plants usually don't continue with ingestion, but keep salivating or keep returning back to salivation (Will *et al.* 2007).

Pressure in the phloem elements is sufficient to drive phloem sap up the food canal of the stylets. Intake is controlled by a piston valve, though aphids do have the ability to suck/pump phloem sap. The phloem sap composition is sampled by gustatory epipharyngeal sensillae located in the pharynx (Dixon 1998). The alimentary tract is reviewed by Dixon (1987).

Evading the host defence

Aphids use various strategies to go unnoticed by plants while probing and feeding. These strategies range from preventing the release and/or synthesis of toxic compounds, to influencing redox poise, manipulating the wounding response, neutralising ROS and preventing induced defence pathways. Some of these approaches can also result in morphological changes in the host – mostly beneficial to the aphid. These strategies used in plant-aphid interactions can be either general or species specific.

Generally, aphid feeding may affect the intact cell layers around the stylet sheaths. These can demonstrate chloroplast degeneration, nucleus enlargement, the loss of starch and an increase in permeability. Free amino acids in aphid saliva were shown to cause increases in permeability and respiration, reduction in photosynthesis and monocot growth and even toxic reactions in plants (Miles 1968 and references). Sheath caused vascular blockages are thought to be responsible for localised pigmentation that can occur above blocked phloem vessels (as auxins and photosynthates accumulate) and wilting or decrease in transpiration in the case of xylem elements (Miles 1968 and references within). Aphid feeding can also cause hypertrophy in plant cells, *i.e.* the formation of leaf rolls (or pseudogalls). These pseudogalls are usually associated with cell growth on the opposite side of the aphid's feeding site and are thought to be as the result of increased auxin activity (Miles 1999).

Aphids need to neutralise any wounding or defence response while probing to enable a successful and long term feeding site. The probing process punctures numerous cells along the way, sometimes intentionally to determine the stylet position (Dixon 1998; Will & van Bel 2006). A small amount of watery saliva is released into the cell when a membrane is punctured (Miles 1968; Martin *et al.* 1997; Miles 1999; Saheed *et al.* 2007a). Salivary enzymes are thought to play a major role in suppressing host defence responses. The watery saliva also diffuse into the cells surrounding the advancing stylet track - no sheath material is produced until the stylet tip is removed (Miles 1999).

The stylet sheath prevents calcium inflow from the apoplast into the cell, thereby preventing a full launch of the wounding response (Yoo *et al.* 2002; Will & van Bel 2006; Will *et al.* 2007). Furthermore, no cell content, *i.e.* cytoplasm or ruptured vacuoles, can leak out into the intercellular spaces, thereby preventing the release/production of compounds that might trigger the defence responses in adjacent tissue (Miles 1999). Wound induced HRs are further suppressed when potentially toxic compounds (*e.g.* phenolics), that may promote necrosis, are absorbed and immobilised within the sheath material (Miles 1999). Sealing the wound with sheath saliva also prevents phloem sap losses and a drop in phloem turgor pressure. This is important since the loss of turgor pressure in the sieve elements (SE) may also induce the wounding response (Will *et al.* 2007/ and references). The small stylet volume and phloem sap flow regulation during feeding prevents a drop in SE turgor pressure (Dixon 1987; Will & van Bel 2006). Turgor pressure is also maintained during stylet penetration of the SE, thereby preventing the activation of stretch-activated calcium channels (Will & van Bel 2006).

Enucleated phloem SE was long thought to function as transporters of nutrients from source to sink tissues. However, this system's role in transport and signalling is far more complex. Molecules involved in plant defence (antioxidants, protease inhibitors, *etc.*), signalling (small RNAs, PK, SA, JA, *etc.*), and various macromolecules (mRNA and proteins), are all translocated *via* this system (Imlau *et al.* 1999; Ruiz-Medrano *et al.* 1999; Yoo *et al.* 2002; Lough & Lucas 2006; Gaupels *et al.* 2008; Kehr & Buhtz 2008; Le Hir *et al.* 2008). There is even indications of controlled transport of macromolecules to specific destinations in the plant (Aoki *et al.* 2005). Furthermore, proteins involved in RNA-binding, mRNA translation, macromolecule trafficking, *etc.* were found in phloem (Lin *et al.* 2008), thus suggesting protein synthesis and turnover – something not thought possible in SEs lacking nuclei and ribosomes (Yoo *et al.* 2002). The complexity of the system further increases when one considers that all the loading and unloading of phloem is done *via* the companion cells (Brett & Waldron 1990; Balachandran *et al.* 1997; Yoo *et al.* 2002; reviewed by Will & van Bel

2006; Le Hir *et al.* 2008; Braun & Slewinski 2009). Therefore, undetected breaching of this system poses numerous problems to aphids. Any breach will result in the rapid sealing of the SEs' sieve plates (as part of the wounding response). This is achieved by P-proteins (Dixon 1998; Miles 1999), or in the case of monocots like the Poaceae, by proteinaceous inclusions (bodies) of the SE plastids (Eleftheriou 1990; Miles 1999; Dinant *et al.* 2003; Will & van Bel 2006). In wheat the plastid envelope ruptures in heavily damaged sieve tubes, releasing the proteinaceous inclusions which seal the sieve-plate pores, which is followed by callose deposition (Bornman & Botha 1973; Brett & Waldron 1990; Eleftheriou 1990). Proteinaceous inclusion blocking is a fast reaction to wounding, while callose plugging of the sieve plates is thought to be a slower process (Will & van Bel 2006).

It has been suggested that the watery saliva prevents these blockages of the sieve plates when the stylet first break through the phloem CW (Dixon 1998). Calcium binding domains on salivary proteins suggest that they interact with Ca^{2+} on a molecular level, thereby preventing the calcium-initiated SE blocking (Will *et al.* 2007). The initial watery saliva injection may inhibit protein coagulation/redox-dependent precipitation, thus preventing blockages (Miles 1999). Aphid watery saliva can also unplug blocked SE (Will *et al.* 2007). However, stylectomy showed that phloem flow eventually stop, usually faster in a resistant cultivar (Miles 1999).

Salivary enzymes: Neutralising potential signals and suppressing defences

Aphid saliva is thought to play a role in overcoming toxic phenolics, ROS, wounding and defence responses. Enzyme composition usually differs between the proteinaceous and watery saliva, though the same enzymes might be in both (Miles 1967; 1968; Miles & Peng 1989; Urbanska *et al.* 1998; Miles 1999; Cherqui & Tjallingii 2000). Hydrolases, involved in degrading CW polysaccharides and oxidoreductase, in disrupting the plant's redox balance, are usually part of aphid salivary enzymes (Miles 1999). Pectinases and other polysaccharide depolymerising enzymes could actually counter wound-induced HR by pre-empting plant pectinases, *i.e.* by converting pectin or pectin-derived signals/elicitors to non-functional messengers that do not induce defence responses. An increase in salivation while probing, without HR, supports this theory (Miles 1999). Aphid pre-conditioning of hosts suggests that watery saliva have a greater effect on the surrounding phloem (Miles 1999). These systemic effects could possibly be attributed to watery saliva movement though the vascular system (Schotzko & Smith 1991; Rafi *et al.* 1996; Miles 1999). Oxidases are likely candidates since they move faster than reducing systems can counter their action through sensitive tissue (Miles 1999).

Salivary enzyme composition have led to the "redox-hypothesis" (Miles & Oertli 1993). This hypothesis suggest that the cellular redox homoeostasis, *i.e.* the soluble redox couples like NADPH, AA and GSH, are regulated by the aphid's salivary enzymes (Miles & Oertli 1993; Foyer & Noctor 2009). Plant phenolics are usually maintained in a reduced state by antioxidants in the cell (*Figure 2.5*). During aphid probing, plant cells oxidise and mobilise phenolic compounds in response to wounding (Miles & Oertli 1993). Plants use polyphenol oxidases (PPO or catechol oxidase) in the plastids to catalyse the oxidation of phenols to quinines (He *et al.* 2009). Phenolic monomers, and their oxidised quinones, are either toxic or may act as feeding deterrents to aphids (Miles & Oertli 1993; Grayer *et al.* 1994). Quinones redox cycling also play a role in defence: reducing equivalents are redirected to superoxide (and hydrogen peroxide) production, which may form either directly or indirectly part of the HR response (Cape *et al.* 2006 & references within). The final oxidised products, *i.e.* polymers and protein-phenol conjugates, are non-toxic cell sealants (Miles & Oertli 1993). Effective sealing and defence response in plant cells demand a controlled rate of phenolic oxidation (*Figure 2.6*) (Miles & Oertli 1993). Aphid salivary enzymes alter the redox state of cells (or SE), thereby interfering with coagulation and oxidation reactions of the phenolic compounds and proteins. This leads to faster oxidation, unordered sieve plate sealing and a reduction in the defence response (Miles & Oertli 1993).

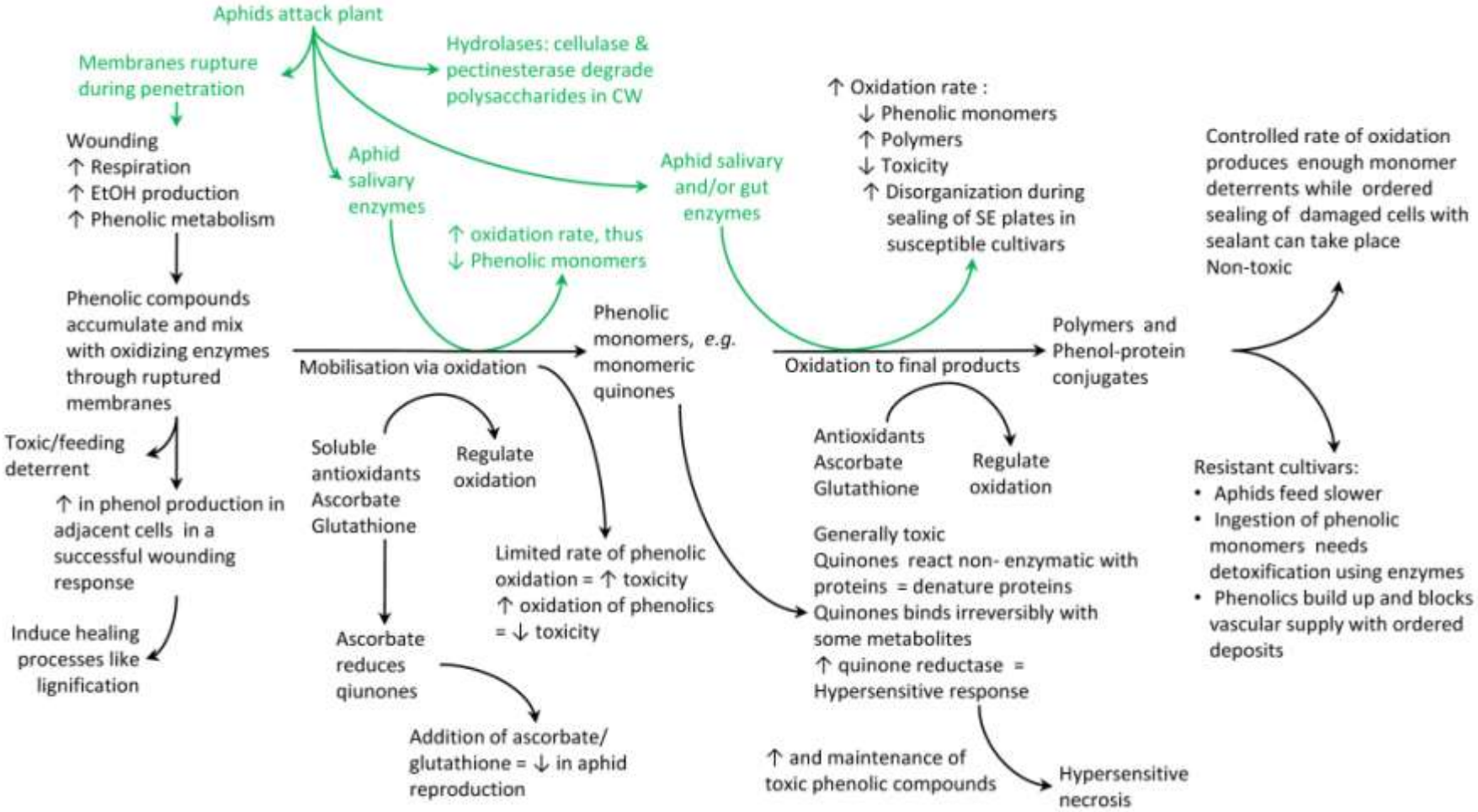


Figure 2.5 An overview of aphid feeding (green) and the responses of the host. Antioxidants in the plant regulate the redox condition of phenolic compounds in the cells by keeping them in a reduced form under normal cellular conditions, while regulating oxidation rates when wounded. Aphid salivary enzymes (green) alter the redox poise of the cell, thereby enhancing the oxidation of toxic phenolics/quinones into non-toxic phenol-protein conjugates and polymers. Refer to *Figures 2.6 & 2.7*. Compiled from Miles & Oertli 1993; Jarabak *et al.* 1997 and Ni *et al.* 2000.

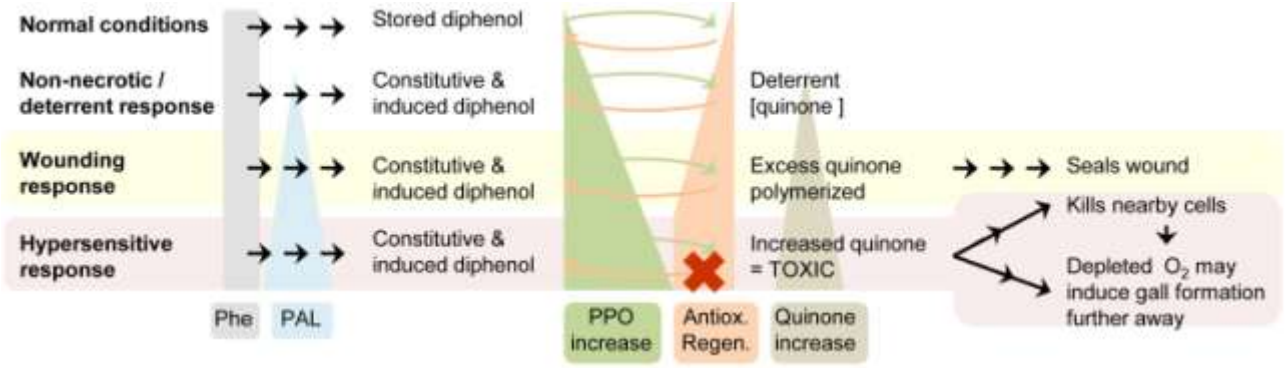


Figure 2.6 Plant defence responses during unsuccessful aphid infestation. Increases in expression or substrate of phenylalanine ammonia lyase (PAL) and polyphenol oxidase (PPO) lead to increases in quinones. The antioxidant regeneration system (Antiox Regen) regulates the conversion from diphenols to quinone, thereby keeping the supply of toxic compounds steady. When antioxidants are used faster than they can be replenished (X), quinones reach the HR level. Phe, phenylalanine. Also see *Figure 2.7*. Compiled from Miles & Oertli 1993 and Miles 1999.

Salivary enzymes of the RWA

Macerated RWA extracts showed no detectable polysaccharidase, phospholipase A, lipase or protease activity, but low levels of aminopeptidase (possibly from lysosomes) were observed (Fouche *et al.* 1984). In another study, total RWA extracts, confirmed with aphid head extracts, showed no amylase, pectinase or peroxidase (PX) activities, though cellulase, pectinesterase (PE), ascorbate oxidase (AO), superoxide dismutase (SOD), catechol oxidase (polyphenol oxidase, PPO) and catalase (CAT) activities were present (*Figure 2.5*) (Ni *et al.* 2000). The hydrolases (cellulase and PE) are involved with CW degradation while the oxidoreductase (AO, CAT, PPO and SOD) disrupts the plant's redox balance (Ni *et al.* 2000). Interestingly, CAT was only found in the RWA while PX was only present in *Rhopalosiphum padi* (Ni *et al.* 2000). The authors suggested that this difference in salivary composition could probably explain the different symptoms observed in susceptible hosts' in responses to these two aphids. CAT are probably more active as it needs no donor like PX when converting H_2O_2 to water and oxygen (Ni *et al.* 2000).

The possible effects of the RWA salivary enzymes on the plant are summarised in *Figure 2.7*. However, little is known about the role of aphid salivary AO. Recent studies on plant AO could possibly shed some light on the role of this enzyme in aphid-plant interactions.

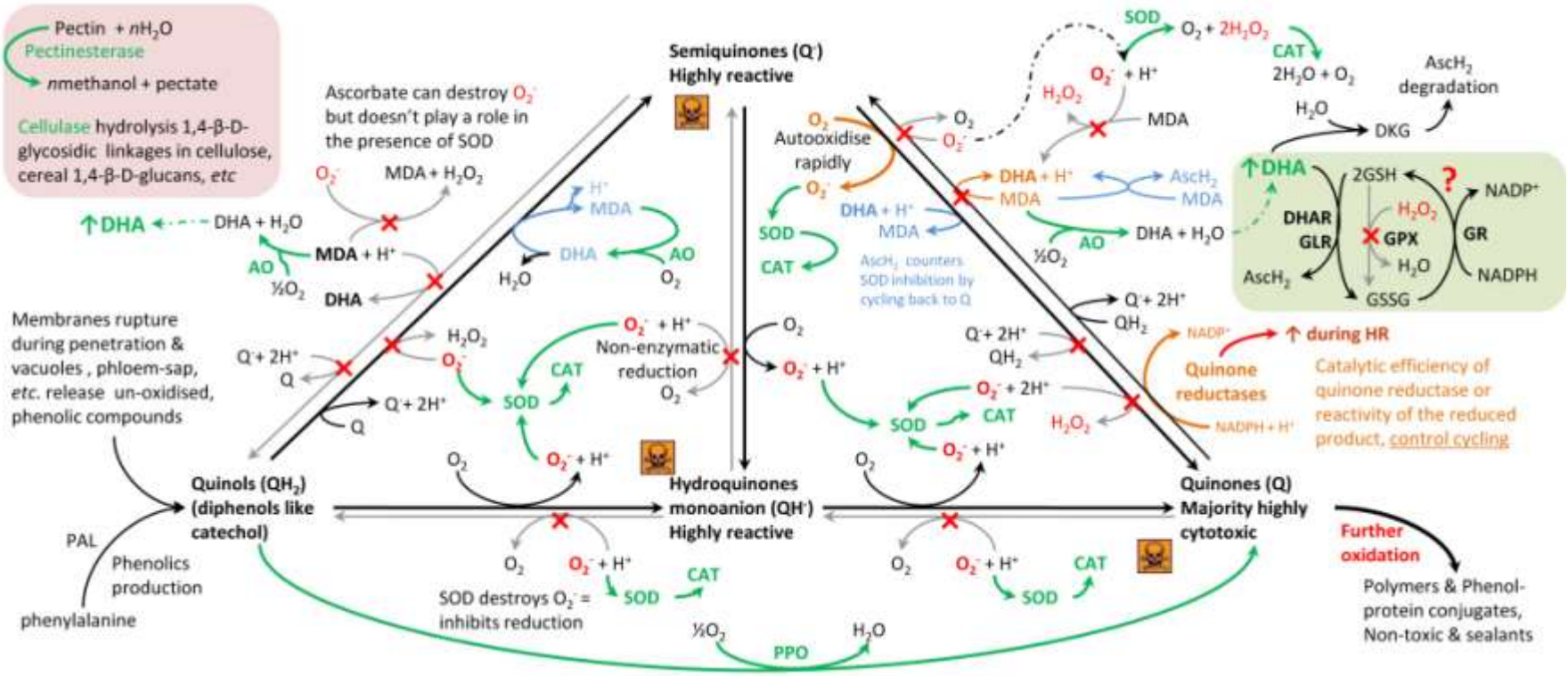


Figure 2.7 Known RWA salivary enzymes and their role in upsetting cellular and apoplastic redox poise in the host. In healthy tissue phenolics are kept in a reduced state by antioxidants and the continual reductase renewal. During wounding, membranes rupture and phenolics are released/produced. Quinone redox-cycling start (orange) and produce O_2^- and H_2O_2 resulting in ROS and HR. Quinone reductase transcription is up-regulated during this HR response, producing more ROS species. Toxic semi- and hydroquinones are also being recycled, thereby controlling the oxidation of quinones to non-toxic substances. Ingestion of un-oxidised phenolics/quinones, in the absence of the renewable antioxidants, autoxidise with proteins in the gut to form toxic substances. Salivary oxidoreductases shift the redox poise of the apoplast/cytosol and may even detoxify phloem sap *en route* to the gut. $AscH_2$, ascorbate; AO, ascorbate oxidase; APX, ascorbate peroxidase; CAT, catalase; DHA, dehydroascorbate; DHAR, DHA reductase; DKG, 2,3-diketo-l-gulonic acid; GR, glutathione reductase; GSH, reduced glutathione; GSSH, oxidised glutathione; GPX, glutathione peroxidase MDA, monodehydroascorbate; PPO, polyphenol oxidase (catechol oxidase); SOD, superoxide dismutase. (Miles & Oertli 1993; Jarabak *et al.* 1997; Miyake & Kurata 1998; Roginsky *et al.* 1999; Ni *et al.* 2000; Cape *et al.* 2006; Foyer & Noctor 2009); IUBMB Enzyme Nomenclature EC 1.6.5.5; Pfam PF01095; www.brenda-enzymes.org (EC3.1.1.11 – pectinesterase).

Ascorbate oxidase in plants is localised to the apoplast/CW (Pignocchi *et al.* 2003; Fotopoulos *et al.* 2006; Fotopoulos *et al.* 2008). The enzyme catalyses the oxidation of AA ($AscH_2$) to monodehydroascorbate (MDA) that disproportionate to AA and dehydroascorbate (DHA) (Fotopoulos *et al.* 2006; Fotopoulos *et al.* 2008; Foyer & Noctor 2009). Studies in plants showed that an increase in AO leads to a lower apoplastic AA (AA_a) *in vivo*, but has no effect on normal AA recycling gene expression (Pignocchi *et al.* 2003; Yamamoto *et al.* 2005; Fotopoulos *et al.* 2006). However, DHA levels increased though the overall AA, *i.e.* $AscH_2$, DHA, MDA, remained the same (Pignocchi *et al.* 2003). This suggests that the transport system could not cope with the high rate of AA oxidation, even though it tried to address the redox poise in the apoplast by increasing the total AA levels (Pignocchi *et al.* 2003). Plants with over-expressed AO subjected to oxidative stress, like an infection or environmental stress, also appear more sensitive to these stresses, with higher chloroplast loss and higher *in situ* H_2O_2 (Yamamoto *et al.* 2005; Fotopoulos *et al.* 2006; Fotopoulos *et al.* 2008). This suggests that an increased AO expression in plants reduce their capacity to up-regulate defences against ROS. Furthermore, Ca^{2+} channel expression decrease under higher AO expression levels, which could have major influences on signalling, elicitors, ROS-induced regulation, *etc.* (Fotopoulos *et al.* 2006). Over-expression of AO in tobacco plants led to partial stomatal closure, increase in water content, decline in peroxide-scavenging enzyme activities and an increase in leaf ABA levels (*Figure 2.6*) (Fotopoulos *et al.* 2008). The authors also suggested that AA_a perceived environmental cues and then use DHA to regulate stomatal dynamics (Fotopoulos *et al.* 2008). Though all these results are based on over-expressed apoplastic plant AO, it does suggest a possible role for salivary AO in the host. The salivary AO could possibly suppress the wounding response initiated by Ca^{2+} movement/signalling, reduce the plants ability to launch defence responses and, depending on its location, could possibly play a role in diet enrichment by increasing AA. The increase observed in apoplastic H_2O_2 under higher AO levels is probably being addressed by RWA salivary CAT.

RWA-wheat interaction

Constitutive plant defences in resistant individuals cause the RWA to engage for longer periods in pre-penetration behaviour (Kindler *et al.* 1992; Webster *et al.* 1993). Epicuticular waxes are known to play a role in plant defences (Powell *et al.* 1999; Peeters 2002), however, cereal leaf waxes have little effect towards RWA resistance (Ni *et al.* 1998; Bahlmann 2002; Bahlmann *et al.* 2003). Furthermore, the average frequencies of leaf probing and the duration thereof, were similar in

both susceptible and resistant cultivars (Kindler *et al.* 1992). Conflicting results exist for trichomes and their influences on aphid feeding (Bahlmann 2002; Peeters 2002; Bahlmann *et al.* 2003). Trichome length, but not density, was shown to be directly linked to RWA resistance in some cultivars (Ni & Quisenberry 1997). Trichomes are usually located on leaf veins, possibly hindering aphids in finding their favourite feeding sites (Ni & Quisenberry 1997; Ni *et al.* 1998; Bahlmann 2002). However, the RWA was shown to have no preferential penetration sites or -patterns on wheat leaves but rather follows a randomised approach (Fouche *et al.* 1984).

The RWA prefers to feed from the thin walled phloem sieve tubes of the leaf basis and new growth, where the availability of photosynthetic assimilates are greater (Kriel *et al.* 1986; Botha & Matsiliza 2004). Aphids cease their feeding on resistant cultivars or non-host plants soon after the phloem element is penetrated, responses attributed to lectins binding to the chitin that surrounds the food canal and foregut of the aphid (Dixon 1998). However, other products, as part of the host defence response upon the recognition of a gene-for-gene interaction, might also deter aphid feeding (Kaloshian 2004; Botha *et al.* 2006; Smith & Boyko 2007). The average duration of salivation in feeding RWA was shown to be the same on both resistant and susceptible wheat cultivars, but less frequent in susceptible lines (Kindler *et al.* 1992). On non-hosts, the RWA salivates more and ingests less, taking up to four times longer to locate the phloem (Girma *et al.* 1992). RWAs also spend significantly longer time feeding on phloem of susceptible accessions than on resistant lines. Together with longer periods in the pre-penetration stage, these suggest that RWAs turned to non-phloem feeding for survival on resistant cultivars (Kindler *et al.* 1992). This hypothesis is also supported by a study of RWA feeding on barley (Webster *et al.* 1993).

There is a decrease in population size and an increase in alate female numbers when wheat becomes unsuitable for RWA feeding at the ear stage (see *Biology*) (Kriel *et al.* 1984; Walters *et al.* 1984). Aphid numbers peak at wheat growth stage 5, when photosynthetic rates are extremely high and products are directed towards the flag leaf and developing ears (Kriel *et al.* 1984). RWA "preconditioning" of its host over time, makes it easier to feed on the same plant for subsequent generations (Schotzko & Smith 1991; Rafi *et al.* 1996), thus two nymphs feeding on the same leaf results in faster development than if the two were to settle on different leaves (Rafi *et al.* 1996; Qureshi & Michaud 2005)

Feeding by the RWA causes leaf curling and the thickening of the epidermal cells in all lignin containing tissues (see *Symptoms*). No such reactions are induced by the same method of feeding as used by other aphids like *R. padi* and *S. graminum* (Fouche *et al*. 1984). Lignification after RWA feeding, as part of the HR, starts at the middle lamella and makes the CW impenetrable (Brett &

Waldron 1990; Mitchell *et al.* 1994; Mitchell *et al.* 1999). Initially, at these sites, it was hypothesized that a phytotoxin is injected that prevents the unfurling of the leaf (Smith *et al.* 1991). This phytotoxin was thought to be involved in the breakdown of the chloroplast (Fouche *et al.* 1984), with chlorophyll reductions of up to 85 % recorded (Kruger & Hewitt 1984), and thus the yellow streaking of the leaves. However, an ultra-structural study on the effects of RWA feeding on a susceptible barley cultivar, suggested that both xylem and phloem (parenchyma, thick-walled sieve tubes and companion cells) were extensively damaged during probing, thus leading to the typical RWA symptoms (yellow streaks, leaf rolling, *etc.*) (Saheed *et al.* 2007b). Similarly, severe RWA damage occurred to longitudinal vein phloem, including to most of the leaf phloem vascular system, of a susceptible wheat cultivar (Botha & Matsiliza 2004). Massive damage that can be observed as wound-related callose, is thought to re-route assimilates (with little or no transverse or longitudinal movement of assimilates) to form sinks at the aphid feeding sites (Botha & Matsiliza 2004). Furthermore, the RWA taps xylem to obtain water and in the process injects large amounts of watery saliva into the vessels. This saliva is electron-dense, smooth to amorphous, and lines the metaxylem thereby sealing the pit membranes between elements and those that connect the xylem vessels and xylem parenchyma (Saheed *et al.* 2007b). Callose further blocks the plasmodesmata of the phloem parenchyma elements and sieve tube-companion cells. A lot of the cells showed different degrees of plasmolysis which can be a result of the probing behaviour of the RWA that puncture cells. This results in oxidative stress with the occurrence of necrosis and chlorosis (Saheed *et al.* 2007b). The damage and sealing of the vascular system would interfere with apoplasmic and symplasmic transfer of water and nutrients, thus the white/yellow streaks and leaf rolling observations in RWA susceptible cultivars (Saheed *et al.* 2007b). These results are in contrast with an earlier study, conducted on *S. graminum* and the RWA, using isotope labelling (Burd 2002). Feeding of both aphids cause similar responses in the plant (Sandstrom *et al.* 2000). However, unlike greenbug, the RWA does not affect the rate of phloem loading and also has no or little influence on phloem translocation that occurs at the feeding site (Burd 2002). This suggests that the RWA feeding does not influence phloem function around the feeding site.

RWA feeding reduces chlorophyll fluorescence and photosynthetic rates of susceptible wheat and barley cultivars (Burd & Elliott 1996; Haile *et al.* 1999). A reduction in chlorophyll *a* and *b* content in susceptible cultivars, but not in resistant lines (Kruger & Hewitt 1984; Burd & Elliott 1996), or electron transport inhibition of the photochemical reaction centre (Miller *et al.* 1994; Burd & Elliott 1996), could explain these observations. Leaves treated with *D. noxia* extracts (2 h) showed that the chloroplasts initially lose their ordered arrangement next to the CW and are

Table 2.2 Morphological and physiological changes in cereals over time in response to RWA infestation.

Time interval (hours)	***Susceptible cultivars infested***	***Resistant cultivars infested***
2	Chloroplasts loosely ordered arrangement next to cell wall. Chloroplasts distributed throughout the cell (*in vitro*)[1]	
4	Chloroplast membranes disintegrate and the contents spread throughout the cell (*in vitro*)[1]	
5	Chloroplast homogenise and are no longer visible (*in vitro*)[1]	
6	Ethylene production begins (1800 nL h^{-1} g^{-1} fresh mass, control 400)[2]	Ethylene production begins (600 nL h^{-1} g^{-1} fresh mass (control 800))[2]
12	Ethylene production[2]	Ethylene production[2]
18	Ethylene production peak[2]	Ethylene production peak[2]
24 (1 d)		Chitinase activity is induced in TugelaDN, MolopoDN & BettaDN[4]
48 (2 d)	No increased inter- and intracellular β-1,3-glucanase activity observed in NILs: Tugela, Betta, Molopo[6]	Increased inter- and intracellular β-1,3-glucanaseactivity observed in *Dn1*-resistant cultivars: TugelaDN, BettaDN, MolopoDN, with 7 isoforms of the enzyme up-regulated intercellular which was background specific[6] Chitinase activity further increases for Tugela, Molopo & Betta[4]
72 (3 d)	Plasma lemma situated further from cell wall, convoluted appearance (*in vivo*)[1]	Ethylene production lower[2] Peroxidase activity starts to increase (TugelaDN,

	Ethylene production lowers[2]	MolopoDN & BettaDN)[4]
	Peroxidase activity starts to increase (Tugela, Molopo & Betta) [4]	Chitinase activity keeps increasing[4]
96 (4 d)	More prominent: plasmalemma situated further from cell wall, convoluted appearance. Some chloroplast membranes showed similar changes. Few grana were slightly swollen (*in vivo*) [1]	Ethylene patterns are the same till day 14[2]
		Peroxidase activity almost double that of the susceptible lines (TugelaDN, MolopoDN & BettaDN) [4]
	Reduction in chlorophyll a and b and total chlorophyll content[5]	Chitinase activity keeps increasing[4]
	Ethylene patterns are the same till day 14[2]	
	Peroxidase activity increases (Tugela, Molopo & Betta) [4]	
120 (5 d)	Wheat (average RWA host): Grana obviously swollen, frets showed swelling, chloroplast shape normal (*in vivo*) [1]	Peroxidase activity increases (TugelaDN, MolopoDN & BettaDN) [4]
	Chloroplast regarded as intermediately susceptible	Chitinase activity keeps increasing[4]
	Barley (good RWA hosts):	
	Convoluted cell membranes, swollen grana and fret disrupted, many osmiophilic globuli, many chloroplast disrupted completely	
	Chloroplast regarded as highly susceptible[1]	
	Oats (poor RWA hosts):	
	No changes in ultra-structure. Chloroplast regarded as resistant (*in vivo*) [1]	
	Peroxidase activity increases (Tugela, Molopo & Betta) [4]	
144 (6 d)	Chloroplast situated further from plasma lemma and cell wall. Grana-fret system swollen[1]	Peroxidase activity increasing (TugelaDN, MolopoDN & BettaDN) [4]
	Peroxidase activity increases for Molopo & Betta, levelling off for Tugela [4]	Chitinase activity peak for MolopoDN[4]
		Proteins (PR?) in IWF detected in resistant Dn1 lines

	Chitinases activity lowest for Molopo & Betta, peak for Tugela [4] Induction of a 53 kD protein in Stephens and reduction of 47, 48 and 49 kD proteins[6]	(upper 8cm used) (*in vivo*) [1]
168 (7 d)	Chloroplast situated further from plasma lemma and cell wall, Grana-fret system swollen. Many osmiophillic globuli present (*in vivo*) [1] Strong burst of chitinase[2] Peroxidase activity increases for Molopo & Betta, decreasing for Tugela [4] Chitinases increasing for Molopo & Betta, decreasing for Tugela [4]	Peroxidase activity increasing (TugelaDN, MolopoDN & BettaDN) [4] Chitinase activity decreasing for TugelaDN & MolopoDN, increasing for BettaDN[4]
192 (8 d)	Double chloroplast membranes and plasma lemma extremely convoluted, Grana-fret system disrupted totally, many osmiophillic globuli present (*in vivo*) [1] Peroxidase activity increases for Molopo & Betta, decreasing for Tugela [4] Chitinases increasing for Molopo & Betta, decreasing for Tugela [4]	Peroxidase activity increasing (TugelaDN, MolopoDN & BettaDN) [4] Chitinase activity decreasing for TugelaDN & MolopoDN, increasing for Betta*Dn*[4]
216 (9 d)	No internal chloroplast definition remains (*in vivo*) [1] Peroxidase activity increases for Molopo & Betta, decreasing for Tugela [4] Chitinases increasing for Molopo & Betta, decreasing for Tugela [4]	Peroxidase activity increasing (TugelaDN, MolopoDN & BettaDN) [4] Chitinase activity decreasing for TugelaDN & MolopoDN, increasing for BettaDN[4]
240 (10 d)	Chloroplast membranes disintegrate, "osmiophillic globuli" present	Induction of chitinase[2] Peroxidase activity increasing for TugelaDN, peaking for

	Purple (cyanidine containing cells) and white (bleached cells) streaking occurs on leaves—NO organelles left in cells (*in vivo*) [1] Peroxidase activity increases for Molopo & Betta, decreasing for Tugela [4] Chitinases increasing for Molopo & Betta, decreasing for Tugela [4]	MolopoDN & BettaDN [4] Chitinase activity decreasing for TugelaDN & MolopoDN, peaking for BettaDN[4]
264 (11 d)	Reductions of 47, 48 and 49 kD proteins[6]	Induction of a 53 kD protein and reduction of a 47, 48 and 49 kD proteins[6]
288 (12 d)	Peroxidase activity increase for Molopo & Betta, decreasing for Tugela [4] Chitinases increasing for Molopo & Betta, decreasing for Tugela [4]	Peroxidase activity increasing for TugelaDN, decreasing for MolopoDN & BettaDN [4] Chitinase activity increasing for TugelaDN, decreasing for MolopoDN & BettaDN[4]
336 (14 d)	Peroxidase activity still increase for Molopo & Betta, decreasing for Tugela [4] Chitinases peaking for Molopo & Betta, decreasing for Tugela[4]	Peroxidase activity increasing for TugelaDN, decreasing for MolopoDN & BettaDN [4] Chitinase activity peaking for TugelaDN, decreasing for MolopoDN & BettaDN[4]

Compiled from: Fouche *et al.* (1984)[1], Nagel (1995)[2], Botha *et al.* (1998)[3], Van der Westhuizen and Pretorius (1996)[4], Burd and Elliott (1996)[5]; Van der Westhuizen *et al.*1998a[6]

distributed through the cell content. After 4 hours (h) the chloroplast membranes started disintegrating and the contents spread throughout the cell, with no visible chloroplast left after 5 h (*Table 2.2*) (Fouche *et al.* 1984). RWA feeding seems to target the electron transport of photosystem II (PSII), leading to a decrease in the integrity of the thylakoid membrane system, which results in changes in the photosynthetic ability (Kruger & Hewitt 1984; Burd & Elliott 1996; Haile *et al.* 1999). However, Miller *et al.* (1994) observed no differences in PSII effectiveness of resistant and susceptible barley cultivars, but a decrease in the recovery of the quinone after illumination in susceptible cultivars. This suggests antioxidant involvement. Indeed, CAT (together with other oxidoreductases) in RWA saliva is thought to be responsible for the necrotic streaking observed in susceptible cultivars and the loss of chloroplast integrity (Ni *et al.* 2000). CAT, unlike PX, needs no donor when dealing with H_2O_2 and together with other oxidoreductases, they change the redox state of the cell, thus affecting the chloroplast electron transport chain (Ni *et al.* 2000). Rubisco subunits (LSU and SSU) also decrease in infected accessions (with the SSU more susceptible to RWA infestation) (Botha & van der Westhuizen 1992; Rafi *et al.* 1996).

RWA infestations also result in a proline increase in susceptible wheat lines faster than it does in resistant lines. However, during later stages, proline concentrations in resistant lines rapidly exceed and are maintained at higher levels than those of the susceptible lines (Botha & van der Westhuizen 1992). Moisture content observations of plants after RWA infection give conflicting results (Kruger & Hewitt 1984; Botha & van der Westhuizen 1992). Various other enzymes are also induced in resistant cultivars during RWA feeding (*Table 2.2 & 2.3*). RWA infestations also increase the accumulation of intercellular β-1,3-glucanase in resistant wheat cultivars (*Dn1* resistant cultivars), though cultivar dependent, but not in susceptible cultivars (Van der Westhuizen *et al.* 1998a; Mohase & van der Westhuizen 2002). Glucanases hydrolyse glucosidic linkages found in glucans, like callose (β-1,3,-glucan), which forms part of CW (Van der Westhuizen *et al.* 1998a). These CW glycoproteins can also activate plant defence responses. Indeed, glycoproteins in resistant wheat cultivars were shown to elicit defence responses after RWA infestation (Mohase & van der Westhuizen 2002). This suggest that RWA saliva could either release eliciting CW components in resistant cultivars (*via* salivary hydrolases) or could contain eliciting compounds that may activate defence responses, thereby resulting in increased PX, β-1,3-glucanase and PR-proteins, like chitinase, expression (Van der Westhuizen *et al.* 1998a; b; Mohase & van der Westhuizen 2002).

Studies on the apoplast of RWA infested cultivars showed many defence related genes up-regulated in the apoplasts of these cultivars after infestation (*Table 2.3*). The genetic background of cultivars may play a significant role in the expression of the same enzymes, as was shown for

peroxidase and chitinase expression in the apoplast of the near isogenic lines (NIL) BettaDN and TugelaDN (Van der Westhuizen *et al.* 1998b). Intercellular washing fluid studies of TugelaDN revealed increases in chitinase activity in the symplastic tissue (Nagel 1995; Botha *et al.* 1998). Nagel (1995) showed these increases in the apoplast to be a result of endochitinase induction in TugelaDN. Eight chitinase isoforms were present in the Tugela and TugelaDN lines, with two extra chitinases induced after RWA infestation. A four-fold up-regulation of a third chitinase was also present in the resistant TugelaDN (Nagel 1995). Furthermore, ethylene induced a similar response to RWA feeding in the susceptible Tugela. Wounding had no effect on the chitinase activity in both lines. This indicated that ethylene and RWA infestations are similar in their effects on chitinase expression, but that wounding could not account for the observed inductions (Nagel 1995; Botha *et al.* 1998). Nagel (1995) further speculated that a chitosan-like substance could be responsible for the elevated chitinase expression observed after RWA infestation, with ethylene and salicylic acid (SA) acting as secondary elicitors. Both SA and ethylene are induced by chitosan (Nagel 1995). An elicitor-induced study, using phenylalanine ammonia lyase- (PAL), peroxidase-, chitinase- and glucanase-activities as measure, demonstrated that chitin and chitosan (and derivatives thereof) do elicit a non-selective defence response in ‘Tugela’ and ‘TugelaDN’ (Van der Westhuizen & Oberholster 1996). Lapitan *et al* (Lapitan *et al.* 2007a) confirmed the latter results when they showed that RWA specific responses were only obtained from the protein portion of aphid extracts, but not from chitin. This supports a gene-for-gene interaction with a protein elicitor (Lapitan *et al.* 2007a).

Table 2.3 Proteins present in the apoplast and vascular system with special focus on wheat.

Protein sizes	***Enzyme and description***
100 kD, 70 kD	Endoproteases (serine-like proteases)[1]
28-33, 22-24, 18.5-19.5 and 15.5-17 kD	Possible PR-proteins. From 3 cultivars with *Dn1* gene[2]
Unknown	Pathogenesis-related (PR) proteins[3]
36, 34, 27, 22 kD (apoplast) 61, 54, 49.5, 35,1 kD (vascular)	Chitinases (hydrolyse chitin) 36, 34 & 22 kD chitinases were induced in a resistant cultivar after RWA infestation[1,4,5,6]
Unknown	ß-1,3-glucanases[1]
100 kD, 70 kD, 63 kD, 55 kD, 40 kD fragments	No gluconase-6-phospate activity, caseinase activity or proteases activity[1]
Unknown	Peroxidase[6]

[1]Pinedo *et al.* (1993), [2]Van der Westhuizen & Pretorius (1996), [3]Bowles (1990), [4]Nagel (1995), [5]Botha *et al.* (1998), [6]Van der Westhuisen *et al.* (1998b).

Conclusion

The development of additional RWA biotypes have open new research areas in RWA-wheat research, especially on how new biotypes may develop in reaction to plant defence strategies. Additionally, RWA biotypes also allow us to investigate the RWA-endosymbiont interaction and the role of the endosymbiont(s) in aphid biotype development. The current theory on RWA-*B. aphidicola* interactions suggests that the RWA is increasingly less dependent on its bacterial endosymbiont, especially regarding essential amino acid production. Furthermore, increases in phloem leucine and tryptophan levels of susceptible wheat cultivars during RWA feeding, implies a highly specialized aphid-plant interaction with a lesser dependency on the endosymbiont. Since resistance wheat cultivars have lower RWA induced leucine and tryptophan levels, the RWA-*B. aphidicola* theory may suggests a change in the aphid that allows biotypes to induce higher levels of the required essential amino acids in resistant wheat cultivars as well. However, it was the aphid-bacterium symbiosis that first allowed phloem feeding. Therefore, if the aphid reverts back to utilizing *B. aphidicola*, which still has the means to produce the required essential amino acids, the RWA may feed on resistant cultivars. This hypothesis suggests that continuous selective pressures of resistant wheat cultivars could have selected for RWA biotypes with more beneficial essential amino acid producing *B. aphidicola* or

an advantages change in endosymbiont composition — in both cases sufficient levels of essential amino acids are produced to maintain a RWA population, and thus overcome host resistance. Furthermore, the various molecular studies on plant responses to RWA infestation have not yet given a definitive answer. Questions still remain on the details of the plant's responses to RWA infestation and how (if at all) different modes of resistance, *i.e.* antibiosis, tolerance and antixenosis, differ in their reaction to infestation. Currently, high-throughput array analyses are used in answering these questions. In order to obtain the differentially regulated genes with these methods, various statistical analyses are done on the datasets. The influences of different statistical normalization and background corrections methods on the dataset may have a large effect on the genes eventually deemed differentially regulated. This hypothesis is tested in the second part of this study.

References

Aalbersberg, Y. K., M. C. Walters and N. J. Van Rensburg (1984) The status and potential biological control studies on *Diuraphis noxia* (Aphididae). Technical Communication Department of Agriculture Republic of South Africa. M. C. Walters. Bloemfontein. **191:** 44-46.

Aoki, K., N. Suzui, S. Fujimaki, N. Dohmae, K. Yonekura-Sakakibara, T. Fujiwara, H. Hayashi, T. Yamaya and H. Sakakibara (2005) Destination-selective long-distance movement of phloem proteins. *The Plant Cell* **17**: 1801-1814.

Archer, T. L. and E. D. Bynum (1992) Economic injury level for the Russian wheat aphid (Homoptera: Aphididae) on dryland winter wheat. *Journal of Economic Entomology* **85**: 987-992.

Bahlmann, L. (2002) Factors affecting the resistance mechanisms of the Russian wheat aphid (*Diuraphis noxia*) on wheat. Faculty of Natural and Agricultural Sciences. Pretoria, University of Pretoria. **MSc:** 124.

Bahlmann, L., P. Govender and A.-M. Botha (2003) Leaf epicuticular wax ultrastructure and trichome presence on Russian wheat aphid (*Diuraphis noxia*) resistant and susceptible leaves. *African Entomology* **11**: 59-64.

Baier, M. and K.-J. Dietz (2005) Chloroplast as source and target of cellular redox regulation: a discussion on chloroplast redox signals in the context of plant physiology. *Journal of Experimental Botany* **56**: 1449-1462.

Baker, D. A., H. D. Loxdale and O. R. Edwards (2003) Genetic variation and founder effects in the parasitoid wasp, *Diaeretiella rapae* (M'intosh) (Hymenoptera:

Braconidae:Aphidiidae), affecting its potential as a biological control agent. *Molecular Ecology* **12**: 3303-3311.

Balachandran, S., Y. Xiang, C. Schobert, G. A. Thompson and W. J. Lucas (1997) Phloem sap proteins from *Cucurbita maxima* and *Ricinus communis* have the capacity to traffic cell to cell through plasmodesmata. *Proceedings of the National Academy of Sciences, USA* **94**: 14150-14155.

Baldwin, I. T. and C. A. Preston (1999) The eco-physiological complexity of plant responses to insect herbivores. *Planta* **208**: 137-145.

Basky, Z. (2003) Biotypic and pest status differences between Hungarian and South African populations of Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Homoptera:Aphididae). *Pest Management Science* **59**: 1152-1158.

Baumann, L., P. Baumann and N. A. Moran (1998) The endosymbiont (*Buchnera*) of the aphid *Diuraphis noxia* contains all the genes of the tryptophan biosynthetic pathway. *Current Microbiology* **37**: 58-59.

Baumann, L., P. Baumann, N. A. Moran, J. Sandstrom and M. L. Thao (1999) Genetic characterization of plasmids containing genes encoding enzymes of leucine biosynthesis in endosymbionts (*Buchnera*) of aphids. *Journal of Molecular Evolution* **48**: 77-85.

Baumann, P., L. C.-Y., L. Baumann, D. Rouhbakhsh, N. A. Moran and M. A. Clark (1995) Mutualistic associations of aphids and prokaryotes: biology of the genus *Buchnera*. *Applied and Environmental Microbiology* **61**: 1-7.

Beers, E. P. and J. M. McDowell (2001) Regulation and execution of programmed cell death in response to pathogens, stress and developmental cues. *Current Opinion in Plant Biology* **4**: 561-567.

Belefant-Miller, H., D. R. Porter, M. L. Pierce and A. J. Mort (1994) An early indicator of resistance in barley to Russian wheat aphid. *Plant Physiology* **105**: 1289-1294.

Bergeson, E. and F. J. Messina (1998) Effect of a co-occurring aphid on the susceptibility of the Russian wheat aphid to lacewing predators. *Entomologia Experimentalis et Applicata* **87**: 103-108.

Birkle, L. M., L. B. Minto and A. E. Douglas (2002) Relating genotype and phenotype for tryptophan synthesis in an aphid-bacterial symbiosis. *Physiological Entomology* **27**: 302-306.

Bizzaro, D., M. Mandrioli, M. Zanotti, M. Giusti and G. C. Manicardi (2000) Chromosome analysis and molecular characterization of highly repeated DNAs in the aphid *Acyrthosiphon pisum* (Aphididae, Hemiptera). *Gentica* **108**: 197-202.

Blackman, R. L. and V. F. Eastop (2000) Aphids on the world's crops. An identification and information guide. London, Wiley.

Bornman, C. H. and C. E. J. Botha (1973) The role of aphids in phloem research. *Endeavour* **32**: 129-133.

Bosque-Perez, N. A., J. B. Johnson, D. J. Schotzko and L. Unger (2002) Species diversity, abundance, and phenology of aphid natural enemies on spring wheats resistant and susceptible to Russian wheat aphid. *BioControl* **47**: 667-684.

Botha, A.-M., L. Lacock, C. van Niekerk, T. M. Matsioloko, F. B. du Preez, S. Loots, E. Venter, K. J. Kunert and C. A. Cullis (2006) Is photosynthetic transcriptional regulation in *Triticum aestivum* L. cv. 'TugelaDN' a contributing factor for tolerance to *Diuraphis noxia* (Homoptera: Aphididae)? *Plant Cell Reports* **25**.

Botha, A.-M., M. A. C. Nagel, A. J. van der Westhuizen and F. C. Botha (1998) Chitinase isoenzymes in near-isogenic wheat lines challenged with Russian wheat aphid, exogenous ethylene, and mechanical wounding. *Botanical Bulletin of Academia Sinica* **39**: 99-106.

Botha, C. E. J. and B. Matsiliza (2004) Reduction in transport in wheat (*Triticum aestivum*) is caused by sustained phloem feeding by the Russian wheat aphid (*Diuraphis noxia*). *South African Journal of Botany* **70**: 249-254.

Botha, F. C. and A. J. van der Westhuizen (1992) Moleculêre biologie van Russiese koringluis weerstand. Bloemfontein, Departement Plantkunde en Genetika, Universiteit van die Oranje-Vrystaat**:** 1-75.

Botha, T. C. (1984) Aspects of the chemical control of *Diuraphis noxia* Technical communication Department of Agriculture Republic of South Africa. M. C. Walters. Bloemfontein. **191:** 63-66.

Bourne, H. R., D. A. Sanders and F. McCormick (1991) The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**: 117-127.

Bowles, D. J. (1990) Defense-related proteins in higher plants. *Annual Review in Biochemistry* **59**: 873-907.

Boyko, E. V., C. M. Smith, V. K. Thara, J. M. Bruno, Y. Deng, S. R. Starkey and D. L. Klaahsen (2006) Molecular basis of plant gene expression during aphid invation: wheat *Pto*- and *Pti*-like sequences are involved in interactions between wheat and Russian wheat aphid (Homoptera: Aphididae). *Journal of Economical Entomology* **99**: 1430-1445.

Braun, D. M. and T. L. Slewinski (2009) Genetic control of carbon partitioning in grasses: Roles of sucrose transporters and tie-dyed loci in phloem loading. *Plant Physiology* **149**: 71-81.

Brennan, E. B., S. A. Weinbaum and K. Pinney (2001) A new technique for studying the stylet tracks of homopteran insects in hand-sectioned plant tissue using light or epifluorescence microscopy. *Biotechnic & Histochemistry* **76**: 59-66.

Brett, C. and K. Waldron (1990) Physiology and biochemistry of plant cell walls. London, Unwin Hyman.

Brewer, M. J., T. Noma and N. C. Elliott (2005) Hymenopteran parasitoids and dipteran predators of the invasive aphid *Diuraphis noxia* after enemy introductions: temporal variation and implication for future aphid invasions. *Biological Control* **33**: 315-323.

Burd, J. D. (2002) Physiological modification of the host feeding site by cereal aphids (Homoptera: Aphididae). *Journal of Economic Entomology* **95**: 463-468.

Burd, J. D. and R. L. Burton (1992) Characterization of plant damage caused by Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **85**: 2017-2022.

Burd, J. D. and N. C. Elliott (1996) Changes in chlorophyll a fluorescence induction kinetics in cereals infested with Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **89**: 1332-1337.

Burd, J. D., D. R. Porter, G. J. Puterka, S. D. Haley and F. B. Peairs (2006) Biotypic variation among North American Russian wheat aphid (Homoptera: Aphididae) populations. *Journal of Economic Entomology* **99**: 1862-1866.

Butts, P. A. and K. W. Pakendorf (1984a) Seed treatment with sistemic insecticides for the control of *Diuraphis noxia* (Aphididae). Technical Communication Department of Agriculture Republic of South Africa. M. C. Walters. Bloemfontein. **191:** 69-71.

Butts, P. A. and K. W. Pakendorf (1984b) The utility of the embryo count method in characterzing cereal crops for resistance to *Diuraphis noxia*. Progress in Russian wheat aphid (*Diuraphis noxia* Mordw.) research in the Republic of South Africa. M. C. Walters, Technical Communication, Department of Agriculture, Republic of South Africa. **191:** 53-57.

Butts, P. A. and K. W. Pakendorf (1984c) Wheat breeding for resistance to *Diuraphis noxia*: methodology and progress. Technical Communication Department of Agriculture Republic of South Africa. M. C. Walters. Bloemfontein. **191:** 53-57.

Cai, D., M. Kleine, S. Kifle, H.-J. Harloff, N. N. Sandal, K. A. Marcker, R. M. Klein-Lankhorst, E. M. J. Salentijn, W. Lange, W. J. Stiekema, U. Wyss, F. M. W.

Grundler and C. Jung (1997) Positional cloning of a gene for Nematode resistance in sugar beet. *Science* **275**: 832-834.

Cannon, S. B. and N. D. Young (2003) The genomic architecture and evolution of (plant) NBS-LRRs. Plant-Microbe Interactions. G. Stacey and N. T. Keen. St. Paul, APS Press.

Cannon, S. B., H. Zhu, A. M. Baumgarten, R. Spangler, G. May, D. R. Cook and N. D. Young (2002) Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. *Journal of Molecular Evolution* **54**: 548-562.

Cape, J. L., M. K. Bowman and D. M. Kramer (2006) Computation of the redox and protonation properties od quinones: towards the prediction of redox cycling natural products. *Phytochemistry* **67**: 1781-1788.

Castro, A. M., S. Ramos, A. Vasicek, A. Worland, D. Giménez, A. A. Clúa and E. Suárez (2001) Identification of wheat chromosomes involved with different types of resistance against greenbug (*Schizaphis graminum*, Rond.) and the Russian wheat aphid (*Diuraphis noxia*, Mordvilko). *Euphytica* **118**: 321-330.

Castro, A. M., A. Vasicek, S. Ramos, A. Worland, E. Suarez, M. Munoz, D. Gimenez and A. A. Clua (1999) Different types of resistance against greenbug, *Schizaphis graminum* Rond, and the Russian wheat aphid, *Diuraphis noxia* Mordvilko, in wheat. *Plant Breeding* **118**: 131-137.

Chassot, C. and J.-P. Metraux (2005) The cuticle as source of signals for plant defense. *Plant Biosystems* **139**: 28-31.

Cherqui, A. and W. F. Tjallingii (2000) Salivary proteins of aphids, a pilot study on identification, separation and immunolocalisation. *Journal of Insect Physiology* **46**: 1177-1186.

Chisholm, S. T., G. Coaker, B. Day and B. J. Staskawicz (2006) Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* **124**: 803-814.

Clark, T. L. and F. J. Messina (1998) Plant architecture and the foraging success of ladybird beetles attacking the Russian wheat aphid. *Entomologia Experimentalis et Applicata* **86**: 153-161.

Dangl, J. L. and J. D. Jones (2001) Plant pathogens and integrated defence responses to infection. *Nature* **411**: 826-833.

Dangl, J. L. and J. M. McDowell (2006) Two modes of pathogen recognition by plants. *Proceedings of the National Academy of Sciences, USA* **103**: 8575-8576.

Delrot, S., R. Atanassova and L. Maurousset (2000) Regulation of sugar, amino acid and peptide plant membrane transporters. *Biochimica et Biophysica Acta* **1465**: 281-306.

Deslandes, L., J. Olivier, F. Theulières, J. Hirsch, D. X. Feng, P. Bittner-Eddy, J. Beynon and Y. Marco (2002) Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by the recessive *RRS1-R* gebe, a member of a novel family of resistance genes. *Proceedings of the National Academy of Sciences, USA* **99**: 2404-2409.

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

Dixon, A. F. G. (1987) The way of life of aphids: host specificity, speciation and distribution. Amsterdam, Elsevier.

Dixon, A. F. G. (1998) Aphid ecology: An optimization approach. London, Chapman & Hall.

Dong, H. and J. S. Quick (1995) Inheritance and allelism of resistances to the Russian wheat aphid in seven wheat lines. *Euphytica* **81**: 299-303.

Dong, H., J. S. Quick and Y. Zhang (1997) Inheritance and allelism of Russian wheat aphid resistance in several wheat lines. *Plant Breeding* **116**: 449-453.

Douglas, A. E. (1998) Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annual Review of Entomology* **43**: 17-37.

Douglas, A. E. (2006) Phloem-sap feeding by animals: problems and solutions. *Journal of Experimental Botany* **57**: 747-754.

Du Toit, F. (1984) The use of two soil systemic insecticides against *Diuraphis noxia* on winter wheat. Progress in Russian wheat aphid (Diuraphis noxia Mordw.) research in the Republic of South Africa. M. C. Walters. Bloemfontein, Technical Communication Department of Agriculture Republic of South Africa. **191:** 58-62.

Du Toit, F. (1986) Economic thresholds for *Diuraphis noxia* (Hemiptera: Aphididae) on winter wheat in the eastern Orange Free State. *Phytophylactica* **18**: 107-109.

Du Toit, F. (1987) Resistance in wheat (*Triticum aestivum*) to *Diuraphis noxia* (Hemiptera: Aphididae). *Cereal Research Communications* **15**: 175-179.

Du Toit, F. (1989) Inheritance of resistance in two *Triticum aestivum* lines to Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **82**: 1251-1253.

Du Toit, F. and M. C. Walters (1984) Damage assesment and economic threshold values for the chemical control of the Russian wheat aphid, *Diuraphis noxia* (Mordvilko) on winter wheat. Progress in Russian wheat aphid (Diuraphis noxia Mordw.) research in the Republic of South Africa. M. C. Walters. Bloemfontein, Technical Communication Department of Agriculture Republic of South Africa. **191:** 58-62.

Durrant, W. E., O. Rowland, P. Piedras, K. E. Hammond-Kosack and J. D. G. Jones (2000) cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. *The Plant Cell* **12**: 963-977.

Eleftheriou, E. P. (1990) Monocotyledons. Sieve elements: comparative structure, induction and development. H.-D. Behnke and R. D. Sjolund. Berlin, Springer-Verlag**:** 139-159.

Ellis, J. G., G. J. Lawrence, J. E. Luck and P. N. Dodds (1999) Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. *The Plant Cell* **11**: 495-506.

Elsidaig, A. A. and P. K. Zwer (1993) Genes for resistance to Russian wheat aphid in PI294994 wheat. *Crop Science* **33**: 998-1001.

Febvay, G., Y. Rahbe, M. Rynkiewicz, J. Guillaud and G. Bonnot (1999) Fate of dietary sucrose and neosythesis of amino acids in the pea aphid, *Acyrthosiphon pisum*, reared on different diets. *Journal of Experimental Biology* **202**: 2639-2652.

Feldman, M. and A. A. Levy (2005) Allopolyploidy – a shaping force in the evolution of wheat genomes. *Cytogenetic and Genome Research* **109**: 250-258.

Fidantsef, A. L., M. J. Stout, J. S. Thaler, S. S. Duffey and R. M. Bostock (1999) Signal interactions in pathogen and insect attack: expression of lipoxygenase, proteinase inhibitor II, and pathogenesis-related protein P4 in the tomato *Lycopersicon esculentum*. *Physiological and Molecular Plant Pathology* **54**: 97-114.

Fischer, W.-N., B. André, D. Rentsch, S. Krolkiewicz, M. Tegeder, K. Breitkreuz and W. B. Frommer (1998) Amino acid transport in plants. *Trends in Plant Science* **3**: 188-195.

Fluhr, R. (2001) Sentinels of disease. Plant resistance genes. *Plant Physiology* **127**: 1367-1374.

Fotopoulos, V., M. C. De Tullio, J. Barnes and A. K. Kanellis (2008) Altered stomatal dynamics in ascorbate oxidase over-expressing tobacco plants suggest a role for dehydroascorbate signalling. *Journal of Experimental Botany* **59**: 729-737.

Fotopoulos, V., M. Sanmartin and A. K. Kanellis (2006) Effect of ascorbate oxidase over-expression on ascorbate recycling gene expression in response to agents imposing oxidative stress. *Journal of Experimental Botany* **57**: 3933-3943.

Fouche, A., R. L. Verhoeven, P. H. Hewitt, M. C. Walters, C. F. Kriel and J. De Jager (1984) Russian aphid (*Diuraphis noxia*) feeding damage on wheat, related cereals and a *Bromus* grass species. Progress in Russian wheat aphid (*Diuraphis noxia* Mordw.) research in the Republic of South Africa. M. C. Walters. Bloemfontein, Technical Communication Department of Agriculture Republic of South Africa. **191:** 22-23.

Foyer, C. H. and G. Noctor (2005) Redox homeostasis and antioxidant signaling: a metabolic interface between stress perception and physiological responses. *Plant Cell* **17**: 1866-1875.

Foyer, C. H. and G. Noctor (2009) Redox regulation in photosynthetic organisms: signaling, acclimation, and practical implications. *Antioxidants & Redox Signaling* **11**: 1-46.

Fritz, A. K., S. Caldwell and W. D. Worrall (1999) Molecular mapping of Russian wheat aphid resistance from triticale accession PI 386156. *Crop Science* **39**: 1707-1710.

Gaupels, F., A. Buhtz, T. Knauer, S. Deshmukh, F. Waller, A. J. E. Van Bel, K.-H. Kogel and J. Kehr (2008) Adaptation of aphid stylectomy for analyses of proteins and mRNAs in barley phloem sap. *Journal of Experimental Botany* **59**: 3297-3306.

Genoud, T. and J.-P. Metraux (1999) Crosstalk in plant cell signaling: structure and function of the genetic network. *Trends in Plant Science* **4**.

Gil, R., A. Latorre and A. Moya (2004) Bacterial endosymbionts of insects: insights from comparative genomics. *Environmental Entomology* **6**: 1109-1122.

Gill, K. S., E. L. Lubbers, B. S. Gill, W. J. Raupp and T. S. Cox (1991) A genetic linkage map of *Triticum tauschii* (DD) and its relationship to the D genome of bread wheat (AABBDD). *Genome* **34**: 362-374.

Girma, M., G. Wilde and J. C. Reese (1990) Influence of temperature and plant growth stage on development, reproduction, life span, and intrinsic rate of increase of the Russian wheat aphid (Homoptera: Aphididae). *Environmental Entomology* **19**: 1438-1442.

Girma, M., G. E. Wilde and J. C. Reese (1992) Russian wheat aphid (Homoptera: Aphididae) feeding behavior on host and nonhost plants *Journal of Economic Entomology* **85**: 395-401.

Gonzalez, D., C. G. Summers and C. O. Qualset (1992) Russian wheat aphid: natural enemies, resistant wheat offer potential control. *California Agriculture* **46**: 32-34.

Grayer, R. J., J. B. Harbone, F. M. Kimmins, P. C. Stevenson and H. N. P. Wijayagunasekera (1994) Phenolics in rice phloem sap as sucking deterrents to the brown planthopper, *Nilaparvata lugens*. *Acta Horticulturae* **381**: 691-694.

Haile, F. J., L. G. Higley, X. Ni and S. S. Quisenberry (1999) Physiological and growth tolerance in wheat to Russian wheat aphid (Homoptera: Aphididae) injury. *Environmental Entomology* **28**: 787-794.

Haley, S. D., F. B. Peairs, C. B. Walker, J. B. Rudolph and T. L. Randolph (2004) Occurrence of a new Russian wheat aphid biotype in Colorado. *Crop Science* **44**: 1589-1592.

Halitschke, R. and I. T. Baldwin (2005) Jasmonates and related compounds in plant-insect interactions. *Journal of Plant Growth Regulation* **23**: 238-245.

Halterman, D., F. Zhou, F. Wei, R. P. Wise and P. Schulze-Lefert (2001) The MLA6 coiled-coil, NBS-LRR protein confers *AvrMla6*-dependent resistance specificity to *Blumeria graminis* f. sp. *hordei* in barley and wheat. *The Plant Journal* **25**: 335-348.

Hammond-Kosack, K. E. and D. A. Jones (1997) Plant disease resistance genes. *Annual Review of Plant Physiology and Plant Molecular Biology* **48**: 575-607.

Hammond-Kosack, K. E. and J. D. G. Jones (1996) Resistance gene-dependent plant defense responses. *The Plant Cell* **8**: 1773-1791.

He, X. Y., Z. H. He, C. F. Morris and X. C. Xia (2009) Cloning and phylogenetic analysis of polyphenol oxidase genes in common wheat and related species. *Genetic Resources & Crop Evolution* **56**: 311-321.

Heath, M. C. (2000a) Hypersensitive response-related death. *Plant Molecular Biology* **44**: 321-334.

Heath, M. C. (2000b) Nonhost resistance and nonspecific plant defenses. *Current Opinion in Plant Biology* **3**: 315-319.

Heie, O. E. (1992) The Aphidoidea (Hemiptera) of Fennoscandia and Denmark. IV. Family Aphididae: Part 1 of tribe Macrosiphini of subfamily Aphidinae. Leiden, The Netherlands, E.J. Brill / Scandinavian Science Press Ltd.

Hewitt, P. H., G. J. J. Van Niekerk, M. C. Walters, C. F. Kriel and A. Fouche (1984) Aspects of the ecology of the Russian wheat aphid, Duiraphis noxia, in the Bloemfontein district. I. The colonization and infestation of sown wheat, identification of summer hosts and cause of infestation symptoms. Progress in Russian wheat aphid (Diuraphis noxia Mordw.) research in the Republic of South Africa. M. C. Walters. Bloemfontein, Technical Communication Department of Agriculture Republic of South Africa. **191:** 3-13.

Imlau, A., E. Truernit and N. Sauer (1999) Cell-to-cell and long-distance trafficking of the green fluorescent protein in the phloem and symplastic unloading of the protein into sink tissue. *The Plant Cell* **11**: 309-322.

Jarabak, R., R. G. Harvey and J. Jarabak (1997) Redox cycling of polycyclic aromatic hydrocarbon *o*-quinones: reversal of superoxide dismutase inhibition by ascorbate. *Archives of Biochemistry and Biophysics* **339**: 92-98.

Jones, D. A., C. M. Thomas, K. E. Hammond-Kosack, P. J. Balint-Kurti and J. D. G. Jones (1994) Isolation of the tomato *Cf-9*gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* **266**: 789-793.

Joshi, B. N., M. N. Sainani, B. Bastawade, V. V. Deshpande, V. S. Gupta and P. K. Ranjekar (1999) Pearl millet cysteine protease inhibitor. Evidence for the presence of two distinct sites responsible for anti-fungal and anti-feedent activities. *European Journal of Biochemistry* **265**: 556-563.

Jyoti, J. L. and J. P. Michaud (2005) Comparative biology of a novel strain of Russian wheat aphid (Homoptera: Aphididae) on three wheat cultivars. *Journal of Economic Entomology* **98**: 1032-1039.

Jyoti, J. L., J. A. Qureshi, J. P. Michaud and T. J. Martin (2006) Virulence of two Russian wheat aphid biotypes to eight wheat cultivars at two temperatures. *Crop Science* **46**: 774-780.

Kaloshian, I. (2004) Gene-for-gene disease resistance: bridging insect pest and pathogen defense. *Journal of Chemical Ecology* **30**: 2419-2438.

Kaloshian, I., W. H. Lange and V. M. Williamson (1995) An aphid-resistance locus is tightly linked to the nematode-resistance gene, *Mi*, in tomato. *Proceedings of the National Academy of Sciences, USA* **92**: 622-625.

Kanazin, V., L. F. Marek and R. C. Shoemaker (1996) Resistance gene analogs are conserved and clustered in soybean. *Proc. Natl. Acad. Sci. USA* **93**: 11746-11750.

Karley, A. J., A. E. Douglas and W. E. Parker (2002) Amino acid composition and nutritional quality of potato leaf phloem sap for aphids. *Journal of Experimental Biology* **205**: 3009-3018.

Kazemi, M. H., P. Talebi-Chaichi, M. R. Shakiba and M. M. Jafarloo (2001) Biological responses of Russian wheat aphid, *Diuraphis noxia* (Mordvilko) (Homoptera: Aphididae) to different wheat varieties. *Journal of Agricultural Science and Technology* **3**: 249-255.

Kehr, J. and A. Buhtz (2008) Long distance transport and movement of RNA through the phloem. *Journal of Experimental Botany* **59**: 85-92.

Kindler, S. D., L. G. Greer and T. L. Springer (1992) Feeding behavior of the Russian wheat aphid (Homoptera: Aphididae) on wheat and resistant and susceptible slender wheatgrass. *Journal of Economic Entomology* **85**: 2012-2016.

Kindler, S. D. and T. L. Springer (1989) Alternate hosts of Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **82**: 1358-1362.

Kobe, B. and J. Deisenhofer (1994) The leucine-rich repeat: a versatile binding motif. *Trends in biochem Sci* **19**: 415-421.

Kobe, B. and J. Deisenhofer (1995) A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* **374**: 183-186.

Kobe, B. and A. V. Kajava (2001) The leucine-rich repeat as a protein recognition motif. *Current Opinion in Structural Biology* **11**: 725-732.

Kogan, M. and E. E. Ortman (1978) Antixenosis - a new term proposed to replace Painter's "non-preference" modality of resistance. *Bulletin of Entomological Society of America* **24**: 175-176.

Kriel, C. F., P. H. Hewitt, J. De Jager, M. C. Walters, A. Fouche and A. J. Van der Westhuizen (1984) Aspects of the ecology of the Russian wheat aphid, *Diuraphis noxia*, in the Bloemfontein district. II. Population dynamics. Progress in Russian wheat aphid (*Diuraphis noxia* Mordw.) research in the Republic of South Africa. M. C. Walters. Bloemfontein, Technical Communication Department of Agriculture Republic of South Africa. **191:** 14-21.

Kriel, C. F., P. H. Hewitt, M. C. van der Westhuizen and M. C. Walters (1986) The Russian wheat aphid *Diuraphis noxia* (Mordvilko): population dynamics and effect on grain yield in the western Orange Free State. *Journal of the Entomological Society of Southern Africa* **49**: 317-335.

Kruger, G. H. J. and P. H. Hewitt (1984) The effect of Russian Wheat Aphid (*Diuraphis noxia*) extract on photosynthesis of isolated chloroplasts: preliminary studies. Progress in Russian wheat aphid (*Diuraphis noxia* Mordw.) research in the Republic of South Africa. M. C. Walters. Bloemfontein, Technical communication Department of Agriculture Republic of South Africa. **191:** 34-37.

Kubalakova, M., J. Vrana, J. Cihalikova, H. Simkova and J. Dolezel (2002) Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* **104**: 1362-1372.

Kunkel, B. N. and D. M. Brooks (2002) Cross talk between signaling pathways in pathogen defense. *Current Opinion in Plant Biology* **5**: 325-331.

La Camera, S., G. Gouzerh, S. Dhondt, L. Hoffmann, B. Fritig, M. Legrand and T. Heitz (2004) Metabolic reprogramming in plant innate immunity: the contributions of pheynylpropanoid and oxylipin pathways. *Immunological Reviews* **198**: 267-284.

Lacock, L., C. van Niekerk, S. Loots, F. du Preez and A.-M. Botha (2003) Functional and comparative analysis of expressed sequences from *Diuraphis noxia* infested wheat obtained utilizing the conserved nucleotide binding site. *African Journal of Biotechnology* **2**: 75-81.

Lagudah, E. S., O. Moullet and R. Appels (1997) Map-based cloning of a gene sequence encoding a nucleotide-binding domain and a leucine-rich region at the *Cre3* nematode resistance locus of wheat. *Genome* **40**: 659-665.

Lai, C.-Y., L. Baumann and P. Baumann (1994) Amplification of *trpEG*: adaptation of *Buchnera aphidicola* to an endosymbiotic association with aphids. *Proceedings of the National Academy of Sciences, USA* **91**: 3819-3823.

Lai, C.-Y., P. Baumann and N. A. Moran (1996) The endosymbiont (*Buchnera* sp.) of the aphid *Diuraphis noxia* contains plasmids consisting of *trpEG* and tandem repeats of *trpEG* pseudogenes. *Applied and Environmental Microbiology* **62**: 332-339.

Lam, E., N. Kato and M. Lawton (2001) Programmed cell death, mitochondria and the plant hypersensitive response. *Nature* **411**: 848-853.

Lamikanra, O. (2002) Fresh-cut fruits and vegetables: science, technology, and market, CRC Press.

Lapitan, N. L. V., Y.-C. Li, J. Peng and A.-M. Botha (2007a) Fractionated extracts of Russian wheat aphid eliciting defense responses in wheat. *Journal of Economic Entomology* **100**: 990-999.

Lapitan, N. L. V., Y.-C. Li, R. S. G. Walters, Y. Peng, F. B. Peairs and A.-M. Botha (2007b) Limited nuclear and mitochondrial DNA variation among Russian wheat aphid (*Diuraphis noxia*) biotypes from the United States and Africa, *American Entomological Society*, San Diego, December 9-12. American Entomological Society, San Diego.

Le Hir, R., J. Beneteau, C. Bellini, F. Vilaine and S. Dinant (2008) Gene expression profiling: keys for investigating phloem functions. *Trends in Plant Science* **13**: 273-280.

Lequeu, J., M.-L. Fauconnier, A. Chammai, R. Bronner and E. Blee (2003) Formation of plant cuticle: evidence for the occurence of the peroxygenase pathway. *The Plant Journal* **36**: 155-164.

Lin, M.-K., Y.-J. Lee, T. J. Lough, B. S. Phinney and W. J. Lucas (2008) Analysis of the pumpkin phloem proteome provides functional insights into angiosperm sieve tube function. *Molecular and Cellular Proteomics*: M800420-MCP200.

Liu, X. M., C. M. Smith and B. S. Gill (2002) Identification of microsatellite markers linked to Russian wheat aphid resistance genes *Dn4* and *Dn6*. *Theoretical and Applied Genetics* **104**: 1042-1048.

Liu, X. M., C. M. Smith, B. S. Gill and V. Tolmay (2001) Microsatellite markers linked to six Russian wheat aphid resistance genes in wheat. *Theoretical and Applied Genetics* **102**: 504-510.

Lough, T. J. and W. J. Lucas (2006) Integrative plant biology: Role of phloem long-distance macromolecular trafficking. *Annual Review in Plant Biology* **57**: 203-232.

Ma, Z.-Q., A. Saidi, J. S. Quick and N. L. V. Lapitan (1998) Genetic mapping of Russian wheat aphid resistance genes *Dn2* and *Dn4* in wheat. *Genome* **41**: 303-306.

Manicardi, G. C., M. Mandrioli, D. Bizzaro and U. Bianchi (2002) Cytogenetic and molecular analysis of heterochromatic areas in the holocentric chromosomes of different aphid spiecies. Some aspects of chromosome structure and functions. R. C. Sobti, G. Obe and R. S. Athwal. New Delhi, India, Narosa publishing house**:** 47-56.

Marais, G. F. and F. Du Toit (1993) A monosomic analysis of Russian wheat aphid resistance in the common wheat PI 294994. *Plant Breeding* **111**: 246-248.

Marais, G. F., M. Horn and F. Du Toit (1994) Intergeneric transfer (rye to wheat) of a gene(s) for Russian wheat aphid resistance. *Plant Breeding* **113**: 265-271.

Martin, B., J. L. Collar, W. F. Tjallingii and A. Fereres (1997) Intracellulat ingestion and salivation by aphids may cause the acquisition and inoculation of non-persistently transmitted plant viruses. *Journal of General Virology* **78**: 2701-2705.

Martin, G. B., S. H. Brommonschenkel, J. Chunwongse, A. Frary, M. W. Ganal, R. Spivey, T. Wu, E. D. Earle and S. D. Tanksley (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262**: 1432-1436.

McConn, M., R. A. Creelman, E. Bell, J. E. Mullet and J. Browse (1997) Jasmonate is essential for insect defense in *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA* **94**: 5473-5477.

Michelmore, R. W. (2000) Genomic approaches to plant disease resistance. *Current Opinion in Plant Biology* **3**: 125-131.

Miles, P. W. (1965) Studies on the salivary physiology of plant-bugs: the salivary secretions of aphids. *Journal of Insect Physiology* **11**: 1261-1268.

Miles, P. W. (1967) The physiological division of labour in the salivary glands of *Oncopeltus fasciatus* (Dall.) (Heteroptera:Lygaeidae). *Australian Journal of Biological Science* **20**: 785-797.

Miles, P. W. (1968) Insect secretions in plants. *Annual Review of Phytopathology* **6**: 137-164.

Miles, P. W. (1999) Aphid saliva. *Biological Review* **74**: 41-85.

Miles, P. W. and J. J. Oertli (1993) The significance of antioxidants in the aphid-plant interaction: the redox hypothesis. *Entomologia Experimentalis et Applicata* **67**: 275-283.

Miles, P. W. and Z. Peng (1989) Studies on the salivary physiology of plant bugs: detoxification of phytochemicals by the salivary peroxidase of aphids. *Journal of Insect Physiology* **35**: 865-872.

Miller, C. A., A. Altinkut and N. L. V. Lapitan (2001) A microsatellite marker for tagging *Dn2*, a wheat gene conferring resistance to the Russian wheat aphid. *Crop Science* **41**: 1584-1589.

Miller, H., D. R. Porter, J. D. Burd, D. W. Mornhinweg and R. L. Burton (1994) Physiological effects of Russian wheat aphid (Homoptera: Aphididae) on resistant and susceptible barley. *Journal of Economic Entomology* **86**: 493-499.

Mitchell, H. J., J. L. Hall and M. S. Barber (1994) Elicitor-induced cinnamyl alcohol dehydrogenase activity in lignifying wheat (*Triticum aestivum* L.) leaves. *Plant Physiology* **104**: 551-556.

Mitchell, H. J., S. A. Hall, R. Stratford, J. L. Hall and M. S. Barber (1999) Differential induction of cinnamyl alcohol dehydrogenase during defensive lignification in wheat (*Triticum aestivum* L.): characterisation of the major inducible form. *Planta* **208**: 31-37.

Mittler, R., E. H. Herr, B. L. Orvar, W. van Camp, H. Willekens, D. Inze and B. E. Ellis (1999) Transgenic tobacco plants with reduced capability to detoxify reactive oxygen intermediates are hyperresponsive to pathogen infection. *Proceedings of the National Academy of Sciences, USA* **96**: 14165-14170.

Miyake, N. and T. Kurata (1998) Possible formation of dehydro-L-ascorbic acid from 2,3-Diketo-L-gulonic acid in an aqueous solution. *Biosciences, Biotechnology and Biochemistry* **62**: 1419-1421.

Mochida, K., Y. Yamazaki and Y. Ogihara (2003) Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Molecular and General Genetics* **270**: 371 - 377.

Moffett, P., G. Farnham, J. Peart and D. C. Baulcombe (2002) Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *The EMBO Journal* **21**: 4511-4519.

Mohase, L. and A. J. van der Westhuizen (2002) Glycoproteins from Russian wheat aphid infested wheat induce defense responses. *Z. Naturforsch* **57c**: 867-873.

Montesano, M., G. Brader and E. T. Palva (2003) Pathogen derived elicitors: searching for receptors in plants. *Molecular Plant Pathology* **4**: 73-79.

Moran, N. A., H. E. Dunbar and J. L. Wilcox (2005) Regulation of transcription in a reduced bacterial genome: nutrient-provisioning genes of the obligate symbiont *Buchnera aphidicola*. *Journal of Bacteriology* **187**: 4229-4237.

Moran, P. J., Y. Cheng, J. L. Cassell and G. A. Thompson (2002) Gene expression profiling of *Arabidopsis thaliana* in compatible plant-aphid interactions. *Archives of Insect Biochemistry and Physiology* **51**: 182-203.

Moran, P. J. and G. A. Thompson (2001) Molecular responses to aphid feeding in *Arabidopsis* in relation to plant defense pathways. *Plant Physiology* **125**: 1074-1085.

Munson, M. A. and P. Baumann (1993) Molecular cloning and nucleotide sequence of a putative *trpDC(F)BA* operon in *Buchnera aphidicola* (endosymbiont of the aphid *Schizaphis graminum*). *Journal of Bacteriology* **175**: 6426-6432.

Munson, M. A., P. Baumann, A. M. Clark, L. Baumann, N. A. Moran, D. J. Voegtlin and B. C. Campbell (1991) Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. *Journal of Bacteriology* **173**: 6321-6324.

Myburg, A. A., M. Cawood, B. D. Wingfield and A.-M. Botha (1998) Development of RAPD and SCAR markers linked to the Russian wheat aphid resistance gene *Dn2* in wheat. *Theoretical and Applied Genetics* **96**: 1162-1169.

Mysore, K. S., M. D. D'Ascenzo, X. He and G. B. Martin (2003) Overexpression of the disease resistance gene *Pto* in tomato induces gene expression changes similar to immune responses in human and fruitfly. *Plant Physiology* **132**: 1901-1912.

Nagel, M. A. C. (1995) Effect of exposure to the Russian wheat aphid on the expression of chitinase. Department of Botany and Genetics. Bloemfontein, University of the Orange Free State**:** 161.

Nakabachi, A., S. Shigenobu, N. Sakazume, T. Shiraki, Y. Hayashizaki, P. Carninci, H. Ishikawa, T. Kudo and T. Fukatsu (2005) Transcriptome analysis of the aphid bacteriocyte, the symbiotic host cell that harbors an endocellular mutualistic bacterium, *Buchnera*. *Proceedings of the National Academy of Sciences, USA* **102**: 5477-5482.

Ni, X. and S. S. Quisenberry (1997) Effect of wheat leaf epicuticular structure on host selection and probing rhythm of Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **90**: 1400-1407.

Ni, X., S. S. Quisenberry, J. Markwell, T. Heng-Moss, L. Higley, F. Baxendale, G. Sarath and R. Klucas (2001) *In vitro* enzymatic chlorophyll catabolism in wheat elicited by cereal aphid feeding. *Entomologia Experimentalis et Applicata* **101**: 159-166.

Ni, X., S. S. Quisenberry, S. Pornkulwat, J. L. Figarola, S. R. Skoda and J. E. Foster (2000) Hydrolase and oxido-reductase activities in *Diuraphis noxia* and *Rhopalosiphum padi* (Hemiptera: Aphididae). *Annals of the Entomological Society of America* **93**: 595-601.

Ni, X., S. S. Quisenberry, B. D. Siegfried and K. W. Lee (1998) Influence of cereal leaf epicuticular wax on *Diuraphis noxia* probing behavior and nymphoposition. *Entomologia Experimentalis et Applicata* **89**: 111-118.

Nkongolo, K. K., J. S. Quick, A. E. Limin and D. B. Fowler (1991a) Sources and inheritance of resistance to Russian wheat aphid in *Triticum* species amphiploids and *Triticum tauschii*. *Canadian Journal of Plant Science* **71**: 703-308.

Nkongolo, K. K., J. S. Quick, F. B. Peairs and W. L. Meyer (1991b) Inheritance of resistance of PI 372129 wheat to the Russian wheat aphid. *Crop Science* **31**: 905-907.

Noctor, G. (2006) Metabolic signalling in defence and stress: the central roles of soluble redox couples. *Plant, Cell & Environment* **29**: 409-425.

Nowierski, R. M. and B. C. Fitzgerald (2002) Supercooling capacity of Eurasian and North American populations of parasitoids of the Russian wheat aphid, *Diuraphis noxia*. *BioControl* **47**: 279-292.

Nürnberger, T. (1999) Signal perception in plant pathogen defense. *Cellular and Molecular Life Sciences* **55**: 167-182.

Odjakova, M. and C. Hadjiivanova (2001) The complexity of pathogen defense in plants. *Bulgarian Journal of Plant Physiology* **27**: 101-109.

Ogihara, Y., K. Mochida, Y. Nemoto, K. Murai, Y. Yamazaki, T. Shin-I and Y. Kohara (2003) Correlated clustering and viral display of gene expression patterns in the wheat life cycle by large-scale statistical analysis of expressed sequence tags. *The Plant Journal* **33**: 1001-1011.

Ortiz-Lopez, A., H.-C. Chang and D. R. Bush (2000) Amino acid transporters in plants. *Biochimica et Biophysica Acta* **1465**: 275-280.

Painter, R. H. (1936) The food of insects and its relation to resistance of plants to insect attack. *The American Naturalist* **70**: 547-566.

Painter, R. H. (1958) Resistance of plants to insects. *Annual Review of Entomology* **3**: 267-290.

Pan, Q., Y.-S. Liu, O. Budai-Hanrian, M. Sela, L. Carmel-Goren, D. Zamir and R. Fluhr (2000a) Comparative genetics of nucleotide binding site-leucine rich repeat resistance gene homologues in the genomes of two dicotyledons: tomato and *Arabidopsis*. *Genetics* **155**: 309-322.

Pan, Q., J. Wendel and R. Fluhr (2000b) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *Journal of Molecular Evolution* **50**: 203-213.

Peeters, P. J. (2002) Correlations between leaf structural traits and the densities of herbivorous insect guilds. *Biological Journal of the Linnean Society* **77**: 43-65.

Pieterse, C. M. J., J. Ton and L. C. Van Loon (2001) Cross-talk between plant defence signalling pathways: boost or burden? *AgBiotechNet* **3**: 1-8.

Pieterse, C. M. J. and L. C. Van Loon (1999) Salicylic acid-independent plant defence pathways. *Trends in Plant Science* **4**: 52-58.

Pignocchi, C., J. M. Fletcher, J. E. Wilkinson, J. D. Barnes and C. H. Foyer (2003) The function of ascorbate oxidase in tobacco. *Plant Physiology* **132**: 1-11.

Pinedo, M. L., C. Segarra and R. D. Conde (1993) Occurrence of two endoproteinases in wheat leaf intercellular washing fluid. *Physiologia Plantarum* **88**: 287 - 293.

Porter, D. R., C. A. Baker and J. A. Webster (1998) Inheritance of Russian wheat aphid resistance in PI 140207 spring wheat. *Plant Breeding* **117**.

Porter, D. R. and J. A. Webster (2000) Russian wheat aphid-induced protein alterations in spring wheat. *Euphytica* **111**: 199-203.

Potgieter, G. F., G. F. Marais and F. Du Toit (1991) The transfer of resistance to the Russian wheat aphid from *Triticum monococcum* L. to common wheat. *Plant Breeding* **106**: 284-292.

Powell, G., S. P. Maniar, J. A. Pickett and J. Hardie (1999) Aphid responses to non-host epicuticular lipids. *Entomologia Experimentalis et Applicata* **91**: 115-123.

Prinsloo, G., Y. Chen, K. L. Giles and M. H. Grenstone (2002) Release and recovery in South Africa of the exotic aphid parasitoid *Aphelinus hordei* verified by the polymerase chain reaction. *BioControl* **47**: 127-136.

Puterka, G. J., J. D. Burd and R. L. Burton (1992) Biotypic variation in a worldwide collection of Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **85**: 1497-1506.

Quick, J. S., K. K. Nkongolo, W. Meyer, F. B. Peairs and B. Weaver (1991) Russian wheat aphid reaction and agronomic and quality traits of a resistant wheat. *Crop Science* **31**: 50-53.

Qureshi, J. A. and J. P. Michaud (2005) Comparative biology of three cereal aphids on TAM 107 wheat. *Environmental Entomology* **34**: 27-36.

Radman, R., T. Saez, C. Bucke and T. Keshavarz (2003) Elicitation of plants and microbial cell systems. *Biotechnology and Applied Biochemistry* **37**: 91-102.

Rafi, M. M., R. S. Zemetra and S. S. Quisenberry (1996) Interaction between Russian wheat aphid (Homoptera: Aphididae) and resistant and susceptible genotypes of wheat. *Journal of Economic Entomology* **89**: 239-246.

Reed, D. K., S. D. Kindler and T. L. Springer (1992) Interactions of Russian wheat aphid, a hymenopterous parasitoid and resistant and susceptible slender wheatgrasses. *Entomologia Experimentalis et Applicata* **64**: 239-246.

Roginsky, V. A., T. K. Barsukova and H. B. Stegmann (1999) Kinetics of redox interaction between substituted quinones and ascorbate under aerobic conditions. *Chemico-Biological Interactions* **121**: 177-197.

Rossi, M., F. L. Goggin, S. B. Milligan, I. Kaloshian, D. E. Ullman and V. M. Williamson (1998) The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *Proceedings of the National Academy of Sciences, USA* **95**: 9750-9754.

Rouhbakhsh, D., C.-Y. Lai, C. D. von Dohlen, M. A. Clark, L. Baumann, P. Baumann, N. A. Moran and D. J. Voegtlin (1996) The tryptophan biosynthetic pathway of aphid endosymbionts (*Buchnera*): genetics and evolution of plasmid-associated anthranilate synthase (*trpEG*) within the Aphididae. *Journal of Molecular Evolution* **42**.

Ruiz-Medrano, R., B. Xoconostle-Cázares and W. J. Lucas (1999) Phloem long-distance transport of *CmNACP* mRNA: implications for supracellular regulation in plants. *Development* **126**: 4405-4419.

Ryals, J. A., U. H. Neuenschwander, M. G. Willits, A. Molina, H.-Y. Steiner and M. D. Hunt (1996) Systemic acquired resistance. *The Plant Cell* **6**: 1809-1819.

Saheed, S. A., C. E. J. Botha, L. Liu and L. Jonsson (2007a) Comparison of structural damage caused by Russian wheat aphid (*Diuraphis noxia*) and Bird cherry-oat aphid (*Rhopalosiphum padi*) in a susceptible barley cultivar, *Hordeum vulare* cv. Clipper. *Physiologia Planatarum* **129**: 429-435.

Saheed, S. A., L. Liu, L. Jonsson and C. E. J. Botha (2007b) Xylem – as well as phloem – sustains severe damage due to feeding by the Russian wheat aphid. *South African Journal of Botany*.

Saidi, A. and J. S. Quick (1996) Inheritance and allelic relationships among Russian wheat aphid resistance genes in winter wheat. *Crop Science* **36**: 256-258.

Sandstrom, J. and N. A. Moran (1999) How nutritionally imbalanced is phloem sap for aphids? *Entomologia Experimentalis et Applicata* **91**: 203-210.

Sandstrom, J., A. Telang and N. A. Moran (2000) Nutritional enhancement of host plants by aphids-a comparison of three aphid species on grasses. *Journal of Insect Physiology* **46**: 33-40.

Schotzko, D. J. and C. M. Smith (1991) Effects of preconditioning host plants on population development of Russian wheat aphids (Homoptera: Aphididae). *Journal of Economic Entomology* **84**: 1083-1087.

Schroeder-Teeter, S., R. S. Zemetra, D. J. Schotzko, C. M. Smith and M. Rafi (1993) Monosomic analysis of Russian wheat aphid (*Diuraphis noxia*) resistance in *Triticum aestivum* line PI137739. *Euphytica (Historical Archive)* **74**: 117-120.

Seo, H. S., J. T. Song, J.-J. Cheong, Y.-H. Lee, Y.-W. Lee, I. Hwang, J. S. Lee and Y. D. Choi (2001) Jasmonic acid carboxyl methyltransferase: a key enzyme for jasmonate-regulated plant responses. *PNAS* **98**: 4788-4793.

Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki and H. Ishikawa (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* **407**: 81-86.

Shirasu, K. and P. Schulze-Lefert (2000) Regulators of cell death in desease resistance. *Plant Molecular Biology* **44**: 371-385.

Shufran, K. A., L. R. Kirkman and G. J. Puterka (2007) Absence of mitochondrial DNA sequence variation in Russian wheat aphid (Hemiptera: Aphididae) populations consistent with a single introduction into United States. *Journal of the Kansas Entomological Society* **80**: 319-326.

Silva, F. J., R. C. H. J. van Ham, B. Sabater and A. Latorre (1998) Structure and evolution of the leucine plasmids carried by the endosymbiont (*Buchnera aphidicola*) from aphids of the family Aphididae. *FEMS Microbiology Letters* **168**: 43-49.

Smith, C. M., T. Belay, C. Stauffer, P. Stary, I. Kubeckova and S. Starkey (2004) Identification of Russian wheat aphid (Homoptera: Aphididae) populations virulent to the *Dn4* resistance gene. *Journal of Economic Entomology* **97**: 1112-1117.

Smith, C. M. and E. V. Boyko (2007) The molecular bases of plant resistance and defense responses to aphid feeding: current status. *Entomologia Experimetalis et Applicata* **122**: 1-16.

Smith, C. M., D. Schotzko, R. S. Zemetra, E. J. Souza and S. Schroeder-Teeter (1991) Identification of Russian wheat aphid (Homoptera: Aphidiae) resistance in wheat. *Journal of Economic Entomology* **84**: 328-332.

Smith, C. M., D. J. Schotzko, R. S. Zemetra and E. J. Souza (1992) Categories of resistance in plant introductions of wheat resistant to the Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **85**: 1480-1484.

Smith, M., K. Shufran, P. Sloderbeck, D. Mornhinweg, M. Mirik, Y. Weng, F. Peairs, L. Hesler, C. Baker, N. Elliot, J. Blodgett, J. Brown, G. Hein, T. Royer, J. P. Michaud and G. Cuperus (2005) Guidelines for identifying biotypic variation and designation of *Diuraphis noxia* biotypes. Fort Collins, Colorado, Aphid ecology and plant-insect interactions subcommittee report, Annual meeting WERA066. **2009:** 4.

Soler, T., A. Latorre, B. Sabater and F. J. Silva (2000) Molecular characterization of the leucine plasmid from *Buchnera aphidicola*, primary endosymbiont of the aphid *Acyrthosiphon pisum*. *Current Microbiology* **40**: 264-268.

Song, W.-Y., G.-L. Wang, L.-L. Chen, H.-S. Kim, L.-Y. Pi, T. Holsten, J. Gardner, B. Wang, W.-X. Zhai, L.-H. Zhu, C. Fauquet and P. Ronald (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* **270**: 1804-1806.

Souza, E., C. M. Smith, D. J. Schotzko and R. S. Zemetra (1991) Greenhouse evaluation of red winter wheat for resistance to the Russian wheat aphid (*Duiraphis noxia*, Mordvilko). *Euphytica* **57**: 221-225.

Srivastava, P. N. (1987) Nutritional physiology. Aphids. Their biology, natural enemies and control. P. Harrewijn. Amsterdam, Elsevier. **2A:** 99-121.

Stary, P. and H. Lukasova (2002) Increase of Russian wheat aphid, *Diuraphis noxia* (Kurdj.) in hot and dry weather (Hom., Aphididae). *Journal of Pest Science* **75**: 6-10.

Stray, P. (2000) On-going expansion of Russian wheat aphid, *Diuraphis noxia* (Kurdj.) in central Europe (Hom.:Aphididae). *Journal of Pest Science* **73**: 75-78.

Stray, P. (2001) Within-field refugiums of Russian wheat aphid, *Diuraphis noxia* (Kurdj.) in cereals (Hom., Aphididae). *Journal of Pest Science* **74**: 124-125.

Tamas, I., L. Klasson, B. Canback, A. K. Naslund, A.-S. Eriksson, J. J. Wernegreen, J. P. Sandstrom, N. A. Moran and S. G. E. Andersson (2002) 50 Million years of genomic stasis in endosymbiotic bacteria. *Science* **296**: 2376-2379.

Telang, A., J. Sandstrom, E. Dyreson and N. A. Moran (1999) Feeding damage by *Diuraphis noxia* results in a nutritionally enhanced phloem diet. *Entomologia Experimentalis et Applicata* **91**: 493-412.

Thao, M. L., L. Baumann, P. Baumann and N. A. Moran (1998) Endosymbionts (*Buchnera*) from the aphids *Schizaphis graminum* and *Diuraphis noxia* have different copy mumbers of the plasmid containing the leucine biosynthetic genes. *Current Microbiology* **36**: 238-240.

Thomas, C. M., D. A. Jones, M. Parniske, K. Harrison, P. J. Balint-Kurti, K. Hatzixanthis and J. D. G. Jones (1997) Characterization of the Tomato *Cf-4* gene for Resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in *Cf-4* and *Cf-9*. *The Plant Cell* **9**: 2209-2224.

Titarenko, E., E. Rojo, J. Leon and J. J. Sanchez-Serrano (1997) Jasmonic acid-dependent and -independent signaling pathways control wound-induced gene activation in *Arabidopsis thaliana*. *Plant Physiology* **115**: 817-826.

Tolmay, V. L., R. C. Lindeque and G. J. Prinsloo (2007) Preliminary evidence of a resistance-breaking biotype of the Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Homoptera: Aphididae), in South Africa. *African Entomology* **15**: 228-230.

Tottman, D. R., R. J. Makepeace and H. Broad (1979) An explanation of the decimal code for the growth stages of cereals, with illustrations. *Annual Applied Biology* **93**: 221-234.

Traut, T. W. (1994) The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. *Eur. J. Biochem.* **222**: 9-19.

Urbanska, A., W. F. Tjallingii, A. F. G. Dixon and B. Leszczynski (1998) Phenol oxidising enzymes in the grain aphid's saliva. *Entomologia Experimentalis et Applicata* **86**: 197-203.

van der Biezen, E. A. and J. D. G. Jones (1998) Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Science* **23**: 454-456.

Van der Westhuizen, A. J. and A.-M. Oberholster (1996) Molekulêre biologie van bestandheid teen die Russiese koringluis, Departement Plantkunde en Genetika.

Van der Westhuizen, A. J. and Z. Pretorius (1996) Protein composition of wheat apoplastic fluid and resistance to the Russian wheat aphid. *Australian Journal of Plant Physiology* **23**: 645-648.

Van der Westhuizen, A. J., X.-M. Qian and A.-M. Botha (1998a) b-1,3-Glucanases in wheat and resistance to the Russian wheat aphid. *Physiologia Planatarum* **103**: 125-131.

Van der Westhuizen, A. J., X.-M. Qian and A.-M. Botha (1998b) Differential induction of apoplastic peroxidase and chitinase activities in susceptible and resistant wheat cultivars by Russian wheat aphid infestation. *Plant Cell Reports* **18**: 132-137.

Van Ham, R. C. H. J., J. Kamerbeek, C. Palacios, C. Rausell, F. Abascal, U. Bastolla, J. M. Fernandez, L. Jimenez, M. Postigo, F. J. Silva, J. Tamames, E. Viguera, A. Latorre, A. Valencia, F. Moran and A. Moya (2003) Reductive genome evolution in *Buchnera aphidicola*. *Proceedings of the National Academy of Sciences, USA* **100**: 581-586.

Van Ham, R. C. H. J., D. Martinez-Torres, A. Moya and A. Latorre (1999) Plasmid-encoded anthranilate synthase (*TrpEG*) in *Buchnera aphidicola* from aphids of the family Pemphigidae. *Applied and Environmental Microbiology* **65**: 117-125.

Van Ham, R. C. H. J., A. Moya and A. Latorre (1997) Putative evolutionary origin of plasmids carrying the genes involved in leucine biosynthesis in *Buchnera aphidicola* (endosymbiont of aphids). *Journal of Bacteriology* **179**: 4768-4777.

Van Niekerk, C. (2003) Analysis of gene expression in *Triticum aestivum* L. cv. 'Tugela DN' after Russian wheat aphid (*Diuraphis noxia* Mordvilko) infestation. Department of Genetics. Pretoria, University of Pretoria**:** 200.

Van Zyl, R. A. (2007) Elucidation of possible virulence factors present in Russian wheat aphid (*Diuraphis noxia*) biotype saliva. Department of Genetics. Pretoria, University of Pretoria**:** 118.

Venter, E. and A.-M. Botha (2000) Development of markers linked to *Diuraphis noxia* resistance in wheat using a novel PCR-RFLP approach. *Theoretical and Applied Genetics* **100**: 965-970.

Voelckel, C., W. W. Weisser and I. T. Baldwin (2004) An analysis of plant–aphid interactions by different microarray hybridization strategies. *Molecular Ecology* **13**: 3187-3195.

Walling, L. L. (2000) The myriad plant responses to herbivores. *Journal of Plant Growth Regulation* **19**: 195-216.

Walling, L. L. (2008) Avoiding effective defenses: strategies employed by phloem-feeding insects. *Plant Physiology* **146**: 859-866.

Walters, M. C., F. Penn, F. Du Toit, T. C. Botha, K. Aalbersberg, P. H. Hewitt and S. W. Broodryk (1984) The Russian wheat aphid Techincal Communication Department of Agriculture Republic of South Africa. M. C. Walters. Bloemfontein. **191:** 72-77.

Webster, J. A., D. R. Porter, C. A. Baker and D. W. Mornhinweg (1993) Resistance to Russian wheat aphid (Homoptera: Aphididae) in barley: effects of aphid feeding. *Journal of Economic Entomology* **86**: 1603-1608.

Weiland, A. A., F. B. Peairs, T. L. Randolph, J. B. Rudolph, S. D. Haley and G. J. Puterka (2008) Biotypic diversity in Colorado Russian wheat aphid (Hemiptera: Aphididae) populations. *Journal of Economic Entomology* **101**: 569-574.

Wernegreen, J. J. and N. A. Moran (2000) Decay of mutualistic potential in aphid endosymbionts through silencing of biosynthetic loci: Buchnera of Diuraphis. *Proceedings of the Biological Society* **267**: 1423-1431.

Wernegreen, J. J. and N. A. Moran (2001) Vertical transmission of biosynthetic plasmids in aphid endosymbionts (*Buchnera*). *Journal of Bacteriology* **183**: 785-790.

Wheeler, A. G. (2001) Biology of the plant bugs (Hemiptera: Miridae): pests, predators, opportunists, Cornell University Press.

Wilkinson, T. L. (1998) The elimination of intracellular microorganisms from insects: an analysis of antibiotic-treatment in the pea aphid (*Acyrthosiphon pisum*). *Comparative Biochemistry and Physiology Part A* **119**: 871-881.

Will, T., W. F. Tjallingii, A. Thönnessen and A. J. E. van Bel (2007) Molecular sabotage of plant defense by aphid saliva. *Proceedings of the National Academy of Sciences, USA* **104**: 10536-10541.

Will, T. and A. J. E. van Bel (2006) Physical and chemical interactions between aphids and plants. *Journal of Experimental Botany* **57**: 729-737.

Xiao, S., S. Ellwood, O. Calis, E. Patrick, T. Li, M. Coleman and J. G. Turner (2001) Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by *RPW8*. *Science* **291**: 118-120.

Yamamoto, A., N. H. Bhuiyan, R. Waditee, Y. Tanaka, M. Esaka, K. Oba, A. T. Jagendorf and T. Takabe (2005) Suppressed expression of the apoplastic ascorbate oxidase gene

increases salt tolerance in tobacco and *Arabidopsis* plants. *Journal of Experimental Botany* **56**: 1785-1796.

Yoo, B.-C., J.-Y. Lee and W. J. Lucas (2002) Analysis of the complexity of protein kinases within the phloem sieve tube system. *The Journal of Biological Chemistry* **277**: 15325-15332.

Young, N. D. (2000) The genetic architecture of resistance. *Current Opinion in Plant Biology* **3**: 285-290.

Zaninotto, F., S. La Camera, A. Polverari and M. Delledonne (2006) Cross talk between reactive nitrogen and oxygen species during the hypersensitive disease resistance response. *Plant Physiology* **141**: 379-383.

Zemetra, R. S., D. Schotzko, C. M. Smith and E. J. Souza (1990) Seedling resistance to the Russian wheat aphid in white wheat germplasm. *Cereal Research Communications* **18**: 223-227.

Zhang, Y., J. S. Quick and S. Liu (1998) Genetic variation in PI 294994 wheat for resistance to Russian wheat aphid. *Crop Science* **38**: 527-530.

Zientz, E., T. Dadndekar and R. Gross (2004) Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiology and Molecular Biology Reviews* **68**: 745-770.

Zwer, P. K., M. G. Mosaad, A. A. Elsidaig and R. W. Rickman (1994) Effect of Russian wheat aphid on wheat root and shoot development in resistant and susceptible genotypes. *Crop Science* **34**: 650-655.

CHAPTER 3

LIMITED ENDOSYMBIONT VARIATION IN *DIURAPHIS NOXIA* (HEMIPTERA: APHIDIDAE) BIOTYPES FROM THE USA AND SOUTH AFRICA

Published in part as:

Swanevelder Z.H., A.K.J. Surridge, E. Venter and A-M. Botha (2010) Limited endosymbiont variation in *Diuraphis noxia* (Hemiptera: Aphididae) biotypes from the USA and South Africa. *Journal of Economic Entomology* **103:** 887-897.

Introduction

Diuraphis noxia Kurdjumov (Russian wheat aphid, RWA), originally from Central Asia through to the Middle East (Stary 2000; Stary & Lukasova 2002), has increased its distribution to all cereal producing countries except for the region of Australasia (Blackman & Eastop 2000; Baker *et al.* 2003). Introduction to these new areas were shown to have low biotypic diversity (Puterka *et al.* 1993), probably due to a small founder effect. Recently new RWA biotypes appeared in the USA and South Africa (Haley *et al.* 2004; Burd *et al.* 2006; Tolmay *et al.* 2007; Weiland *et al.* 2008) implying either new introductions or adaptation and diversification of existing populations. Molecular analysis of the new biotypes supports a diversification theory (Lapitan *et al.* 2007). However, these new aphid biotypes still have little nuclear and mitochondrial sequence variation (Lapitan *et al.* 2007; Shufran *et al.* 2007). RWA biotypes are distinguished on their ability to overcome resistance, their fecundity and the damage that they cause to a plant differential, but not on aphid morphology (Puterka *et al.* 1992; Jyoti & Michaud 2005; Burd *et al.* 2006; Jyoti *et al.* 2006; Weiland *et al.* 2008). However, little information on these biotypes' bacterial endosymbionts exists.

Symbiosis enabled aphids to exploit a nutritionally imbalanced food source such as phloem (Srivastava 1987; Douglas 1998). The usual partner in this relationship, the bacterium *Buchnera aphidicola*, is maintained transovarially and maternally between generations inside aphid produced cells called mycetocytes (bacteriocytes) (Munson *et al.* 1991; Baumann *et al.* 1995; Dixon 1998; Douglas 1998). The *Buchnera*-host relationship is not only complex but also obligatory to both parties, with this gram-negative bacterium found in almost all aphid species. Removal of the endosymbiont from its host usually leads to sterile aphid offspring (Munson *et al.* 1991; Baumann *et al.* 1995; Douglas 1998). The dependency of the endosymbiont on its host was clearly illustrated by genome sequencing that showed many genes from key pathways were absent in the bacterium (Shigenobu *et al.* 2000; Tamas *et al.* 2002; Van Ham *et al.* 2003; Zientz *et al.* 2004).

Pathways retained by *B. aphidicola*, like the essential amino acid biosynthetic pathway, suggest that this endosymbiont still plays a crucial role in aphid nutrition (Shigenobu *et al.* 2000; Tamas *et al.* 2002; Van Ham *et al.* 2003; Nakabachi *et al.* 2005). The type of pathway genes retained appear to be diet dependent (Tamas *et al.* 2002; Zientz *et al.* 2004), *i.e. B. aphidicola* is responsible for the production and recycling of certain essential amino acids found in low quantities in the aphid's diet (Mittler 1971; Douglas & Prosser 1992; Douglas 1998). The production of these essential amino acids in sufficient quantities is accomplished by duplicating genes and by moving rate limiting enzymes (anthranilate synthase, *ptrpEG*), or even whole

pathways (leucine biosynthetic pathway, *pleuABCD*), to single or multiple copy plasmids (Lai *et al.* 1994; Van Ham *et al.* 1997; Baumann *et al.* 1999). These plasmid copy numbers are known to differ between biotypes of the same species (Moran *et al.* 2003).

However, lower plasmid copy numbers and the presence of pseudogenes have led to the belief that the symbiotic relationship between *D. noxia* and *B. aphidicola* is concurrently diverging from essential amino acid production (Lai *et al.* 1996; Thao *et al.* 1998). This hypothesis is further supported by observations that the RWA alters protein profiles and up-regulates leucine and tryptophan levels in the phloem of susceptible wheat cultivars, but not in resistant lines (Van der Westhuizen & Pretorius 1996; Telang *et al.* 1999; Porter & Webster 2000; Sandstrom *et al.* 2000; Ni *et al.* 2001).

In addition to the primary endosymbiont, *B. aphidicola*, aphids may contain an independently acquired secondary (facultative) symbiont(s) (Unterman *et al.* 1989; Chen *et al.* 1996; Darby *et al.* 2001; Fukatsu 2001). Evidence suggests that these can be horizontally transferred when aphids feed on contaminated host plants (Sandstrom *et al.* 2001; Russell *et al.* 2003; Russell & Moran 2005) and that even horizontal plasmid transfer to *B. aphidicola* could be possible (Van Ham *et al.* 2000). In *Bemisia tabaci* (sweet potato whitefly) certain secondary symbionts are biotype specific (Chiel *et al.* 2007). Facultative symbionts can contribute to the well being of their aphid hosts through inferring resistance against predators such as parasitic wasps (Oliver *et al.* 2003; Ferrari *et al.* 2004) or pathogenic fungi (Scarborough *et al.* 2005), an ability to withstand higher temperatures (Chen *et al.* 2000; Montllor *et al.* 2002), higher fecundity (Chen *et al.* 2000) and in some instances could even compensate for the loss of *B. aphidicola* (Koga *et al.* 2003). Similarly, some may have a negative impact on aphids, manifested as slower growth rate, a decrease in reproduction and a shorter life span (Chen *et al.* 2000; Leonardo 2004; Ferrari *et al.* 2007).

The importance of this symbiont-host relationship was recently illustrated in the *B. aphidicola* found in *Acyrthosiphon pisum*, where a single point mutation proved to be crucial for the survival of the host at different temperatures (Dunbar *et al.* 2007). Changes in symbiont diversity or small changes within the symbiont(s) can therefore have major effects on the aphid's viability and adaptation potential. The possible role of symbiont(s) on RWA biotype development and the reverse, *i.e.* the effect of biotype formation on the symbiont(s), have to date not been investigated. Here these interactions are investigated in ten *D. noxia* biotypes by determining the complete endosymbiont assemblages, as well as sequence variation and copy numbers of the leucine and tryptophan plasmids from their primary endosymbiont, *B. aphidicola*.

Materials and Methods

Diuraphis species and biotypes

The originally introduced South African biotype (SA) (Du Toit 1989) was obtained from a colony established on field collected parthenogenetic females at the ARC-Small Grains Institute, Bethlehem, South Africa and sustained on a RWA susceptible wheat cultivar, Scheepers. A mutated form of the South African biotype (SAM), laboratory induced by *Dn* resistant selective pressure, was maintained on the resistant cultivar, TugelaDN (Van Zyl & Botha 2008). Females of both biotypes were kept in insect cages at 20 ± 2 ºC with continuous artificial fluorescent lighting. Samples of the eight USA biotypes were obtained from Prof. N.L.V. Lapitan (Colorado State University, Fort Collins, USA). Biotype USA1 is the original USA 1986 RWA introduction and USA2 is characterized by its virulence to *Dn4* resistant winter wheat cultivars (Haley *et al.* 2004). The first five USA biotypes were from cultivated wheat; USA6 from both volunteer wheat and downy brome; USA7 collected from volunteer wheat, crested wheatgrass, Canada wildrye, green foxtail and intermediate wheatgrass, while USA8 was collected from crested wheatgrass and smooth brome (Lapitan *et al.* 2007; Weiland *et al.* 2008). USA biotypic identifiers are described in detail by Weiland *et al*. (2008). All USA biotypes were isofemale lines maintained on a mixed diet of susceptible wheat and barley cultivars under greenhouse conditions (Lapitan *et al.* 2007). Individuals of *D. mexicana* and *D. tritici* were collected on mountain brome in Colorado, USA, and kindly provided by Dr. G.J. Puterka (USDA-ARS, Stillwater, Oklahoma, USA). Apterae (wingless) morphs were collected with a brush, washed with 70 % (v/v) ethanol and rinsed twice with sterilized distilled water before DNA extraction. Total DNA was extracted using the DNAzol extraction protocol (Molecular Research Centre). DNA of the USA biotypes was extracted as described in Lapitan *et al*. (2007) and shipped at -20 °C to South Africa. All DNA samples were treated with RNase, cleaned (DNeasy cleanup kit, Qiagen, USA) and quantified (Nanodrop ND-1000 Spectrophotometer, Thermo Scientific, RSA).

Biotypic endosymbiont investigation

A portion of the *16S rDNA* gene was amplified for use in denaturing gradient gel electrophoresis (DGGE) analysis using two universal primers, one containing a GC-clamp (*Table Appx 3.1*, Appendix Chapter 3). All PCR reagents and conditions used in this research are described in *Table Appx 3.1*. The GC-clamp increases fragment separation, enabling the detection of up to single point mutation (Myers *et al.* 1985; Sheffield *et al.* 1989; Muyzer & Smalla 1998). Amplicons, after verification on agarose gels, were run in duplicate on 8 %

polyacrylamide gels with a denaturing gradient of 25-55 % urea-formamide. DGGE was performed as described in Surridge (2007) and stained with SYBR Gold (Molecular Probes, USA). Bands were viewed under a blue light trans-illuminator (model CCR DR-88M DR, Inqaba-Biotec), excised, re-amplified and sequenced.

Buchnera aphidicola sequence variation amongst biotypes

The leucine plasmid, *pleuABCD* and parts of the *ptrpEG* plasmid, *16S rDNA* and *trpB*, were amplified and sequenced (ABI BigDye v3.1. System, Applied Biosystems, USA) (*Table Appx 3.1*). Fragments were aligned using ContigExpress (Vector NTI Advance 9, Invitrogen, USA) (Lu & Moriyama 2004), with cloned fragments first vector clipped using VecScreen (http://www.ncbi.nlm.nih.gov/VecScreen/). Inconsistencies in assemblies were manually investigated on the corresponding chromatograms and resolved with further sequencing. All BLAST analyses (Altschul *et al.* 1990; Altschul *et al.* 1997) were against the non-redundant Genbank database (NCBI, http://www.ncbi.nlm.nih.gov/).

Structural analysis

Bacterial promoters on the leucine plasmid were predicted with BPROM (http://softberry.com). Quickfold (Markham & Zuker 2005) and the Kinefold server (Xayaphoummine *et al.* 2005), using simulations of stochastic folding pathways from different random seed events, were used to predict the free energy values of the inverted repeat region upstream of *leuA*. The plasmids were screened for *Rho*-independent terminators using FindTerm (http://softberry.com).

Plasmid copy numbers

Plasmid copy numbers were determined according to Plague *et al*. (2003). External standards were amplified (*Table Appx 3.1*), cleaned (QIAquick PCR purification kit, Qiagen) and cloned into pGEM-T Easy vector (Promega, USA). Plasmid isolations were done with the Concert Rapid Plasmid Miniprep System (GibcoBRL, Life Technologies, USA) and single copy amplicon inserts confirmed with sequencing. Plasmid copy number for cloned external standards were determined spectrophotometrically in triplicate and converted to copy number per µg DNA (Plague *et al.* 2003). Real-time qPCRs were performed using forward and nested reverse primers of the different target genes (*Table Appx 3.1*). Primer optimization and qPCR reactions were done in accordance with the manufacturer's protocols (Bio-Rad, USA). The iQ SYBR Green Supermix (Bio-Rad) was used for quantitative PCRs and reactions performed on an iCycler with

an iQ Real-Time PCR Detection System (Bio-Rad). Single amplicons were confirmed for each qPCR with melting curve analysis and on 2 % TAE (40 mM Tris, 20 mM acetic acid, 1 mM EDTA (pH 8.3)) agarose gels. The iCycler Optical System Interface v2.3 was used to calculate the copy number for each sample using calibrated external standard curves of each gene. Plasmid number per chromosome gene was calculated for the leucine plasmid using *leuB*:*trpB* ratio (Plague *et al.* 2003).

Results

Biotypic endosymbiont investigation

Most aphids contain one or more endosymbionts. In order to establish the relationship between the aphid *D. noxia* and its particular endosymbiont, we have investigated the occurrence of mono-or-multisymbiosis in RWA biotypes. DGGE of the bacterial-*16S rDNA* (*Table 3.1*) amplicons showed that the ten *D. noxia* biotypes, *D. tritici* and *D. mexicana*, were all monosymbiotic.

The obtained *16S rDNA* DGGE gene fragments, from both the biotypes and other *Diuraphis* species, were sequenced to confirm the identity of the endosymbiont. Genbank database searches (Altschul *et al.* 1990; Altschul *et al.* 1997) had highly significant homologies with *B. aphidicola* (E-values: 0 to $7e^{-140}$), *Sodalis glossinidius* ($1e^{-67}$), *Erwinia chrysanthemi* ($2e^{-66}$) and an *Arsenophonus* endosymbiont ($3e^{-66}$). A Ribosomal Database (release 9.51) (Cole *et al.* 2006) search and phylogenetic analysis (*Figure Appx 3.1*) confirmed the endosymbiont as *B. aphidicola*, although the biotypes had homologous sequences that differed by four indels from the Genbank accession for *B. aphidicola* of *D. noxia* (*Table 3.1*).

Buchnera aphidicola sequence variation amongst biotypes

The importance of *Buchnera* in aphid nutrition is well known. However, the current hypothesis states that the relationship between the RWA and its endosymbiont is degrading. We have therefore targeted genes or plasmids involved in nutrition to investigate the aphid-endosymbiont relationship. To this end more than 10 kb was sequenced for *B. aphidicola* from each *D. noxia* biotype (*Table 3.1*). Although numerous differences to sequences on Genebank were found (*Table 3.1*), only a single CCC insertion on the leucine plasmid differs between the *Buchnera* sequences of RWA biotypes.

This CCC-insert was upstream of the *leuA* gene (*pleuABCD*) in an inverted repeat region of biotypes SA, SAM, USA3 and USA7. The biotypes' *pleuABCD* sequences also increased in

length, from 7768 (Genbank) to 7771 (without insert) and 7774 bp (with insert). Other differences were also observed between the leucine plasmid sequences of the biotypes and previously submitted Genbank accessions (AF041837, NC001911 and *Figure Appx 3.2*). These include various insertions, deletions or point mutations, for example a single nucleotide insertion in a non-coding region between the *leuA* and *leuB* genes transposed them into the same open reading frame (*Figure Appx 3.2*). The differences in the sequenced leucine plasmids also altered the AT-contents causing an increase from 74.11 % to 74.39 % and 74.36 %, when compared with the previously deposited Genbank accessions (AF041837, NC001911).

Table 3.1 Sequences analyzed and the differences observed against homologues on Genbank and between the different biotypes.

Fragment	***Size in base pairs***	***Genbank differences***	***Biotype differences***
Bacterial chromosome			
16S rDNA[#]	460	4	0
trpB[#]	342	0	0
Bacterial plasmids			
pleuABCD[#]	7768-7774	60	1 (CCC-insert)
ptrpEG[#]	~2100	3	0
Aphid mitochondrion sequences			
COI across the distribution range	Variable	1	-

[#]Fragments sequenced in this project and used in the comparisons (Genbank accession numbers FJ705277-FJ705296; FJ705299-FJ705318). Genebank accessed November 2008.

Primary and secondary structural analysis

The primary and secondary structures of DNA are known to play a crucial role in gene expression. Thus, changes in it may alter secondary structures and therefore functionality of the gene region. To date, no differences within a species were shown for this region (*Figure 3.1*). We therefore simulated the secondary folding patterns to predict free energy values for the inverted repeat region upstream of *leuA*.

Analyses revealed that the most stable stemloop structure produced by this inverted repeat region is formed in biotypes without the insert, with a predicted free energy of -63.4 kcal mol^{-1} (Kinefold server, Xayaphoummine *et al.* 2005), while biotypes containing the CCC-insert gave a free energy value of -57.9 kcal mol^{-1}. Similarly, Quickfold (Markham & Zuker 2005) also predicted lower free energy values for DNA from the biotypes without the CCC-insert (*Figure 3.2*), though both values were lower than those predicted by Kinefold server.

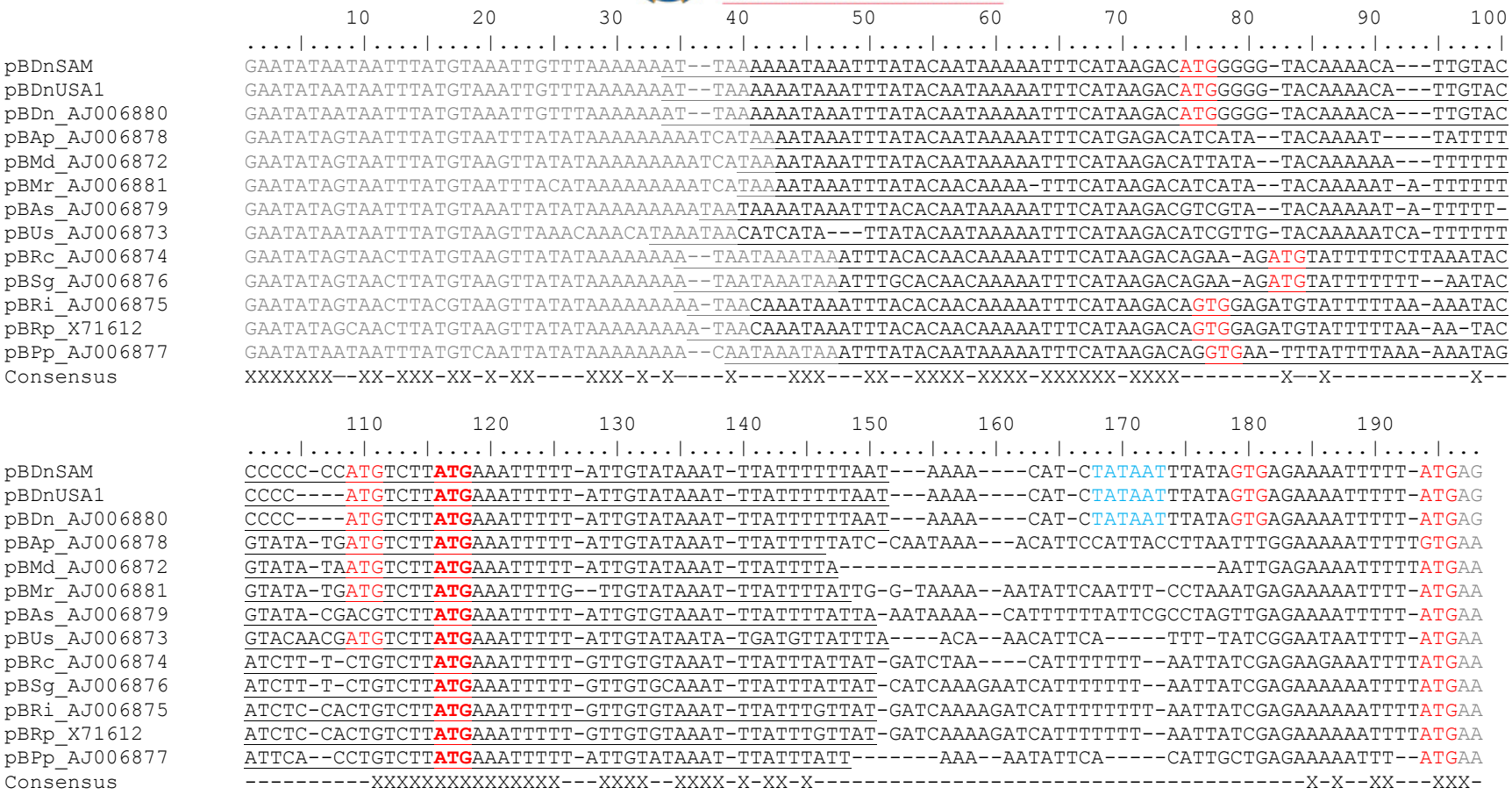


Figure 3.1 The inverted repeat region between *repA2* and *leuA* on leucine plasmids of the Aphididae. The end of *repA2* and the start of *leuA* are indicated in grey, with possible start codons indicated in red. Abbreviations: 'pB' indicates the leucine plasmid of *B. aphidicola* which is then followed by the aphid species, 'Dn' and biotype 'SAM': pBDnSAM, *D. noxia* biotype SAM; pBDnUSA1, *D. noxia* biotype USA1; pBDn, *D. noxia* (Genbank); pBAp, *Acyrthosiphon pisum*; pBMd, *Metapolophium dirhodum*; pBMr, *Macrosiphum rosae*; pBAs, *Aulacorthum solani*; pBUs, *Uroleucon sonchi*; pBRc, *Rhopalosiphum cerasifoliae*; pBSg, *Schizaphis graminum*; pBRi, *R. insertum*; pBRp, *R. padi*; pBPp, *Ptercomma populeum*.

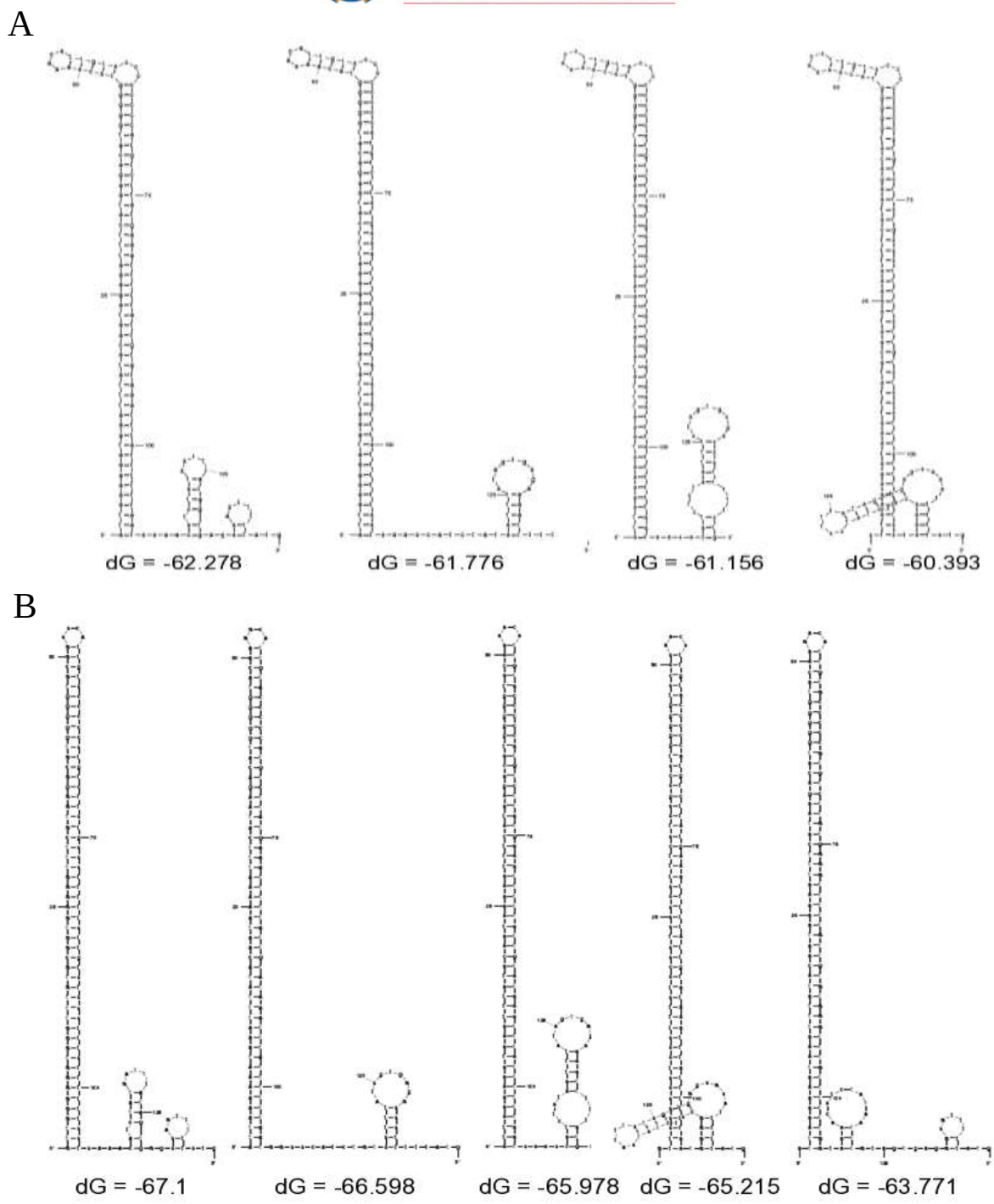


Figure 3.2 Minimum free energy diagrams for the inverted repeat region up to the start codon of *leuA* of *D. noxia* biotypes as predicted with Quickfold (DNA and temperature 25 °C, DINAMelt Server) (Markham & Zuker 2005), with (*Figure 3.2A*) and without (*Figure 3.2B*) CCC-insert. Similar structures and energy values were also obtained for predicted RNA structures of the same region. In both cases the CCC-insert resulted in an increase of ~5 kcal mol^{-1} free energy between the predicted structures. All free energy values (dG) are given as kcal mol^{-1}.

Interestingly, *B. aphidicola* has only two predicted sigma factors, σ^{32} (*rpoH*) and σ^{70} (*rpoD*) (Shigenobu *et al.* 2000). A BPROM promoter search of the regions adjacent to the CCC-insert predicted that it could form a new transcription factor (TF) binding site with homology to *rpoH3* TF binding site. The first predicted BPROM promoter has homology with two σ^{70} TF binding sites and a better linear discriminant function (LDF) score than the new σ^{32} binding site that is formed by the CCC-insert.

The leucine plasmid has two predicted *Rho*-independent terminators, one following *leuD* and the other, following the *repA2* gene. FindTerm (http://softberry.com) predictions high-lighted the second inverted repeat region, the one in which the CCC-insert is located, as the best *Rho*-independent terminator on the plasmid. The CCC-insert is located upstream of the *leuA* gene (*Figure 3.3*) and downstream of a predicted Sigma70 promoter (BPROM, http://softberry.com), TATAAT box and start codon, and could therefore be transcribed.

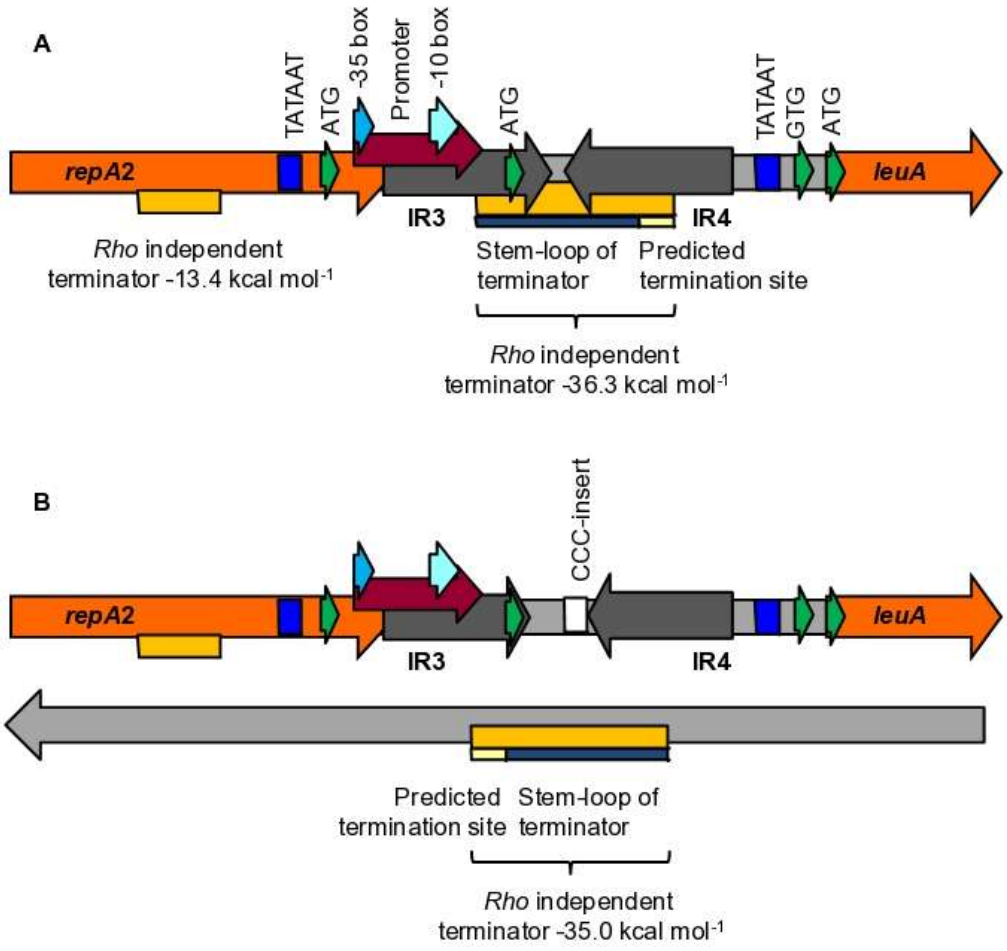


Figure 3.3 The region between *repA2* and *leuA* indicating predicted promoter and *Rho* independent terminator regions of RWA biotypes without (*Figure 3.3A*) and with (*Figure 3.3B*) the CCC-insert. The CCC-insert shifts the predicted *Rho* independent terminator from the coding to the non-coding strand (*B*). This increases the predicted free energy value of the terminator structure, suggesting a less stable terminator.

Plasmid copy numbers

In order to investigate the possible role of the endosymbiont in the aphid-endosymbiont relationship, we have studied the plasmid copy numbers in *Buchnera*. It is suggested that individuals with higher copy numbers may have advantages in terms of fitness, as higher copy numbers supply the aphid with more essential amino acids, while lower copy numbers and pseudogenes indicate a degenerating relationship with reduced access to essential amino acid availability (Lai *et al.* 1996; Thao *et al.* 1998).

We have found that an average of 0.9 copies/bacterial chromosome for *pleuABCD*. The South African RWA biotypes had lower *pleuABCD* copies than the USA biotypes (*Figure 3.4*), while USA8, had more than double the previously published number of copies (2 copies per bacterial genome, *Figure 3.4*).

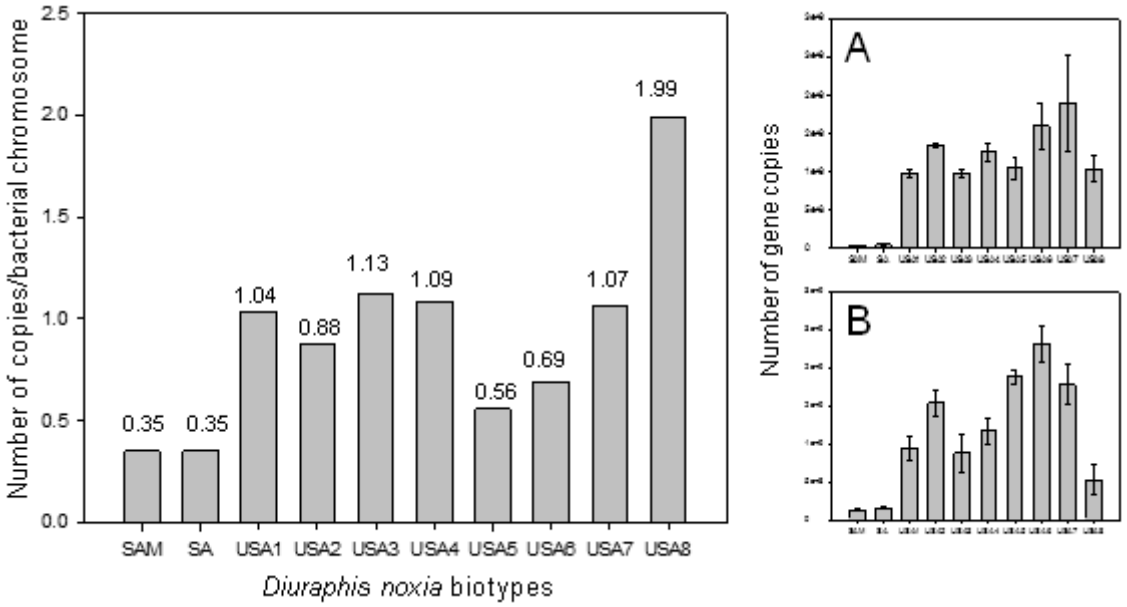


Figure 3.4 The ratio *leuB:trpB* was used to calculate the copy number of the different biotypes. The mean copy numbers for *leuB* (insert A, n=4) and of *trpB* (insert B, n=4) were determined for the different biotypes and then used to calculate the *leuB:trpB* ratio. South African biotypes had lower copy numbers than their USA counterparts which is supported by literature. The average across all the biotypes was 0.91 copies per chromosomal gene.

Discussion

Biotypic endosymbiont investigation

DGGE has been used to investigate associated microbial endosymbionts in a variety of insects, including wasps (Reeson *et al.* 2003), whiteflies (Gottlieb *et al.* 2006), ants (Stoll *et al.* 2007), ticks (Moreno *et al.* 2006) and aphids (Haynes *et al.* 2003). In the present study the DGGE fragments supplied sufficient phylogenetic resolution to determine *B. aphidicola* accession up to the host level (*Figure Appx 3.1*). Although aphids have been shown to contain secondary endosymbionts in addition to *B. aphidicola*, the contribution of a secondary symbiont in RWA biotype development was not supported by DGGE-based data since the *Diuraphis* species and biotypes analyzed were all monosymbiotic.

Buchnera aphidicola sequence variation amongst biotypes

A single CCC insertion on the leucine plasmid differs between the sequences of the RWA biotypes after more than 10 kb was sequenced for *B. aphidicola* from each *D. noxia* biotype. This insert is located upstream of the *leuA* gene (*pleuABCD*) in an inverted repeat region in the South African biotypes and two US biotypes, namely USA3 and USA7. Interestingly, the CCC-insert is located within the variable region of the inverted repeat stem and is consistent with sequence variation found within the family Aphididae (Silva *et al.* 1998). The insert also results in an increase in sequence length, this increase in size is in sharp contrast with other publications where it was shown that plasmids were shrinking in accordance with the genome reduction usually observed in *B. aphidicola* (Gil *et al.* 2006). Mutations, *e.g.* single nucleotide insertion in non-coding region between the *leuA* and *leuB* genes that transposed them into the same open reading frame, were also observed. The CCC-insertion could change the expression levels of *leuA*, and thus would also change the expression levels of the other portions of the operon, *i.e. leuB*.

The rate of sequence change in endosymbiotic bacteria is higher than in free living bacteria (Baumann 2005), therefore more sequence divergence is expected. This is supported by the genome comparison of *B. aphidicola* from different aphid species that showed extremely stable genomes with no re-arrangements or gene acquisitions, but with substantial sequence evolution and few gene losses (Tamas *et al.* 2002; Van Ham *et al.* 2003). In contrast, very little sequence variation is usually found within *B. aphidicola* from the same population (Funk *et al.* 2001; Abbot & Moran 2002; Van Ham *et al.* 2003). These observations are supported by data in the present study where only a single insert has been found to vary among the different biotypes.

Similarly, limited sequence variation was previously found for these biotypes' mitochondrial cytochrome oxidase subunit I (*COI*) gene (*Table 3.1*) (Lapitan *et al.* 2007; Shufran *et al.* 2007). Phylogenetic analysis of these and sequences in *Table 3.1* support the same relationships as determined by aphid and mitochondrial genes (*Figure 3.5*).

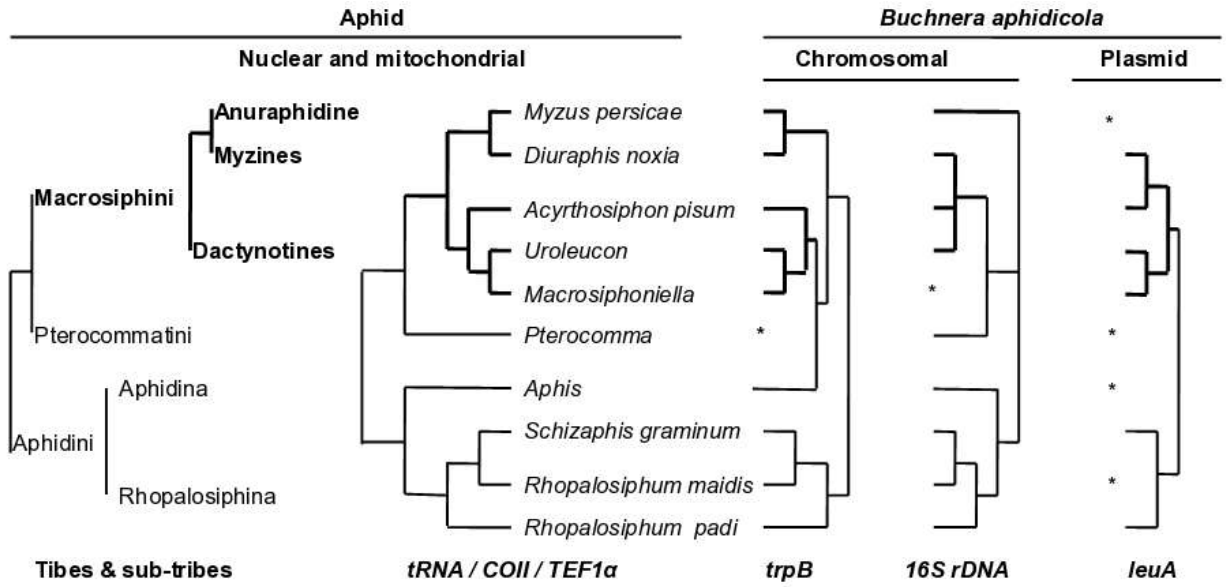


Figure 3.5 Phylogenetic relationships between tribal representatives of the subfamily Aphidinae. The first two phylogenies are an abbreviated representation of a combined nuclear and mitochondrial dataset of *tRNA*, *COII* and *TEF1α*, from von Dohlen *et al.* (2006). Bold lines indicate members of the tribe Macrosiphini. Species names are given where data are available for the phylogenies and genus names when different species of the same genus were used. An asterisk indicates absent genera.

Primary and secondary structural analysis

The inverted repeat region, where the insert is located, has known secondary structural differences and is conserved within the family Aphididae (Silva *et al.* 1998). The most stable secondary structure is the one with the lowest free energy, which is usually the structure with the most bases paired. To this end, most secondary structure prediction algorithms identify the best secondary structure by trying to minimize free energy (Wuchty *et al.* 1999). Since both these prediction algorithms (Quickfold and Kinefold server) awarded higher free energy values for the CCC-insert containing regions, it is highly likely that this region forms a less stable structure.

Conservation of this entire region between *repA2* and *leuA*, including the inverted repeats, within *B. aphidicola* found associated with the family Aphididae (*Figure 3.1*) suggests a functional constraint (Silva *et al.* 1998). A start codon (methionine), preserved in all the

members of the family, is located in the conserved region of the stemloop just following the CCC-insert (*Figure 3.1*). Destablization of this region by the CCC-insert, as predicted by free energy values (*Figure 3.2*) could thus allow easier access to this start codon.

BPROM predictions suggest that the first promoter has homology to two σ^{70} TF binding sites and a better linear discriminant function (LDF) score than the new σ^{32} binding site that is formed by the CCC-insert. However, predicting promoters accurately in AT-rich genomes, such as those of *Buchnera*, is extremely difficult (Baumann *et al.* 1995). If the new CCC-insert-induced σ^{32} binding site is functional, then biotypes with this insert could produce more leucine. Since the CCC-insert also forms an imperfect stemloop (*Figure 3.2*) which destabilizes the secondary DNA structure, the TF binding site would be easier accessed during transcription.

Interestingly, the FindTerm program predicted that the terminating site of this *Rho*-independent terminator following the *repA2* gene would change from the coding strand, in the plasmids without the CCC-insert, to the non-coding strand in the presence of the CCC-insert (*Figure 3.3*). If the CCC-insert moves the termination site to the non-coding strand, transcription termination could be deactivated and upstream promoters could come into play in regulating *leuA*. Because of the location of the CCC-insert, and the possibility that this region may be transcribed, the stemloop/hairpin structure formed by the inverted repeats could act in stabilising the mRNA of either *repA2* or *leuA*. Both 5'- and 3'-terminal hairpin structures are known to act as mRNA stabilizers in bacteria (Wong & Chang 1986; Emroy *et al.* 1992). In all cases this region could play a role in gene expression, mRNA stabilization or translational regulation.

Plasmid copy numbers

The observed average of 0.9 plasmid copies/bacterial chromosome for *pleuABCD* was higher than the previously published value of 0.8 (Thao *et al.* 1998). Interestingly, the South African biotypes had lower *pleuABCD* copies than the US biotypes. This, together with the CCC-insert that is absent in the original US biotype, suggests that the US RWA introduction did not originate from South Africa, thus supporting previous findings of multiple introductions (Lapitan *et al.* 2007; Smith 2009). In contrast, USA8 a RWA biotype only collected from species other than wheat, *i.e. Agrophyron cristatum*, had more than double the previously published number of copies (2 copies per bacterial genome, *Figure 3.4*). The data on the copy numbers of the leucine plasmid supports the hypothesis that the CCC-insert is a functional mutation. The South African biotypes, with the CCC-insert, have lower leucine copy numbers than their US counterparts. One may argue that the US biotypes with the inserts (USA3 and USA7) have only recently developed, and as such have not yet had sufficient time to lower their copy numbers.

Gene amplification allows for adaptation by enabling the over expression of specific gene products necessary for survival in a changing environment or biological interaction (Romero & Palacios 1997). This state would only remain while selective conditions were in effect (Lai *et al.* 1994; Romero & Palacios 1997). Furthermore, having only a single transcriptional regulator for essential amino acid biosynthesis known to exist in a single *B. aphidicola* accession, plasmid copy number is the most likely regulatory mechanism (Moran *et al.* 2005). Plasmid copy numbers are highly variable between and within aphid species (Plague *et al.* 2003). The same tendency was observed for the *pleuABCD* plasmids in *D. noxia*.

Wide-host range selective pressure may explain the high *pleuABCD* copy number of the USA8 biotype. SAM, though kept under selective pressure by feeding aphids resistant wheat cultivars, developed out of the original SA biotype. The low copy number observed here shows the close relationship between SA and SAM. The observed *pleuABCD* copy numbers for the South African biotypes are similar to the 0.3 copies per bacterial genome obtained by Moran *et al*. (2003) for the wild type SA biotype. Copy number suggests that the symbiosis could be involved when the RWA feeds on hosts other than wheat, but no differences in copy number or sequence structure exist between biotypes feeding on resistant and susceptible wheat cultivars.

The observation that genetic diversity of *B. aphidicola* cannot explain the ecological diversity observed between aphids holds here for the different biotypes, and *B. aphidicola*'s initial contribution in *Diuraphis* adaptation is still supported (Tamas *et al.* 2002; Van Ham *et al.* 2003). The CCC-insert and lower copy numbers may support the idea of symbiotic degradation between *D. noxia* and *B. aphidicola* for the South African biotypes, as suggested by Wernegreen and Moran (2000), but the opposite can be argued for one of the US biotypes (USA8) where higher copy numbers for the leucine plasmid have been observed, and for those biotypes where the CCC-insert occurs.

Conclusions

Aphids feeding on different plant species or cultivars have different requirements of their endosymbionts since the essential amino acid content of plants varies (Sandstrom & Moran 1999). This research shows that *B. aphidicola* of different RWA biotypes showed little variation in sequence, but differed in plasmid copy numbers. However, small variations in *B. aphidicola* have major implications for host viability (Dunbar *et al.* 2007). Varying *B. aphidicola* plasmid copy numbers may allow some measure of adaptation to the host where essential amino acids cannot be altered in the plant. Therefore, the role of the CCC-insert is unclear: it may confirm the suggested degeneration of the symbiotic relationship between the RWA and *B. aphidicola* if it

causes down-regulation of subsequent genes; or the opposite could occur and the genes may be up-regulated. It could also be non-functional indicating only normal variation for the region as observed within the family. Keeping symbiosis in mind may still prove to be the key in understanding biotype development within aphids.

Acknowledgements

The authors would like to thank Prof N.L.V. Lapitan (Colorado State University), Dr. V.L. Tolmay (ARC-Small Grains Institute, Bethlehem, South Africa) and Dr. G.J. Puterka (USDA-ARS, Stillwater, Oklahoma, USA) for aphid samples. This work was funded by the Winter Cereal Trust and the National Research Foundation of South Africa.

References

Abbot, P. and N. A. Moran (2002) Extremely low levels of genetic polymorphism in endosymbionts (*Buchnera*) of aphids (*Pemphigus*). *Molecular Ecology* **11**: 2649-2660.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389-3402.

Baker, D. A., H. D. Loxdale and O. R. Edwards (2003) Genetic variation and founder effects in the parasitoid wasp, *Diaeretiella rapae* (M'intosh) (Hymenoptera: Braconidae:Aphidiidae), affecting its potential as a biological control agent. *Molecular Ecology* **12**: 3303-3311.

Baumann, L., P. Baumann, N. A. Moran, J. Sandstrom and M. L. Thao (1999) Genetic characterization of plasmids containing genes encoding enzymes of leucine biosynthesis in endosymbionts (*Buchnera*) of aphids. *Journal of Molecular Evolution* **48**: 77-85.

Baumann, P. (2005) Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annual Review Microbiology* **59**: 55-89.

Baumann, P., C.-Y. Lai, L. Baumann, D. Rouhbakhsh, N. A. Moran and A. M. Clark (1995) Genetics, physiology and evolutionary relationships of the genus *Buchnera*: intracellular symbionts of aphids. *Annual Review of Microbiology* **49**: 55-94.

Blackman, R. L. and V. F. Eastop (2000) Aphids on the world's crops. An identification and information guide. London, Wiley.

Burd, J. D., D. R. Porter, G. J. Puterka, S. D. Haley and F. B. Peairs (2006) Biotypic variation among North American Russian wheat aphid (Homoptera: Aphididae) populations. *Journal of Economic Entomology* **99**: 1862-1866.

Chen, D.-Q., B. C. Campbell and A. H. Purcell (1996) A new *Rickettsia* from a herbivorous insect, the pea aphid *Acyrthosiphon pisum* (Harris). *Current Microbiology* **33**: 123-128.

Chen, D.-Q., C. B. Montllor and A. H. Purcell (2000) Fitness effects of two facultative endosymbiotic bacteria on the pea aphid, *Acyrthosiphon pisum*, and the blue alfalfa aphid, *A. kondoi*. *Entomologia Experimentalis et Applicata* **95**: 315-323.

Chiel, E., Y. Gottlieb, E. Zchori-Fein, N. Mozes-Daube, N. Katzir, M. Inbar and M. Ghanim (2007) Biotype-dependent secondary symbiont communities is sympatric populations of *Bemisia tabaci*. *Bulletin of Entomological Research* **97**: 407-413.

Cole, J. R., B. Chai, R. J. Farris, Q. Wang, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, A. M. Bandela, E. Cardenas, G. M. Garrity and J. M. Tiedje (2006) The ribosomal database project (RDP-II): introducing *myRDP* space and quality controlled public data. *Nucleic Acids Research* **35**: D169-D172.

Darby, A. C., L. M. Birkle, S. L. Turner and A. E. Douglas (2001) An aphid-borne bacterium allied to the secondary symbionts of whitefly. *FEMS Microbiology Ecology* **36**: 43-50.

Dixon, A. F. G. (1998) Aphid ecology: An optimization approach. London, Chapman & Hall.

Douglas, A. E. (1998) Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annual Review of Entomology* **43**: 17-37.

Douglas, A. E. and W. A. Prosser (1992) Synthesis of the essential amino acid tryptophan in the pea aphid (*Acyrthosiphon pisum*) symbiosis. *Journal of Insect Physiology* **38**: 565-568.

Du Toit, F. (1989) Inheritance of resistance in two *Triticum aestivum* lines to Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **82**: 1251-1253.

Dunbar, H., A. C. C. Wilson, N. R. Ferguson and N. A. Moran (2007) Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLOS Biology* **5**: e96 [1006-1015].

Emroy, S. A., P. Bouvet and J. G. Belasco (1992) A 5'-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes & Development* **6**: 135-148.

Ferrari, J., A. C. Darby, T. J. Daniell, H. C. J. Godfray and A. E. Douglas (2004) Linking the bacterial community in pea aphids with host-plant use and natural enemy resistance. *Ecological Entomology* **29**: 60-65.

Ferrari, J., C. L. Scarborough and H. C. J. Godfray (2007) Genetic variation in the effect of a facultative symbiont on host-plant use by pea aphids. *Oecologia* **153**: 323-329.

Fukatsu, T. (2001) Secondary intracellular symbiotic bacteria in aphids of the genus *Yamatocallis* (Homoptera: Aphididae: Drepanosiphinae). *Applied and Environmental Microbiology* **67**: 5315-5320.

Funk, D. J., J. J. Wernegreen and N. A. Moran (2001) Intraspecific variation in symbiont genomes: bottlenecks and the aphid-*Buchnera* association. *Genetics* **157**: 477-489.

Gil, R., B. Sabater-Munoz, V. Perez-Brocal, F. J. Silva and A. Latorre (2006) Plasmids in the aphid endosymbiont *Buchnera aphidicola* with the smallest genomes. A puzzling evolutionary story. *Gene* **370**: 17-25.

Gottlieb, Y., M. Ghanim, E. Chiel, D. Gerling, V. Portnoy, S. Steinberg, G. Tzuri, A. R. Horowitz, E. Belausov, N. Mozes-Daube, S. Kontsedalov, M. Gershon, S. Gal, N. Katzir and E. Zchori-Fein (2006) Identification and localization of a *Rickettsia* sp. in *Bemisia tabaci* (Homoptera: Aleyrodidae). *Applied and Environmental Microbiology* **72**: 3646-3652.

Haley, S. D., F. B. Peairs, C. B. Walker, J. B. Rudolph and T. L. Randolph (2004) Occurrence of a new Russian wheat aphid biotype in Colorado. *Crop Science* **44**: 1589-1592.

Haynes, S., A. C. Darby, T. J. Daniell, G. Webster, F. J. F. van Veen, H. C. J. Gofdfray, J. I. Prosser and A. E. Douglas (2003) Diversity of bacteria associated with natural aphid populations. *Applied and Environmental Microbiology* **69**: 7216-7223.

Jyoti, J. L. and J. P. Michaud (2005) Comparative biology of a novel strain of Russian wheat aphid (Homoptera: Aphididae) on three wheat cultivars. *Journal of Economic Entomology* **98**: 1032-1039.

Jyoti, J. L., J. A. Qureshi, J. P. Michaud and T. J. Martin (2006) Virulence of two Russian wheat aphid biotypes to eight wheat cultivars at two temperatures. *Crop Science* **46**: 774-780.

Koga, R., T. Tsuchida and T. Fukatsu (2003) Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the essential endosymbiont *Buchnera* in an aphid. *Proceedings of the Royal Society of London* **270**: 2543-2550.

Lai, C.-Y., L. Baumann and P. Baumann (1994) Amplification of *trpEG*: adaptation of *Buchnera aphidicola* to an endosymbiotic association with aphids. *Proceedings of the National Academy of Sciences, USA* **91**: 3819-3823.

Lai, C.-Y., P. Baumann and N. A. Moran (1996) The endosymbiont (*Buchnera* sp.) of the aphid *Diuraphis noxia* contains plasmids consisting of *trpEG* and tandem repeats of *trpEG* pseudogenes. *Applied and Environmental Microbiology* **62**: 332-339.

Lapitan, N. L. V., Y.-C. Li, R. S. G. Walters, Y. Peng, F. B. Peairs and A.-M. Botha (2007) Limited nuclear and mitochondrial DNA variation among Russian wheat aphid (*Diuraphis noxia*) biotypes from the United States and Africa, *American Entomological Society*, San Diego, December 9-12. American Entomological Society, San Diego.

Leonardo, T. E. (2004) Removal of a specialization-associated symbiont does not affect aphid fitness. *Ecology Letters* **7**: 461-468.

Lu, G. and E. N. Moriyama (2004) Vector NTI, a balanced all-in-one sequence analysis suite. *Briefings in Bioinformatics* **5**: 378-388.

Markham, N. R. and M. Zuker (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Research* **33**: 577-581.

Mittler, T. E. (1971) Dietary amino acid requirements of the aphid *Myzus persicae* affected by antibiotic uptake. *The Journal of Nutrition* **101**: 1023-1028.

Montllor, C. B., A. Maxmen and A. H. Purcell (2002) Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. *Ecological Entomology* **27**: 189-195.

Moran, N. A., H. E. Dunbar and J. L. Wilcox (2005) Regulation of transcription in a reduced bacterial genome: nutrient-provisioning genes of the obligate symbiont *Buchnera aphidicola*. *Journal of Bacteriology* **187**: 4229-4237.

Moran, N. A., G. R. Plague, J. P. Sandstrom and J. L. Wilcox (2003) A genomic perspective on nutrient provisioning by bacterial symbionts of insects. *PNAS* **100**: 14543-14548.

Moreno, C. X., F. Fred Moy, T. J. Daniels, H. P. Godfrey and F. C. Cabello (2006) Molecular analysis of microbial communities identified in different developmental stages of *Ixodes scapularis* ticks from Westchester and Dutchess Counties, New York. *Environmental Microbiology* **8**: 761-772.

Munson, M. A., P. Baumann and M. G. Kinsey (1991) *Buchnera* gen. nov. and *Buchnera aphidicola* sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. *International Journal of Systematic Bacteriology* **41**: 566-568.

Muyzer, G. and K. Smalla (1998) Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie van Leeuwenhoek* **73**: 127-141.

Myers, R. M., S. G. Fisher, L. L.S. and T. Maniatis (1985) Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. *Nucleic Acids Research* **13**: 3131-3145.

Nakabachi, A., S. Shigenobu, N. Sakazume, T. Shiraki, Y. Hayashizaki, P. Carninci, H. Ishikawa, T. Kudo and T. Fukatsu (2005) Transcriptome analysis of the aphid

bacteriocyte, the symbiotic host cell that harbors an endocellular mutualistic bacterium, *Buchnera*. *Proceedings of the National Academy of Sciences, USA* **102**: 5477-5482.

Ni, X., S. S. Quisenberry, J. Markwell, T. Heng-Moss, L. Higley, F. Baxendale, G. Sarath and R. Klucas (2001) *In vitro* enzymatic chlorophyll catabolism in wheat elicited by cereal aphid feeding. *Entomologia Experimentalis et Applicata* **101**: 159-166.

Oliver, K. M., J. A. Russell, N. A. Moran and M. S. Hunter (2003) Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proceedings of the National Academy of Sciences, USA* **100**: 1803-1807.

Plague, G. R., C. Dale and N. A. Moran (2003) Low and homogeneous copy number of plasmid-borne symbiont genes affecting host nutrition in *Buchnera aphidicola* of the aphid *Uroleucon ambrosiae*. *Molecular Ecology* **12**: 1095-1100.

Porter, D. R. and J. A. Webster (2000) Russian wheat aphid-induced protein alterations in spring wheat. *Euphytica* **111**: 199-203.

Puterka, G. J., W. C. Black, W. M. Steiner and R. L. Burtpn (1993) Genetic variation and phylogenetic relationships among worldwide collections of the Russian wheat aphid, *Diuraphis noxia* (Mordvilko), inferred from allozyme and RAPD-PCR markers. *Heredity* **70**: 604-618.

Puterka, G. J., J. D. Burd and R. L. Burton (1992) Biotypic variation in a worldwide collection of Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **85**: 1497-1506.

Reeson, A. F., T. Jankovic, M. L. Kasper, S. Rogers and A. D. Austin (2003) Application of *16S rDNA*-DGGE to examine the microbial ecology associated with a social wasp *Vespula germanica*. *Insect Molecular Biology* **12**: 85-91.

Romero, D. and R. Palacios (1997) Gene amplification and genome plasticity in prokaryotes. *Annual Review of Genetics* **31**: 91-111.

Russell, J. A., A. Latorre, B. Sabater-Munoz, A. Moya and N. A. Moran (2003) Side-stepping secondary symbionts: widespread horizontal transfer across and beyond the Aphidoidea. *Molecular Ecology* **12**: 1061-1075.

Russell, J. A. and N. A. Moran (2005) Horizontal transfer of bacterial symbionts: heritability and fitness effects in a novel aphid host. *Applied and Environmental Microbiology* **71**: 7987-7994.

Sandstrom, J. and N. A. Moran (1999) How nutritionally imbalanced is phloem sap for aphids? *Entomologia Experimentalis et Applicata* **91**: 203-210.

Sandstrom, J., A. Telang and N. A. Moran (2000) Nutritional enhancement of host plants by aphids-a comparison of three aphid species on grasses. *Journal of Insect Physiology* **46**: 33-40.

Sandstrom, J. P., J. A. Russell, J. P. White and N. A. Moran (2001) Independent origins and horizontal transfer of bacterial symbionts of aphids. *Molecular Ecology* **10**: 217-228.

Scarborough, C. L., J. Ferrari and H. C. J. Godfray (2005) Aphid protected from pathogen by endosymbiont. *Science* **310**: 1781.

Sheffield, V. C., D. R. Cox, L. S. Lerman and R. M. Myers (1989) Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proceedings of the National Academy of Sciences, USA* **86**: 232-236.

Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki and H. Ishikawa (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* **407**: 81-86.

Shufran, K. A., L. R. Kirkman and G. J. Puterka (2007) Absence of mitochondrial DNA sequence variation in Russian wheat aphid (Hemiptera: Aphididae) populations consistent with a single introduction into United States. *Journal of the Kansas Entomological Society* **80**: 319-326.

Silva, F. J., R. C. H. J. van Ham, B. Sabater and A. Latorre (1998) Structure and evolution of the leucine plasmids carried by the endosymbiont (*Buchnera aphidicola*) from aphids of the family Aphididae. *FEMS Microbiology Letters* **168**: 43-49.

Smith, C. M. (2009) Global phylogenetics of an invasive species: Evidence for multiple invasions into North America. Joint meeting of the Southwestern Branch of the Entomological Society of America and WERA066 (Western Extension/Education Research Activity), Stillwater, Oklahoma.

Srivastava, P. N. (1987) Nutritional physiology. Aphids. Their biology, natural enemies and control. P. Harrewijn. Amsterdam, Elsevier. **2A:** 99-121.

Stary, P. (2000) On-going expansion of Russian wheat aphid, *Diuraphis noxia* (Kurdj.) in central Europe (Hom.:Aphididae). *Journal of Pest Science* **73**: 75-78.

Stary, P. and H. Lukasova (2002) Increase of Russian wheat aphid, *Diuraphis noxia* (Kurdj.) in hot and dry weather (Hom., Aphididae). *Journal of Pest Science* **75**: 6-10.

Stoll, S., J. Gadau, R. Gross and H. Feldhaar (2007) Bacterial microbiota associated with ants of the genus *Tetraponera*. *Biological Journal of the Linnean Society* **90**: 399-412.

Surridge, A. K. J. (2007) Denaturing gradient gel electrophoresis characterisation of microbial communities in polycyclic aromatic hydrocarbon and polychlorinated biphenyl contaminated soil. Microbiology. Pretoria, University of Pretoria**:** 183.

Tamas, I., L. Klasson, B. Canback, A. K. Naslund, A.-S. Eriksson, J. J. Wernegreen, J. P. Sandstrom, N. A. Moran and S. G. E. Andersson (2002) 50 Million years of genomic stasis in endosymbiotic bacteria. *Science* **296**: 2376-2379.

Telang, A., J. Sandstrom, E. Dyreson and N. A. Moran (1999) Feeding damage by *Diuraphis noxia* results in a nutritionally enhanced phloem diet. *Entomologia Experimentalis et Applicata* **91**: 493-412.

Thao, M. L., L. Baumann, P. Baumann and N. A. Moran (1998) Endosymbionts (*Buchnera*) from the aphids *Schizaphis graminum* and *Diuraphis noxia* have different copy mumbers of the plasmid containing the leucine biosynthetic genes. *Current Microbiology* **36**: 238-240.

Tolmay, V. L., R. C. Lindeque and G. J. Prinsloo (2007) Preliminary evidence of a resistance-breaking biotype of the Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Homoptera: Aphididae), in South Africa. *African Entomology* **15**: 228-230.

Unterman, B. M., P. Baumann and D. L. McLean (1989) Pea aphid symbiont relationships established by analysis of *16S rRNA*s. *Journal of Bacteriology* **171**: 2970-2974.

Van der Westhuizen, A. J. and Z. Pretorius (1996) Protein composition of wheat apoplastic fluid and resistance to the Russian wheat aphid. *Australian Journal of Plant Physiology* **23**: 645-648.

Van Ham, R. C. H. J., F. Gonzalez-Candelas, F. J. Silva, B. Sabater, A. Moya and A. Latorre (2000) Postsymbiotic plasmid acquisition and evolution of the *repA1*-replicon in *Buchnera aphidicola*. *Proceedings of the National Academy of Sciences, USA* **97**: 10855-10860.

Van Ham, R. C. H. J., J. Kamerbeek, C. Palacios, C. Rausell, F. Abascal, U. Bastolla, J. M. Fernandez, L. Jimenez, M. Postigo, F. J. Silva, J. Tamames, E. Viguera, A. Latorre, A. Valencia, F. Moran and A. Moya (2003) Reductive genome evolution in *Buchnera aphidicola*. *Proceedings of the National Academy of Sciences, USA* **100**: 581-586.

Van Ham, R. C. H. J., A. Moya and A. Latorre (1997) Putative evolutionary origin of plasmids carrying the genes involved in leucine biosynthesis in *Buchnera aphidicola* (endosymbiont of aphids). *Journal of Bacteriology* **179**: 4768-4777.

Van Zyl, R. A. and A.-M. Botha (2008) Eliciting proteins from *Diuraphis noxia* biotypes differ in size and composition, *IPRI-18th Biennial Workshop*, Ft. Collins, USA, February 10-13. IPRI-18th Biennial Workshop, Ft. Collins, USA.

von Dohlen, C. D., C. A. Rowe and O. E. Heie (2006) A test of morphological hypotheses for tribal and subtribal relationships of Aphidinae (Insecta: Hemiptera: Aphididae) using DNA sequences. *Molecular Phylogenetics and Evolution* **38**: 316-329.

Weiland, A. A., F. B. Peairs, T. L. Randolph, J. B. Rudolph, S. D. Haley and G. J. Puterka (2008) Biotypic diversity in Colorado Russian wheat aphid (Hemiptera: Aphididae) populations. *Journal of Economic Entomology* **101**: 569-574.

Wernegreen, J. J. and N. A. Moran (2000) Decay of mutualistic potential in aphid endosymbionts through silencing of biosynthetic loci: *Buchnera* of *Diuraphis*. *Proceedings of the Royal Society of London* **267**: 1423-1431.

Wong, H. C. and S. Chang (1986) Identification of a positive retroregulator that stabilizes mRNAs in bacteria. *Proceedings of the National Academy of Sciences, USA* **83**: 3233-3237.

Wuchty, S., W. Fontana, I. L. Hofacker and P. Schuster (1999) Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers* **49**: 145-165.

Xayaphoummine, A., T. Bucher and H. Isambert (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Research* **33**: W605-W610.

Zientz, E., T. Dadndekar and R. Gross (2004) Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiology and Molecular Biology Reviews* **68**: 745-770.

CHAPTER 4

SYMBIOSIS: VARIATION IN *BUCHNERA APHIDICOLA*'S LEUCINE PLASMID MAY CONFER ADVANTAGE TO RUSSIAN WHEAT APHID BIOTYPE

Introduction

Aphids' success as pests can be attributed to a symbiotic alliance with *Buchnera aphidicola*, a bacterium that allows them to exploit dietary imbalanced phloem sap as food source (Srivastava 1987; Munson *et al.* 1991; Douglas 1998; Blackman & Eastop 2000). This ancient relationship between aphid and symbiont is crucial for the survival of both organisms, with the elimination of either leading to death of its partner (Munson *et al.* 1991; Lai *et al.* 1994; Douglas 1998; Sandstrom & Moran 1999; Baumann *et al.* 2006). Indeed, this relationship where the symbiont is located within host produced cells (bacteriocytes/ mycetocytes) are regarded as an advanced stage of symbiosis (Dixon 1998; Douglas 1998; Braendle *et al.* 2003). Here the bacterial symbiont is responsible for the synthesis and recycling of specific essential amino acids that is either present at low concentrations or absent in the dietary phloem (Douglas & Prosser 1992; Prosser & Douglas 1992; Febvay *et al.* 1994; Lai *et al.* 1994; Douglas 1998; Thao *et al.* 1998; Sandstrom & Moran 1999; Baumann *et al.* 2006).

Diuraphis noxia (Russian wheat aphid, RWA) is a major pest found in all but a few of the cereal producing countries (Blackman & Eastop 2000). The RWA only contains *B. aphidicola* as its endosymbiont (Swanevelder *et al.* 2010). However, in this relationship the contribution of the endosymbiont in the maintenance of certain essential amino acids is questionable when compared to other cereal feeding aphids. Lower gene copy numbers and the presence of pseudogenes (Lai *et al.* 1996; Wernegreen & Moran 2000; 2001), together with lower plasmid copy numbers (Baumann *et al.* 1995; Lai *et al.* 1996; Rouhbakhsh *et al.* 1996; Silva *et al.* 1998; Thao *et al.* 1998; Baumann *et al.* 1999; Soler *et al.* 2000) and higher non-synonymous substitutions rates in functional amino acid biosynthetic genes (Wernegreen & Moran 2000), all suggest a reduced contribution towards essential amino acid biosynthesis by the endosymbiont and/or a lower dependency of the aphid host on its symbiotic partner. This degradation in the mutualistic relationship is attributed to the ability of the RWA to induce higher levels of the selected essential amino acids in the phloem sap of susceptible host plants, thereby removing the selective pressure from the endosymbiont to retain as many functional gene copies of the required essential amino acid biosynthetic genes (Telang *et al.* 1999; Porter & Webster 2000; Sandstrom *et al.* 2000; Ni *et al.* 2001).

The recent appearance of new RWA biotypes in the USA and South Africa (Haley *et al.* 2004; Burd *et al.* 2006; Tolmay *et al.* 2007; Weiland *et al.* 2008) allowed for the investigation of aphid biotype variation and development in the field (Lapitan *et al.* 2007; Shufran *et al.* 2007). RWA biotypes are not anatomically or morphologically distinguishable from each other, but are discernable based on their ability to overcome specific host resistances in a plant differential

study (Puterka *et al.* 1992; Jyoti & Michaud 2005; Burd *et al.* 2006; Jyoti *et al.* 2006; Weiland *et al.* 2008). Furthermore, both symbiotic partners showed little sequence variation between different biotypes (Lapitan *et al.* 2007; Shufran *et al.* 2007; Swanevelder *et al.* 2010). However, in the aphid-endosymbiont relationship, a small change in a symbiont could have dire consequences for the aphid host, *e.g.* a single point mutation in *Buchnera* can determine aphid heat tolerance (Dunbar *et al.* 2007).

In a previous study on RWA *B. aphidicola*, the only sequence variation identified from the different *D. noxia* biotypes was a CCC-insert located between the *repA2* and *leuA* genes of the leucine plasmid (Swanevelder *et al.* 2010). The insert, though in a non-coding region, not only introduces a new predicted *rpoH3* (σ^{32}) transcription factor binding site (TFBs) to the leucine plasmid (Swanevelder *et al.* 2010), but is also located within an Aphididae conserved inverted repeat region. This may suggest some functional constraints for the genome (Silva *et al.* 1998). Furthermore, the newly predicted *rpoH* is one of only two sigma TFs predicted from sequenced *B. aphidicola* genomes, *i.e.* σ^{32} and σ^{70} (Shigenobu *et al.* 2000). To date only a single regulatory gene involved in essential amino acid regulation was identified from the many *Buchnera* genomes sequenced. Also, *B. aphidicola* only has one regulatory pathway that could regulate essential amino acid biosynthesis (Moran *et al.* 2005). It is thus plausible that transcriptional regulation via plasmid copy number is the most likely regulatory mechanism.

Variations in leucine plasmid copy number was found for *Buchnera* from different RWA biotypes (Moran *et al.* 2003; Swanevelder *et al.* 2010). The location of the CCC-insert, *i.e.* in the conserved inverted repeat and near predicted promoters, and the newly introduced TFBs, then suggest that this CCC-insert may play a role during transcription regulation of the leucine plasmid genes. Here we investigate the functionality of the CCC-insert different RWA biotypes. We also examine the sequence variation of the inverted repeat region located between the *repA2* and *leuA* of the leucine plasmid as an indicator of the potential regulatory utilization within other species of the family Aphididae.

Materials and methods

Aphids

The original South African *D. noxia* biotype (SA) was obtained from the ARC-Small Grains Institute, Bethlehem, South Africa, and maintained on a susceptible wheat cultivar Scheepers. A mutated form of the SA biotype (SAM) was maintained on a resistant wheat cultivar TugelaDN (Van Zyl & Botha 2008). Females of both South African biotypes were kept

in insect cages at 20 ± 2 °C with continuous artificial fluorescent lighting. The USA biotypes were obtained from Prof. N. Lapitan (Colorado State University, Fort Collins, USA) where they were maintained on a mixed diet of susceptible wheat and barley cultivars under greenhouse conditions (Lapitan *et al.* 2007). Samples of *Diuraphis mexicana* and *D. tritici*, collected in Colorado, were kindly provided by Dr. G.J. Puterka (USDA-ARS, Stillwater, Oklahoma, USA) and specimens of *Brevicoryne brassicae* (MF1435), *Hyalopterus pruni* (MF1422), *Macrosiphum rosae* (MF1408), *Myzus persicae* and *Uroleucon sonchi* (ACAM937) were a donation from Mr. I. Millar (South African National Collection of Insects, PPRI, Pretoria, South Africa).

DNA and RNA extraction

Aphid total DNA was extracted using the DNAzol extraction protocol (Molecular Research Centre, Cincinnati, USA) and cleaned with the DNeasy cleanup kit (Qiagen, USA) that included the on column RNase treatment (Qiagen). All samples were quantified using the Nanodrop ND-1000 Spectrophotometer (Thermo Scientific, RSA). RNA extractions were performed on a 100 individuals collected with a soft brush and immediately frozen with liquid nitrogen. All extractions were done in accordance to the manufactures' protocols using the RNeasy Mini Kit (Qiagen) which included the on column DNase I treatment (RNase-free DNase set, Qiagen), before spectrophotometric quantification.

Leader sequence determination

The leader sequences for *Buchnera* of the different RWA biotypes, with the insert and without the CCC-insert, were determined using the 5' RACE (rapid amplification of cDNA ends) system (Version2.0E, Invitrogen, USA). Two reverse primers were designed from the Genbank accession FJ705299, LeuA_RACE1_R (5'-CATTGCATCACCTGCTACCT-3') 244 bp into the *leuA* gene, and a nested primer (LeuA_RACE2_R, 5'-GTAATGCTTGTTCACCATC ACG-3') 36 bp from the *leuA* start codon. Cycling conditions for both PCRs consisted of an initial denaturing step at 94 °C for 2 min, followed by 30 cycles of 94 °C for 30 sec, 55 °C for 1 min with a 1 sec/cycle decrease in time, and 72 °C for 2 min; with a final extension step of 72 °C for 15 min. The nested PCR fragments were cleaned with sodium acetate/ethanol precipitation and cloned into pGEM-T Easy Vector (Promega, USA). Inserts were confirmed with colony PCR and the clones for each biotype with the longest fragments were sequenced as prescribed by the manufacturer using ABI BigDye v3.1. System (Applied Biosystems, USA) on an ABI 3100 automated sequencer (Applied Biosystems). The longest sequenced fragment that

correctly aligned with the leucine plasmid was regarded as the leader sequence and submitted to Genbank (GU145279 and GU145280).

The inverted repeat region in the Aphididae

The large inverted repeat region on the *Buchnera* leucine plasmid was isolated for the different members of the Aphididae using degenerate primers (Aphididae_repA2_For, 5'-GAAT TAACDAAAATWGGYCCKMARGG-3' and Aphididae_leuA_Rev, 5'-CCATTACTAG-TATCCTAATGCTTGRTCNCCATCNCG-3') designed from multiple sequence alignments of the region between the *repA2* and *leuA* using ClustalW (Larkin *et al.* 2007). Total DNA (30 ng), with 0.5 U ExSel High Fidelity DNA polymerase (Southern Cross Biotechnology, RSA), 1 × reaction buffer with $MgSO_4$ (Southern Cross Biotechnology), 100 µM of each dNTP and 0.4 µM of each primer, was used in 25 µL reaction volumes to amplify the inverted repeat region using a GeneAmp 9700 thermocycler (Applied Biosystems). Cycling conditions consisted out of an initial denaturing step at 94 °C for 2 min, followed by 12 touchdown cycles (15 sec at 94 °C, 30 sec at 60 °C, decreasing 1 °C /cycle, 30 sec at 72 °C), 25 standard cycles (10 sec at 94 °C, 10 sec at 50 °C, 30 sec at 72 °C) and a final 7 min step at 72 °C. The PCR products were cleaned through ethanol precipitation and cloned (pGEM-T Easy Vector, Promega) before unidirectional sequencing (ABI BigDye v3.1. System, Applied Biosystems). Sequence assemblies of the original chromatograms were produced using the default settings of ContigExpress (Vector NTI Advance 9, Invitrogen) (Lu & Moriyama 2004) and identities confirmed with a BLAST (Altschul *et al.* 1990; 1997) analysis against the non-redundant Genbank database (NCBI, http://www.ncbi.nlm.nih.gov/). Only unique sequences obtained for each species were submitted to Genbank (GU145281-GU145289).

Software analysis

Secondary structural analysis of the sequences were performed using the Quickfold on the DINAMelt server at 25 °C (Markham & Zuker 2005) and the Kinefold server (Xayaphoummine *et al.* 2005). Bacterial promoters with their sigma factor binding sites were determined using BPROM (http://softberry.com). Promoter candidates were investigated using neural network promoter predictions (http://www.fruitfly.org/seq_tools/promoter.html) (Reese & Eeckman 1995; Reese 2001) and Hidden Markov Models (http://bioinformatics.biol.rug.nl/websoftware/ppp/ppp_start.php). The Suite for Computational identification Of Promoter Elements (SCOPE, http://genie.dartmouth.edu/scope/) (Carlson *et al.* 2007; Chakravarty *et al.* 2007) and the pattern

discovery tools of the RSA webpage (http://www.bi.up.ac.za/rsa-tools/) (van Helden *et al.* 2000), were used to identify possible *cis* and other regulatory elements. WebLogo 3 (http://weblogo.threeplusone.com/) (Crooks *et al.* 2004) was used to convert consensus sequences to sequence logos. In all the analysis, *B. aphidicola*, if present in the option list, was always selected as part of an analysis, otherwise *E. coli* or prokaryote was used.

RT-qPCR

The iScript One-Step RT-PCR kit with SYBR Green (Qiagen) was used for the RT-qPCR reactions, with a final volume of 25 µL, 50 ng total aphid RNA, and PCR conditions in accordance to the manufacturer's recommendations. Primer and PCR optimizations were done in accordance to the manufacturers protocols (Bio-Rad, USA). All reactions were performed on an iCycler with an iQ Real-Time PCR Detection System (Bio-Rad). Single amplicons were confirmed with melting curve analysis and agarose gels. Primers for the RT-qPCRs were designed using Primer 3 (Rozen & Skaletsky 2000). Previously, Moran *et al.* (2005) used *rpsL* (ribosomal protein) as a RT-qPCR standard since it showed constant transcription. Based on this we've selected similar genes/subunits available on Genbank for *B. aphidicola* of the RWA, *i.e.* RNA polymerase β-subunit (*rpoB*, AF465521) (rpoB_F, 5'-TACAACGCACGCATTATTCC-3' and rpoB_R, 5'-ACGGTGACTGGAAGTTTTCG-3') and 16S ribosomal RNA (*16S rRNA*, M63251) (16SBuDn_F, 5'-TGTAGCGGTGAAATGCGTAG-3' and 16SBuDn_R, 5'-CC-TCCAAGTCGACATCGTTT-3').

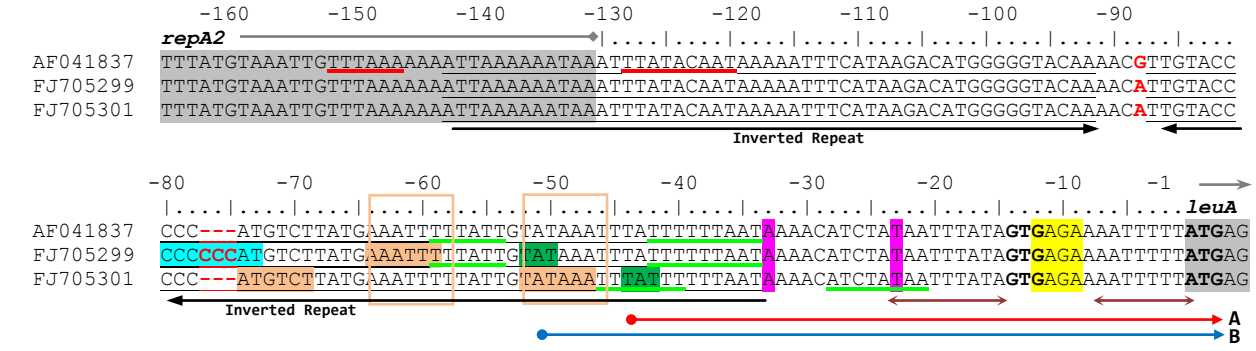
Previous results on the leucine plasmid (pLeu-Dn) indicated that the only difference between the RWA biotypes investigated here, resided within an inverted repeat region upstream of *leuA* (Swanevelder *et al.* 2010). Primers were therefore designed for *leuA* based on the Genebank accession FJ705299 (leuA_Buch_Dn_F, 5'-TGCATTTTCACATTCTTCTGG-3' and leuA_Buch_Dn_R, 5'-CTGCAGCTCTTCCTGATCGT-3'). The same study showed that *leuA* and *leuB* was in the same open reading frame (ORF), though *leuB* has its' own TATAAT box and predicted promoter (Swanevelder *et al.* 2010). Testing this hypothesis and possible secondary structures due to the leader sequence, we've included *leuB* (leuB_Buch_Dn_F, 5'-TGAATGTGCCATGATTACAGG-3' and leuB_Buch_Dn_R, 5'-CCCTGAATATCAGGAG-CTGAAC-3') in the analysis.

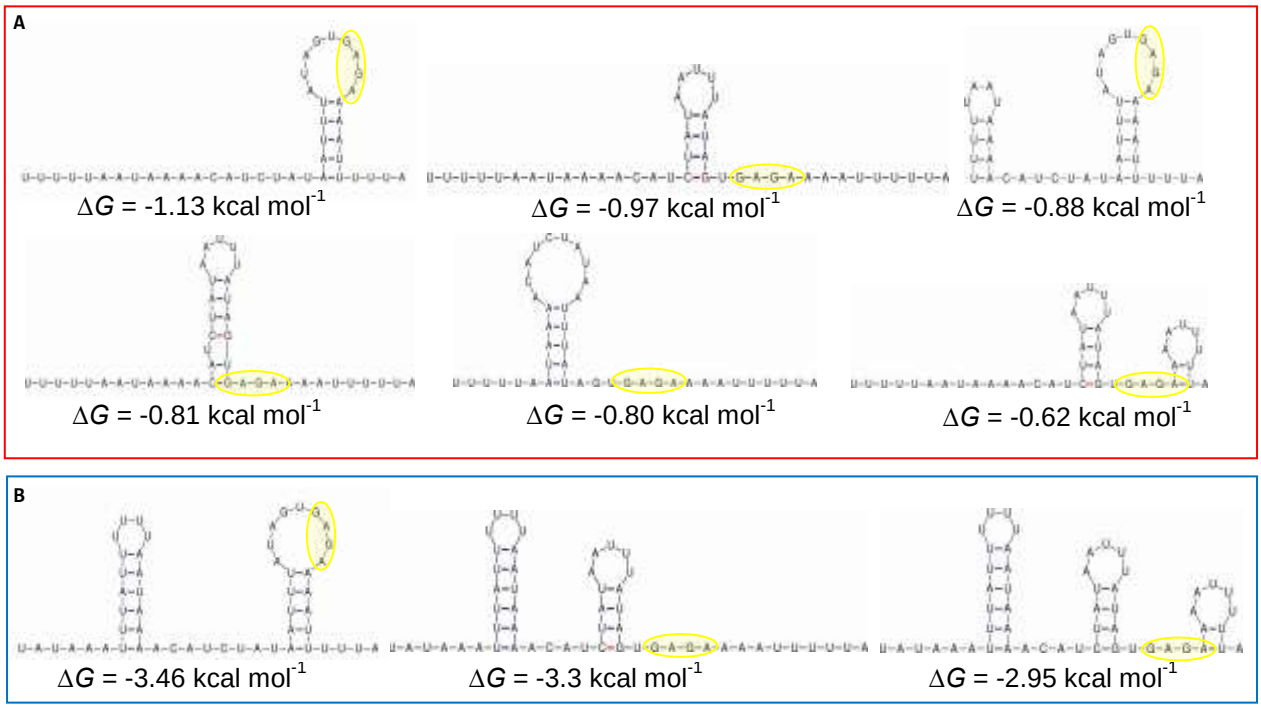
Results

The leader sequence

Several cloned RACE fragments for the *Buchnera* plasmids with and without the CCC-insert were sequenced and used to produce multiple alignments. The longest sequence in each case was submitted to Genbank (GU145279 and GU145280). The sequenced RACE fragments indicated that the mRNA 5' untranslated transcription region (5' UTR) starts on the second stem region of the inverted repeat, just upstream of the *leuA* gene and not as previously suggested to be located before the stemloop structure (Swanevelder *et al.* 2010). It was also found that the leader sequences differ in length (*Figure 4.1*), with longer sequence for the *B. aphidicola* plasmid with the CCC-insert. This finding implies a change to the previously predicted transcription start site. In the new prediction the transcription start site moves upstream, making it part of the conserved stem towards the loop and nearer to the CCC-insert (*Figure 4.1*). This further implies that the predicted *rpoH* (σ^{32}) binding site may be involved in the regulation (Swanevelder *et al.* 2010). However, testing the different predicted promoters (BPROM, Softberry) in a reporter plasmid (pGlow-TOPO Reporter Kit, Invitrogen) within *E. coli* rendered no expression data (data not shown), leading to the conclusion that the high AT-rich *B. aphidicola* predicted promoters tested were either unrecognisable in *E. coli* or alternatively there were too many "recognisable" *E. coli* promoters that interfered with the expression.

The differences in the lengths of the leader sequences produced another possible reason for the observed variance in the transcript expression levels. Secondary structures are known to stabilise the mRNA transcripts of bacteria, thereby increasing their half-life (Emroy *et al.* 1992). We therefore included the predictions of both leader sequences as part of the analysis. The leader sequences produced different 5' secondary structures (Quikfold, DINAMelt Server) (*Figure 4.1*), with the longer 5'UTR producing structures (*Figure 4.1*, insert *B*) that have more than double the free energy values than those from the short fragment (*Figure 4.1*, insert *A*).





Inverted repeat
leuA start
repA2 end
RBS Ribosomal Binding Site
TEXT Species variation
TAT Leader region start site
5' UTR CCC-insert
5' UTR no insert

TEXT Predicted *rpoH3* (δ^{32}) binding site
TEXT BPROM -10 & -35 box predictions for the whole region, *repA2-leuA*
TEXT BPROM predicted -10 & -35 boxes from the loop region to *leuA*
TEXT Neural Network promoter predictions, transcription start sites
TEXT Proposed TATAAT or -35 boxes based on the 5' UTR sequences
ATG/GTG Possible translational start sites
AU-rich regions preceding the coding sequence (CDS)

Figure 4.1 *Buchnera aphidicola* of *D. noxia* with the leader regions (5' UTR) indicated by a blue line for accessions with an upstream CCC-insert and a red line for those without the CCC-insert. Predicted promoters, sigma factor binding sites, AU-rich region and ribosomal binding sites (RBS), including possible start sites, are indicated. Predicted secondary structures (Quickfold, DINAMelt Server, 25 °C) for the leader sequences are given in the inserts A and B. Six structures were predicted for the short 5' UTR (insert A) and three for the longer leader region of the CCC-insert containing *Buchnera* (insert B). The longer leader sequence produced predicted structures that are approximately -2.3 kcal mol^{-1} more stable than those of the shorter 5' UTR. These structures could increase the half life of the RNA molecules by preventing RNA degradation.

The inverted repeat in the Aphididae

The preceding results indicate that the second stem of the inverted repeat region, just upstream of the *leuA* gene, could be part of a functional area within the RWA biotypes investigated. We know from previous work that the inverted region in which the insert is located, is conserved within the family Aphididae (Silva *et al.* 1998). This suggests that the inverted repeat region, especially the second stem, plays a regulatory role or is preserving a regulatory region. We have attempted to gather support for this hypothesis by sequencing the inverted repeat region of a number of aphid individuals, from one or more localities, in order to obtain data that can indicate the extent to which this region is utilized within species/genera as a regulatory mechanism and possibly help indicating the underlining mechanism.

The region, though conserved in most samples, did show some variation within *Buchnera* from the same host species (*Figure 4.2*, in bold red), *i.e.* for *Buchnera* from aphids originating from the same population/sampling site, hosts *B. brassicae* (GU145288-9) and *H. pruni* (GU145286-7) and from different regions, hosts *D. noxia* (AF041837, FJ705299 and FJ705301), *M. rosae* (GU145283 and AJ006881) and *U. sonchi* (GU145285 and AJ006873). Changes observed in the conserved stem regions usually produce imperfect hairpins, *e.g. B. aphidicola* from *B. brassicae*, *D. noxia* and *U. sonchi*, resulting in an increase in the Gibbs free energy (ΔG) of app. 5 kcal mol^{-1} and less stable structures. The only exception was two point mutations in the stem regions of *B. aphidicola* of *H. pruni* that had little effect on the predicted structure or stability of the hairpins (GU145286 vs. GU145287, ΔG difference of 0.57 kcal mol^{-1}). These two point mutations are predicted to interact as T-G base pairings in the stem region (nearest-neighbor effect), thereby preventing major structural changes. Within-hosts differences were also obtained for the variable loop regions of *B. aphidicola* of *D. noxia* and *M. rosae*. A single nucleotide variation in the loop region of the endosymbiont of *D. noxia* (AF041837, FJ705301) alters the predicted stability of the hairpin by -0.57 kcal mol^{-1}, while the stem and loop region lengths are maintained. Changes in loop length and composition, together with varying stem lengths, can have major implications on the predicted structural stability – though a perfect hairpin is maintained. This is illustrated in the two *B. aphidicola* accessions of *M. rosae* where the hairpins' ΔGs differ by 4.23 kcal mol^{-1}, as a result of the changes.

A comparison of this region within a *Buchnera* from the same aphid genus (*i.e. Diuraphis* and *Rhopalosiphum*) showed that the variations between the species seemed to be mainly focused to variable regions (*Figure 4.2*: bold blue text), *i.e.* the loop region and the variable region between the stemloop and the *leuA* start site, and to the edges of the stem regions, *i.e.* a core region is conserved within the stems. Indeed, the conserved cores located in the stem region of the hairpin are preserved across all three tribes of the Aphididae (*Figure 4.2*) just as

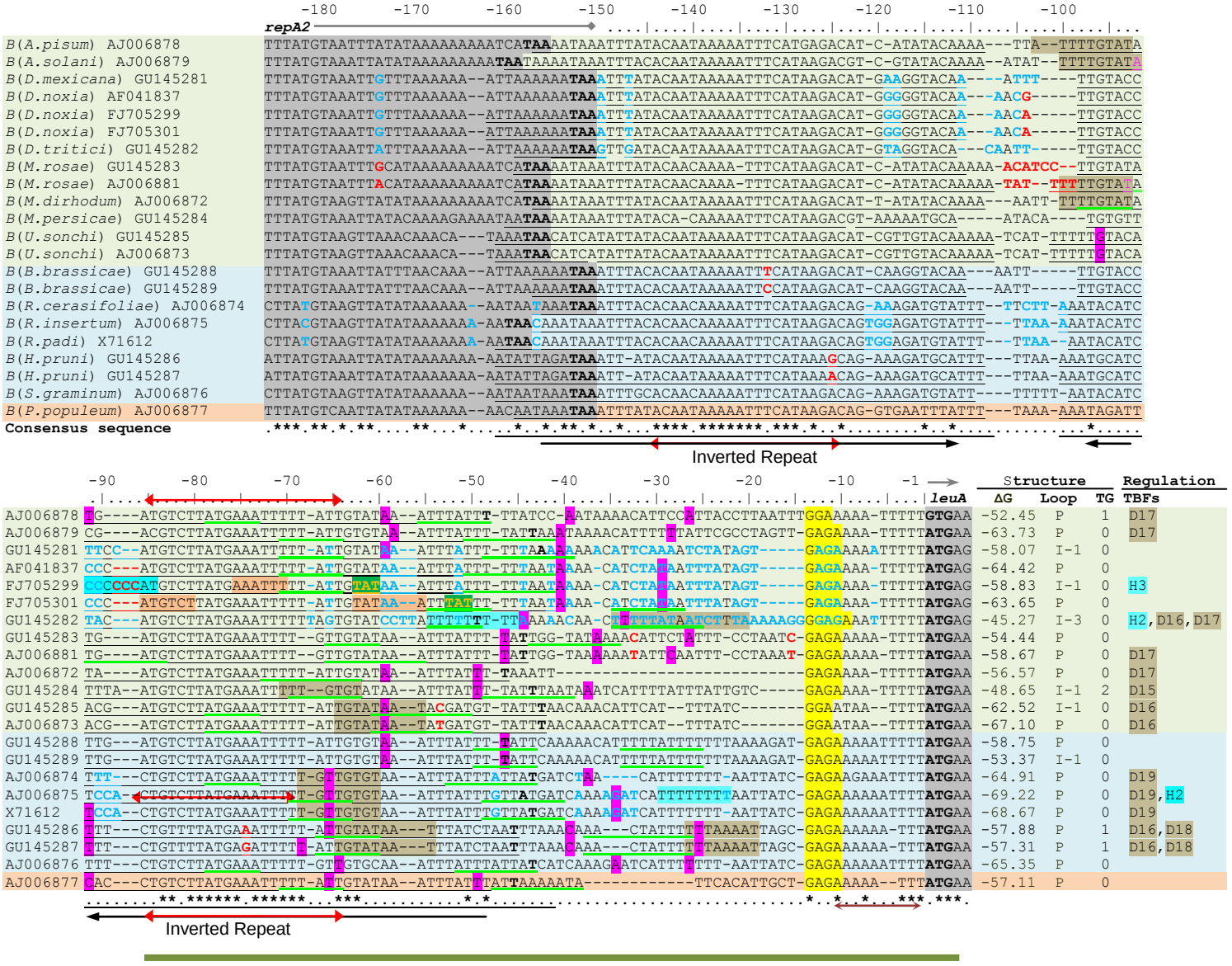

repA2
B(A.pisum) AJ006878
B(A.solani) AJ006879
B(D.mexicana) GU145281
B(D.noxia) AF041837
B(D.noxia) FJ705299
B(D.noxia) FJ705301
B(D.tritici) GU145282
B(M.rosae) GU145283
B(M.rosae) AJ006881
B(M.dirhodum) AJ006872
B(M.persicae) GU145284
B(U.sonchi) GU145285
B(U.sonchi) AJ006873
B(B.brassicae) GU145288
B(B.brassicae) GU145289
B(R.cerasifoliae) AJ006874
B(R.insertum) AJ006875
B(R.padi) X71612
B(H.pruni) GU145286
B(H.pruni) GU145287
B(S.graminum) AJ006876
B(P.populeum) AJ006877
Consensus sequence
Inverted Repeat
Structure
Regulation
ΔG
Loop
TG
TBFs
leuA
AJ006878 -52.45 P 1 D17
AJ006879 -63.73 P 0 D17
GU145281 -58.07 I-1 0
AF041837 -64.42 P 0
FJ705299 -58.83 I-1 0 H3
FJ705301 -63.65 P 0
GU145282 -45.27 I-3 0 H2,D16,D17
GU145283 -54.44 P 0
AJ006881 -58.67 P 0 D17
AJ006872 -56.57 P 0 D17
GU145284 -48.65 I-1 2 D15
GU145285 -62.52 I-1 0 D16
AJ006873 -67.10 P 0 D16
GU145288 -58.75 P 0
GU145289 -53.37 I-1 0
AJ006874 -64.91 P 0 D19
AJ006875 -69.22 P 0 D19,H2
X71612 -68.67 P 0 D19
GU145286 -57.88 P 1 D16,D18
GU145287 -57.31 P 1 D16,D18
AJ006876 -65.35 P 0
AJ006877 -57.11 P 0
Inverted Repeat

Figure 4.2 cont.

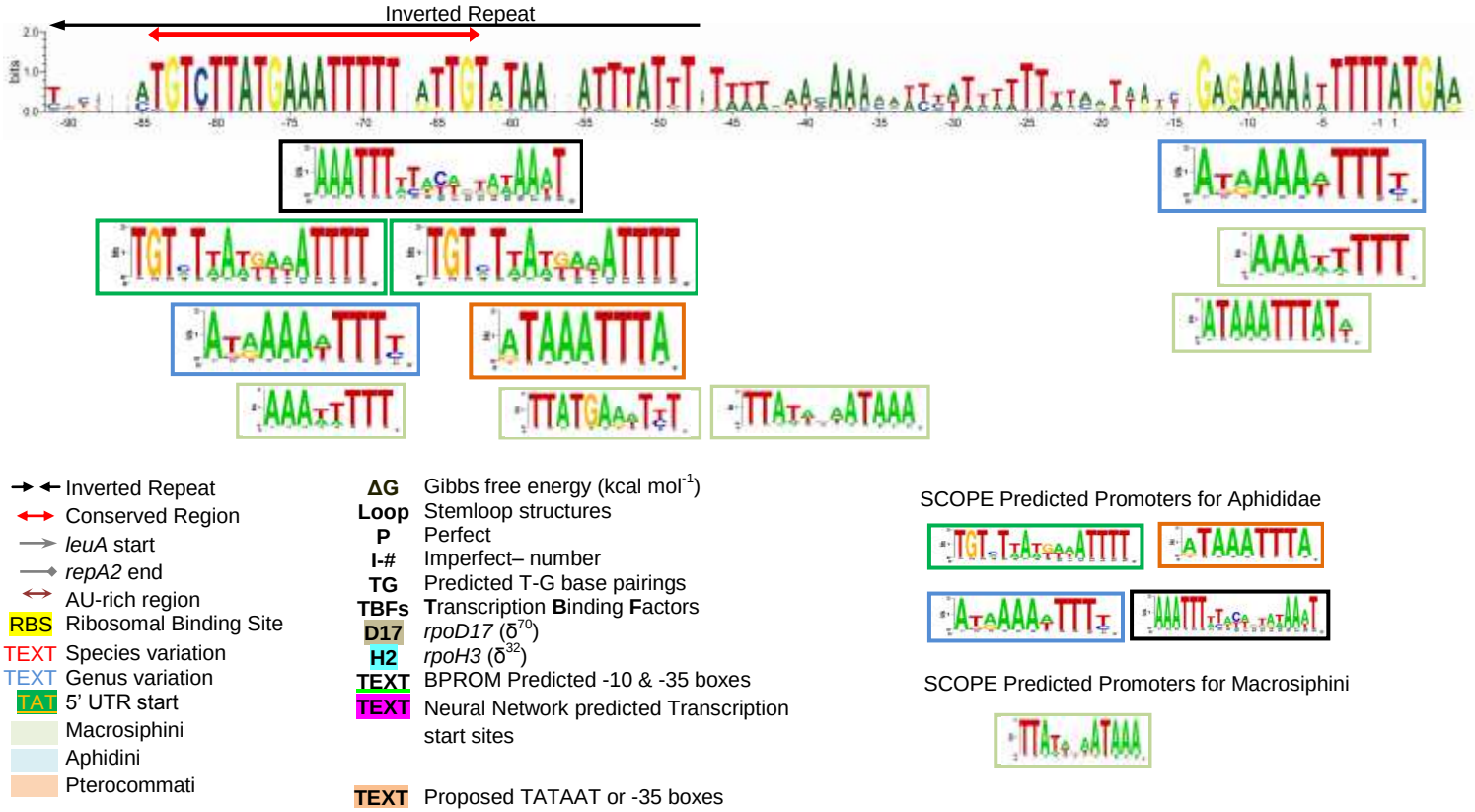


Figure 4.2 The region between the *repA2* and *leuA* on the leucine plasmid of *B. aphidicola* accessions originating from various aphid hosts. Variations within a species is indicated by bold red lettering, variations within a genus by bold blue lettering and the start site of the 5' UTR leader sequence by bold gold lettering. The genes *repA2* and *leuA* are shaded in grey with their end and start codons, respectively, in bold. The inverted repeat regions of the family Aphididae are underlined on the consensus sequence and the regions indicated by black arrows. The conserved core region (red double arrows) of the inverted repeat region (black) includes predicted promoters, transcription factor binding sites and start codons (see legend). The Gibbs' free energy (ΔG) values for the stemloop structures were calculated with Quickfold (DINAMelt Server, at 25 $^{\circ}$C in kcal mol^{-1}). Predicted stemloop structures (Loop) is given as perfect (P) or imperfect (IP), with the number of additional loops indicated. The number of suboptimal T-G pairs (TG) predicted is also listed. RSA-tools' Consensus and Convert-matrix programs were used to obtain the core regions of the stemloop structures that are conserved within the family.

previously shown (Silva *et al.* 1998). This suggests that changes to these variations within the variable loop region and the edges of the inverted repeats could destabilize the stemloop structure in a similar way to the CCC-insert and could theoretically regulate gene expression.

RT-qPCR

Relative gene expression levels of two genes on the *pleuABCD* plasmid of *B. aphidicola*, *leuA* and *leuB*, were quantified to assess whether the presence of the CCC-insert had any effect on the expression of these genes (*Figure 4.3*). A difference found in the expression levels has the potential to contribute to RWA adaptation to new hosts via the endosymbiont. It is known that *B. aphidicola* of the South African RWA biotypes have lower plasmid copy numbers (0.35 copies/bacterial chromosome) than their US counterparts (1.04 and 0.88 copies/bacterial chromosome for RWA-US1 and RWA-US2 respectively) (Chapter 3: *Figure* 3.4, Swanevelder *et al.* 2010). Significantly higher expression levels were obtained for *leuA* and *leuB* after normalization with *rpoB* (*Figure 4.3*) and *16S rRNA* (not shown) in the South African *B. aphidicola* accessions with the CCC-insert than *B. aphidicola* without the insert (*i.e.* RWA-US1 and RWA-US2).

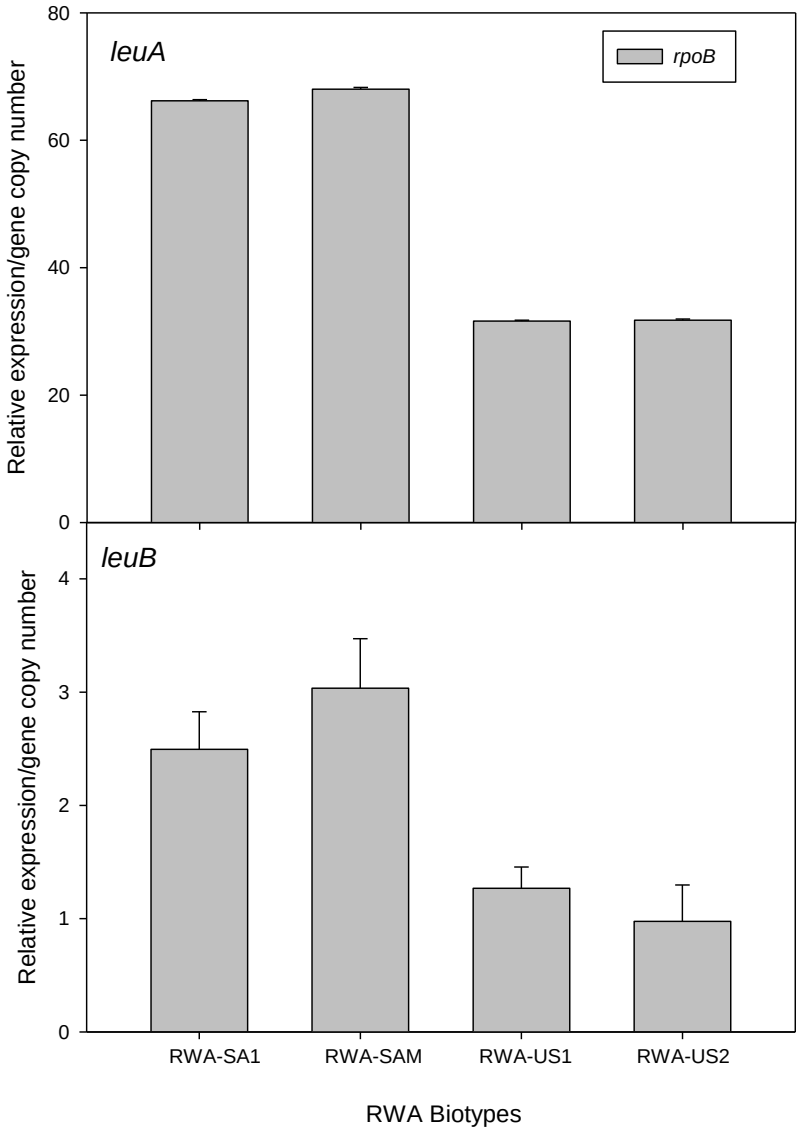


Figure 4.3 The relative gene expression levels per plasmid copy of *leuA* and *leuB* after normalization with *rpoB.*

Discussion

The role of the endosymbiont in an aphid's adaptability to new environments has been rarely investigated, though the two organisms are totally inter-dependent on each other for their survival (Munson *et al.* 1991; Lai *et al.* 1994; Douglas 1998; Sandstrom & Moran 1999; Baumann *et al.* 2006). Recently Dunbar *et al.* (2007) clearly illustrated the influence that the endosymbiont may have on its host by showing that a single point mutation in the bacterium's genome determined the aphid host's tolerance to environmental temperature changes. In our previous study, we identified a single mutation (CCC-insert) that differs between RWA biotypes which was located in a proposed regulatory region situated on the leucine plasmid upstream from the *leuA*-*leuB* ORF (Silva *et al.* 1998; Chapter 3; Swanevelder *et al.* 2010). RWA biotypes and their endosymbionts are known to have little sequence variation (Lapitan *et al.* 2007; Shufran *et al.* 2007); thereby suggesting that this CCC-insert may play a functional role in the development of at least some RWA biotypes.

Here we showed that this insert does result in higher transcript copy numbers (*Figure 4.3).* From our predicted models for the sequenced 5'UTR of *leu*A (Chapter 3), it was shown that the transcriptional start of the region followed the loop in the second inverted repeat upstream from the *leuA* gene on the *leuABCD* plasmid. The sequenced fragments also illustrated that the CCC-insert caused and increased 5'UTR length. This suggests a change in the transcription start sites as a result of the insert, forming a new predicted *rpoH* (σ^{32}) binding site. Only two transcription factors in *B. aphidicola* are known, *i.e. rpoH* that encodes the sigma factor σ^{32} and *rpoD* that encodes σ^{70} (Shigenobu *et al.* 2000).

The *rpoH* (σ^{32}) has a functional significance in *E. coli* as a heat shock transcription factor (TF) (Erickson *et al.* 1987). In order to understand how a heat shock/stress response TF could regulate gene expression under normal conditions, we investigated the *Buchnera* genome and the implications it may have on protein function/stability. *Buchnera* genomes are drastically reduced, have little or no recombination; occur as multiple chromosomal copies per cell; are AT-rich with no codon bias; accrue detrimental mutations and have elevated evolutionary rates with increased nucleotide substitution rates (Moran 1996; Clark *et al.* 1999; Komaki & Ishikawa 1999; Itoh *et al.* 2002; Moya *et al.* 2002). These genomic conditions have resulted in amino acids sequences that produce structurally destabilized or miss-shaped proteins (Moran 1996; Van Ham *et al.* 2003; Wilcox *et al.* 2003). The destabilized/deformed proteins initiate the release of the σ^{32} factor that binds and redirect RNA polymerase to σ^{32} promoters, thereby regulating the expression of genes that

encodes chaperones, proteins associated with cellular homeostasis restoration and proteases (Wilcox *et al.* 2003). It is suggested that the endosymbiont attempts to retain protein functionality by stabilising the protein structures using "heat shock" chaperonins. However, unlike a temporarily protein destabilizing environment, it seems that the genome of *Buchnera* continuously produce miss-formed proteins, thereby necessitating a steady level of protein chaperonin production under normal conditions. Indeed, the chaperonin GroEL constitutes about 10 % of the cellular proteins under normal growth conditions in the *Buchnera* of *Schizaphis graminum* (Baumann *et al.* 1996). The role of this chaperonin in protein stabilization is not only underlined by the positive selection pressure under which the protein is maintained (Fares *et al.* 2002), but also by experimental evidence that showed that with its (GroEL) co-expression *Buchnera* enzyme function is enhanced (Huang *et al.* 2008). *Buchnera*, even though it has lost the ability to regulate most of the heat stress proteins under heat stress conditions (Baumann *et al.* 1996; Sato & Ishikawa 1997a; b; Wilcox *et al.* 2003), still retains a σ^{32} TF that has the necessary binding domains needed for functionality (Wilcox *et al.* 2003). A low, but continuous expression of *rpoH* under normal growth conditions is therefore necessary to maintain the steady production of the various required protein chaperonins to ensure protein stability. Indeed, *rpoH* is only slightly up-regulated from its norm under heat shock conditions (Wilcox *et al.* 2003).

In order to understand how this new putatively formed TF could regulate gene expression, we investigated the expression of *leuA* and *leuB* in RWA biotypes, with and without the CCC-insert. We found a twofold difference in the expression of *leuA* and *leuB* as expressed per copy number (*Figure 4.3*).

The newly formed *rpoH* (σ^{32}) binding site may explain the higher levels of transcripts observed in the biotypes with the introduced TFBs. However, the length differences in the leader sequences produced another possibility for the higher transcript levels. Secondary structures are known to play a role in stabilising mRNA transcripts in bacteria (Emroy *et al.* 1992). Here the predicted 5' secondary structures of the longer 5'UTR (*Figure 4.1* insert *B*) are all more than twice as stable as those produced by the shorter leader sequence (*Figure 4.1* insert *A*). This suggests that secondary structures in the leader sequence may provide more stability to *Buchnera* transcripts originating from plasmids with the CCC-insert, thereby increasing its half life, and thus the total transcript levels. Secondary structures in the translation initiation region is also known to play a role in translation regulation, even within *B. aphidicola* (Tchufistova *et al.* 2003). However, the conserved translational control that features in the leader sequence necessary for S1 protein

translational control (Tchufistova *et al.* 2003) is absent from the leader sequences identified here (*Figure 4.1*).

The functional mutation in the inverted repeat region of *B. aphidicola,* together with the conservation of the structure within the Aphididae (Silva *et al.* 1998), supports a wider functional role within the family. The observed changes never occurred within the conserved "core" region of the inverted repeats, but were kept to variable regions within the stemloop structure. This low level of variation observed could be due to the small species sample size used here. However, changes observed usually did alter the structures of predicted hairpins, thereby increasing/decreasing its stability, *e.g.* the *B. aphidicola* of *M. rosae* had hairpin ΔGs that differ with 4.23 kcal mol^{-1} even though a perfect stemloop was maintained. Within a genus, changes were also never within the conserved core. If the 5' UTRs obtained for the *B. aphidicola* of *D. noxia* biotypes are the norm for the family, the highly conserved core of the second inverted repeat is a likely promoter region. This would suggest that the stemloop structure is used to protect the promoter region, *i.e.* the core regions, while mutations that affect its structure and stability is used to control gene expression or the half-life of the transcript. Therefore, it can be argued that *Buchnera* utilizes structural changes in this region to control gene expression via structural stability, *i.e.* easier access to start sites, or via structural changes in leader sequences that may increase transcript half-lives.

Conclusion

The initial transfer of the leucine biosynthetic genes, from the chromosome to a plasmid, was probably one of the main reasons for a successful aphid-*Buchnera* symbiosis (Latorre *et al.* 2005). However, different aphids require different levels of this essential amino acid. Indeed, the RWA has the ability to increase certain essential amino acids, including leucine, in susceptible hosts' phloem, but cannot achieve the same in a resistant host (Telang *et al.* 1999; Porter & Webster 2000; Sandstrom *et al.* 2000; Ni *et al.* 2001). Previously it was believed that the fine tuning of leucine production in an aphid species was done through changes in the plasmid copy number (Thao *et al.* 1998; Plague *et al.* 2003; Latorre *et al.* 2005). However, the increase in the *leuA-leuB* ORF transcripts relative to the known plasmid copy numbers, suggest that this could be a form of regulation within the species. The existence of variable regions within an aphid species and differences in structural stabilities of either the plasmid or leader sequences within *Buchnera* plasmid could arguably support a regulatory mechanism for leucine control. The fact that the same

insertion occurred twice, independently and involved multiple nucleotides that are not part of the *Buchnera*'s genome bias (AT-rich genome) (Chapter 3; Swanevelder *et al.* 2010), further supports this hypothesis. We therefore propose that the variation within the inverted repeat region, together with plasmid copy numbers, are used by *Buchnera* to control gene expression, either through higher expression levels or via 5' UTR mRNA stabilization.

We showed that copy number is not necessary the same as expression level in *Buchnera*. This suggests that the *Buchnera* endosymbiont is employed by the RWA to compensate for lower leucine levels when it is feeding on resistant cultivars. Since RWA biotypes are characterized based on their ability to overcome different hosts' resistances rather than aphid anatomy and morphology, an endosymbiont mutation that allows feeding on previously resistant cultivars, could result in a classifiable "new RWA biotype".

Acknowledgements

We would like to thank Prof N.L.V. Lapitan (Colorado State University) for the USA RWA biotypes, Dr. V. Tolmay (Small Grains Institute) for the SA RWA biotype, Dr. G.J. Puterka (USDA-ARS, Stillwater, Oklahoma, USA) for the other *Diuraphis* species and Mr. I. Millar (South African National Collection of Insects (SANC), PPRI) for the other members of the Aphididae. This research was funded by THRIP (Technology and Human Resources for Industry Programme, South Africa) and the Winter Cereal Trust.

References

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389-3402.

Baumann, L., P. Baumann, N. A. Moran, J. Sandstrom and M. L. Thao (1999) Genetic characterization of plasmids containing genes encoding enzymes of leucine biosynthesis in endosymbionts (*Buchnera*) of aphids. *Journal of Molecular Evolution* **48**: 77-85.

Baumann, P., L. Baumann and M. A. Clark (1996) Levels of *Buchnera aphidicola* chaperonin GroEL during growth of the aphid *Schizaphis graminum*. *Current Microbiology* **32**: 279-285.

Baumann, P., L. C.-Y., L. Baumann, D. Rouhbakhsh, N. A. Moran and M. A. Clark (1995) Mutualistic associations of aphids and prokaryotes: biology of the genus *Buchnera*. *Applied and Environmental Microbiology* **61**: 1-7.

Baumann, P., N. A. Moran and L. Baumann (2006) Bacteriocyte-associated endosymbionts of insects. The prokaryotes.

Blackman, R. L. and V. F. Eastop (2000) Aphids on the world's crops. An identification and information guide. London, Wiley.

Braendle, C., T. Miura, R. Bickel, A. W. Shingleton, S. Kambhampati and D. L. Stern (2003) Developmental origin and evolution of Bacteriocytes in the aphid–*Buchnera* symbiosis. *PLoS Biology* **1**: 70-76.

Burd, J. D., D. R. Porter, G. J. Puterka, S. D. Haley and F. B. Peairs (2006) Biotypic variation among North American Russian wheat aphid (Homoptera: Aphididae) populations. *Journal of Economic Entomology* **99**: 1862-1866.

Carlson, J. M., A. Chakravarty, C. E. DeZiel and R. H. Gross (2007) SCOPE: a web server for practical *de novo* motif discovery. *Nucleic Acids Reseach* **35**: w259-w264.

Chakravarty, A., J. M. Carlson, R. S. Khetani and R. H. Gross (2007) A novel ensemble learning method for *de novo* computational indentification of DNA binding sites. *BMC Bioinformatics* **8**: 249.

Clark, M. A., N. A. Moran and P. Baumann (1999) Sequence evolution in bacterial endosymbionts having extreme base compositions. *Molecular Biology and Evolution* **16**: 1586-1598.

Crooks, G. E., G. Hon, J. M. Chandonia and S. E. Brenner (2004) WebLogo: a sequence logo generator. *Genome Research* **14**: 1188-1190.

Dixon, A. F. G. (1998) Aphid ecology: An optimization approach. London, Chapman & Hall.

Douglas, A. E. (1998) Nutritional interactions in insect-microbial symbioses: Aphid and their symbiotic bacteria Buchnera. *Annual Review of Entomology* **43**: 17-37.

Douglas, A. E. and W. A. Prosser (1992) Synthesis of the essential amino acid tryptophan in the pea aphid (*Acyrthosiphon pisum*) symbiosis. *Journal of Insect Physiology* **38**: 565-568.

Dunbar, H., A. C. C. Wilson, N. R. Ferguson and N. A. Moran (2007) Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLOS Biology* **5**: e96 [1006-1015].

Emroy, S. A., P. Bouvet and J. G. Belasco (1992) A 5'-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes & Development* **6**: 135-148.

Erickson, J. W., V. Vaughn, W. A. Walter, F. C. Neidhardt and C. A. Gross (1987) Regulation of the promoters and transcripts of *rpoH*, the *Escherichia coli* heat shock regulatory gene. *Genes & Development* **1**: 419-432

Fares, M. A., E. Barrio, B. Sabater-Munoz and A. Moya (2002) The evolution of the heat-shock protein GroEL from *Buchnera*, the primary endosymbiont of aphids, is governed by positive selection. *Molecular Biology and Evolution* **17**: 1162-1170.

Febvay, G., I. Liadouze, J. Guillaud and G. Bonnot (1994) Analysis of energetic amino acid metabolism in *Acyrthosiphon pisum*: A multidimensional approach to amino acid metabolism in aphids. *Archives of Insect Biochemistry and Physiology* **29**: 45-69.

Haley, S. D., F. B. Peairs, C. B. Walker, J. B. Rudolph and T. L. Randolph (2004) Occurrence of a new Russian wheat aphid biotype in Colorado. *Crop Science* **44**: 1589-1592.

Huang, C.-Y., C.-Y. Lee, H.-C. Wu, M.-H. Kuo and C.-Y. Lai (2008) Interactions of chaperonin with a weakly active anthranilate synthase from the aphid endosymbiont *Buchnera aphidicola*. *Microbial Ecology* **56**: 696-703.

Itoh, T., W. Martin and M. Nei (2002) Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. *PNAS* **99**: 12944-12948.

Jyoti, J. L. and J. P. Michaud (2005) Comparative biology of a novel strain of Russian wheat aphid (Homoptera: Aphididae) on three wheat cultivars. *Journal of Economic Entomology* **98**: 1032-1039.

Jyoti, J. L., J. A. Qureshi, J. P. Michaud and T. J. Martin (2006) Virulence of two Russian wheat aphid biotypes to eight wheat cultivars at two temperatures. *Crop Science* **46**: 774-780.

Komaki, K. and H. Ishikawa (1999) Intracellular bacterial symbionts of aphids possess many genomic copies per bacterium. *Journal of Molecular Evolution* **48**: 717-722.

Lai, C.-Y., L. Baumann and P. Baumann (1994) Amplification of *trpEG*: Adaptation of *Buchnera aphidicola* to an endosymbiotic association with aphids. *Proceedings of the National Academy of Sciences, USA* **91**: 3819-3823.

Lai, C.-Y., P. Baumann and N. A. Moran (1996) The endosymbiont (*Buchnera* sp.) of the aphid *Diuraphis noxia* contains plasmids consisting of *trpEG* and tandem repeats of *trpEG* pseudogenes. *Applied and Environmental Microbiology* **62**: 332-339.

Lapitan, N. L. V., Y.-C. Li, R. S. G. Walters, Y. Peng, F. B. Peairs and A.-M. Botha (2007) Limited nuclear and mitochondrial DNA variation among Russian wheat aphid (*Diuraphis noxia*) biotypes from the United States and Africa, *American Entomological Society*, San Diego, December 9-12. American Entomological Society, San Diego.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

Latorre, A., R. Gil, F. J. Silva and A. Moya (2005) Chromosomal stasis versus plasmid plasticity in aphid endosymbiont *Buchnera aphidicola*. *Heredity* **95**: 339-347.

Lu, G. and E. N. Moriyama (2004) Software review: VectorNTI, a balanced all-in-one sequence analysis suite. *Briefings in Bioinformatics* **5**: 378-388.

Markham, N. R. and M. Zuker (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Research* **33**: 577-581.

Moran, N. A. (1996) Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proceedings of the National Academy of Sciences, USA* **93**: 2873-2878.

Moran, N. A., H. E. Dunbar and J. L. Wilcox (2005) Regulation of transcription in a reduced bacterial genome: nutrient-provisioning genes of the obligate symbiont *Buchnera aphidicola*. *Journal of Bacteriology* **187**: 4229-4237.

Moran, N. A., G. R. Plague, J. P. Sandstrom and J. L. Wilcox (2003) A genomic perspective on nutrient provisioning by bacterial symbionts of insects. *PNAS* **100**: 14543-14548.

Moya, A., A. Latorre, B. Sabater-Muñoz and F. J. Silva (2002) Comparative molecular evolution of primary (*Buchnera*) and secondary symbionts of aphids based on two protein-coding genes. *Journal of Molecular Evolution* **55**: 127-137.

Munson, M. A., P. Baumann, A. M. Clark, L. Baumann, N. A. Moran, D. J. Voegtlin and B. C. Campbell (1991) Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. *Journal of Bacteriology* **173**: 6321-6324.

Ni, X., S. S. Quisenberry, J. Markwell, T. Heng-Moss, L. Higley, F. Baxendale, G. Sarath and R. Klucas (2001) *In vitro* enzymatic chlorophyll catabolism in wheat elicited by cereal aphid feeding. *Entomologia Experimentalis et Applicata* **101**: 159-166.

Plague, G. R., C. Dale and N. A. Moran (2003) Low and homogeneous copy number of plasmid-borne symbiont genes affecting host nutrition in *Buchnera aphidicola* of the aphid *Uroleucon ambrosiae*. *Molecular Ecology* **12**: 1095-1100.

Porter, D. R. and J. A. Webster (2000) Russian wheat aphid-induced protein alterations in spring wheat. *Euphytica* **111**: 199-203.

Prosser, W. A. and A. E. Douglas (1992) A test of the hypothesis that nitrogen is upgraded and recycled in an aphid (*Acyrthosiphon pisum*) symbiosis. *Journal of Insect Physiology* **38**: 93-99.

Puterka, G. J., J. D. Burd and R. L. Burton (1992) Biotypic variation in a worldwide collection of Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **85**: 1497-1506.

Reese, M. G. (2001) Application of a time-dalay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Computational Chemistry* **26**: 51-6.s.

Reese, M. G. and F. H. Eeckman (1995) Novel neural network algorithms for improved eukaryotic promoter site recognition. The seventh international genome sequencing and analysis conference. Hyatt Regency, Hilton Head Island, South Carolina. September 16-20.

Rouhbakhsh, D., C.-Y. Lai, C. D. von Dohlen, M. A. Clark, L. Baumann, P. Baumann, N. A. Moran and D. J. Voegtlin (1996) The tryptophan biosynthetic pathway of aphid endosymbionts (*Buchnera*): genetics and evolution of plasmid-associated anthranilate synthase (*trpEG*) within the Aphididae. *Journal of Molecular Evolution* **42:** 414-421.

Rozen, S. and H. J. Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. Bioinformatics methods and protocols: Methods in molecular biology. S. Krawetz and S. Misener. Totowa, NJ, Humana Press**:** 365-386.

Sandstrom, J. and N. Moran (1999) How nutritionally imbalanced is phloem sap for aphids? *Entomologia Experimentalis et Applicata* **91**: 203-210.

Sandstrom, J., A. Telang and N. A. Moran (2000) Nutritional enhancement of host plants by aphids — a comparison of three aphid species on grasses. *Journal of Insect Physiology* **46**: 33-40.

Sato, S. and H. Ishikawa (1997a) Expression and control of an operon from an intracellular symbiont which is homologous to the *groE* operon. *Journal of Bacteriology* **179**: 2300-2304.

Sato, S. and H. Ishikawa (1997b) Structure and expression of the *dnaKJ* operon of *Buchnera*, an intracellular symbiotic bacteria of aphid. *Journal of Biochemistry* **122**: 41-48.

Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki and H. Ishikawa (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* **407**: 81-86.

Shufran, K. A., L. R. Kirkman and G. J. Puterka (2007) Absence of mitochondrial DNA sequence variation in Russian wheat aphid (Hemiptera: Aphididae) populations consistent with a single introduction into United States. *Journal of the Kansas Entomological Society* **80**: 319-326.

Silva, F. J., R. C. H. J. van Ham, B. Sabater and A. Latorre (1998) Structure and evolution of the leucine plasmids carried by the endosymbiont (*Buchnera aphidicola*) from aphids of the family Aphididae. *FEMS Microbiology Letters* **168**: 43-49.

Soler, T., A. Latorre, B. Sabater and F. J. Silva (2000) Molecular characterization of the leucine plasmid from *Buchnera aphidicola*, primary endosymbiont of the aphid *Acyrthosiphon pisum*. *Current Microbiology* **40**: 264-268.

Srivastava, P. N. (1987) Nutritional Physiology. Aphids. Their biology, natural enemies and control. A. K. Minks and P. Harrewijn. Amsterdam, Elsevier. **2A:** 99-121.

Swanevelder, Z. H., A. K. J. Surridge, E. Venter and A.-M. Botha (2010) Limited endosymbiont variation in *Diuraphis noxia* (Hemiptera: Aphididae) biotypes from the United States and South Africa. *Journal of Economical Entomology* **103**: 887-897.

Tchufistova, L. S., A. V. Komarova and I. V. Boni (2003) A key role for the mRNA leader structure in translational control of ribosomal protein S1 synthesis in γ-proteobacteria. *Nucleic Acids Reseach* **31**: 6996-7002.

Telang, A., J. Sandstrom, E. Dyreson and N. A. Moran (1999) Feeding damage by *Diuraphis noxia* results in a nutritionally enhanced phloem diet. *Entomologia Experimentalis et Applicata* **91**: 493-412.

Thao, M. L., L. Baumann, P. Baumann and N. A. Moran (1998) Endosymbionts (*Buchnera*) from the aphids *Schizaphis graminum* and *Diuraphis noxia* have different copy numbers of the plasmid containing the leucine biosynthetic genes. *Current Microbiology* **36**: 238-240.

Tolmay, V. L., R. C. Lindeque and G. J. Prinsloo (2007) Preliminary evidence of a resistance-breaking biotype of the Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Homoptera: Aphididae), in South Africa. *African Entomology* **15**: 228-230.

Van Ham, R. C. H. J., J. Kamerbeek, C. Palacios, C. Rausell, F. Abascal, U. Bastolla, J. M. Fernandez, L. Jimenez, M. Postigo, F. J. Silva, J. Tamames, E. Viguera, A. Latorre, A. Valencia, F. Moran and A. Moya (2003) Reductive genome evolution in *Buchnera aphidicola*. *Proceedings of the National Academy of Sciences, USA* **100**: 581-586.

van Helden, J., B. André and J. Collado-Vides (2000) A website for the computational analysis of yeast regulatory sequences. *Yeast* **16**: 177-187.

Van Zyl, R. A. and A.-M. Botha (2008) Eliciting proteins from *Diuraphis noxia* biotypes differ in size and composition, *IPRI-18th Biennial Workshop*, Ft. Collins, USA, February 10-13. IPRI-18th Biennial Workshop, Ft. Collins, USA.

Weiland, A. A., F. B. Peairs, T. L. Randolph, J. B. Rudolph, S. D. Haley and G. J. Puterka (2008) Biotypic diversity in Colorado Russian wheat aphid (Hemiptera: Aphididae) populations. *Journal of Economic Entomology* **101**: 569-574.

Wernegreen, J. J. and N. A. Moran (2000) Decay of mutualistic potential in aphid endosymbionts through silencing of biosynthetic loci: *Buchnera* of *Diuraphis*. *Proceedings of the Royal Society of London* **267**: 1423-1431.

Wernegreen, J. J. and N. A. Moran (2001) Vertical transmission of biosynthetic plasmids in aphid endosymbionts (*Buchnera*). *Journal of Bacteriology* **183**: 785-790.

Wilcox, J. L., H. E. Dunbar, R. D. Wolfinger and N. A. Moran (2003) Consequences of reductive evolution for gene expression in an obligate endosymbiont. *Molecular Microbiology* **48**: 1491-1500.

Xayaphoummine, A., T. Bucher and H. Isambert (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Research* **33**: W605-W610.

CHAPTER 5

JUST HOW DO AFFYMETRIX NORMALIZATION METHODS COMPARE? STATISTICS CONTEMPLATE BIOLOGY

Published in part as:

Botha A-M., Z.H. Swanevelder and N.L.V. Lapitan (2010) Transcript profiling of wheat genes expressed during feeding by two different biotypes of *Diuraphis noxia*. *Journal of Environmental Entomology* **39(4):** 1206-1231.

Introduction

In the last decade researchers in biology has embraced high-throughput systems to measure gene expression at a transcriptome level. In these information intensive techniques, like cDNA microarrays and Affymetrix arrays, thousands of data points are screened and analyzed, with even larger amounts of data points being produced during an analysis. As in all technologies there are ambiguities that need to be corrected for, *i.e.* the intensities of hybridized spots do not necessarily reflects the correct level of an expressed transcript mRNA but also contains errors as a result of the technology – like PCR biases, hybridization effects, *etc*. (Kriel & Russell 2005; Harr & Schlötterer 2006; Steinhoff & Vingron 2006). Slide comparisons, biological replications, experimental designs and the identification of differentially expressed transcripts, further complicate these analyses.

High-throughput array analyses attempt to counter these possible ambiguities in three different rectifying steps (Kriel & Russell 2005). An initial image analysis deals with image inaccuracies like background corrections, hybridization artefacts, *etc*. (Kriel & Russell 2005; Steinhoff & Vingron 2006). Data pre-processing and normalization are usually done together and attempts to get rid of technical uncertainties while standardizing (*i.e.* normalizing) different hybridization events (*i.e.* slides) to the same scale that will allow direct comparisons between slides (Steinhoff & Vingron 2006). All these normalization methods utilize the Affymetrix GeneChip (Affymetrix, USA) arrays perfect match (*PM*) and/or mismatch (*MM*) probe pairs in some way (Affymetrix 2002, http://www.affymetrix.com/support/technical/whitepapers/sadd_whitepaper.pdf). In Affymetrix GeneChip arrays, the mRNA target molecules are represented on the array by probe sets consisting of 11-20 probe pairs. Each probe pair contains a *PM* probe to the target and a *MM* probe that has an altered nucleotide at the middle position, *i.e.* nucleotide 13 of the 25 bp length probe (Irizarray *et al.* 2003). The mismatch probes are used to measure non-specific hybridization. Expression levels of the transcript is calculated using the intensities of the hybridized target molecules to the various probe sets (Irizarray *et al.* 2003). After all these corrections, higher-level analyses are used to obtain statistical significant answers to biological questions (Kriel & Russell 2005). The analyses needed at each of these steps have prompted the development of numerous statistical applications (Smyth & Speed 2003; Smyth 2004; see Table 1 of Irizarray *et al.* 2006 for an extensive list; Elo *et al.* 2009).

However, with such a wide variety of statistical techniques suitable for application to a dataset, each with its own biases, assumptions and ambiguities, choosing the best approach that will produce statistical significant, biologically relevant results, can quickly become a nightmare.

Various comparisons of different normalization and pre-processing methods attempt to identify the best approach, but only add to the confusion with conflicting solutions to this problem (Bolstad *et al.* 2003; Irizarray *et al.* 2003; Shedden *et al.* 2005; Harr & Schlötterer 2006; Irizarray *et al.* 2006; Lim *et al.* 2007). Highly controlled dilution and/or spike-in calibration datasets, used in these comparisons, are blamed by some for the conflicting answers (Shedden *et al.* 2005).

As an alternative to approaches followed in all previous literature/studies, we propose to test the feasibility of using more than one normalization method when identifying the statistically significant genes expressed using an Affymetrix gene array set. Our hypothesis is that a subset of the Affymetrix gene identities (geneIDs)/probe sets, ascertained as differentially regulated, will be the same regardless of the normalization method utilized if all the other analyses are kept identical. This subset is therefore normalization-method-bias-independent and reflects more accurately significant differentially regulated transcripts. We've selected five different background correction and normalization methods to investigate their influences on the identification of specific, differentially expressed gene transcripts. The rest of the statistical methods were kept the same to investigate the relevance of these normalization and background correction methods. This allowed us to compare the unique Affymetrix geneIDs/probe sets ascertained as differentially regulated, at three statistical significant cut-offs: $p \leq 0.5, 0.1, 0.001$, and thus allowing us to assess how many times a specific geneID/probe set was identified by the different methods. We've also explored the effectiveness of false discovery rate (FDR) and family-wise type I error rate (FWER) approaches in nullifying any normalization-biases in order to produce more reliable differentially expressed gene identifications. During the final analyses to determine whether a specific gene was differentially expressed, we applied the criterion that a gene must be present at a significant level ($p \leq 0.5, 0.1, 0.001$) in at least four of the five normalization methods to fulfil the significant selection criteria.

Materials and methods

Experimental design

Two experiments were independently conducted by different researchers, with three different institutions used for the labelling, probing and scanning of the slides (*i.e.* the Centre for Proteomic & Genomic Research (CPGR), Cape Town, South Africa; the Microarray Core Lab (Aurora, CA) and the Virginia Bioinformatics Institute Core facility). A fractional design for the comparison analysis

of both the datasets was used (*Figure 5.1*). The two experiments used in this comparative analysis employed 12 and 18 GeneChip Wheat Genome Arrays (http://www.affymetrix.com/support/technical/datasheets/wheat_datasheet.pdf) and are referred to as the 12 and 18 slide experiments.

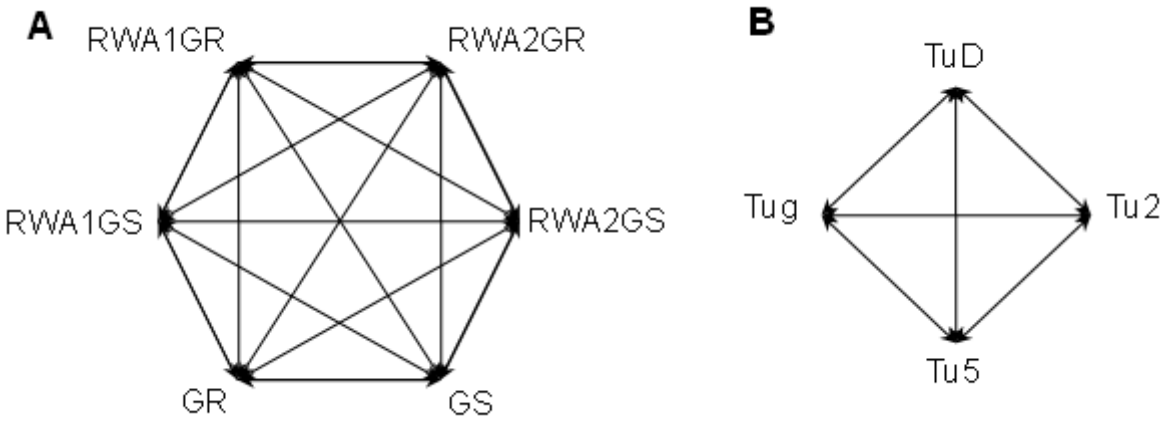


Figure 5.1 A 2 × 2 fractional design used in the analysis of the two experiments, with (A) representing the 18 slide experimental comparison and (B) the 12 slide comparison. Abbreviations (A): RWA1GR, *Dn7* resistant cultivar 94M370 infested with US RWA biotype 1; RWA1GS, the susceptible cultivar Gamtoos infested with US RWA biotype 1; RWA2GR, the *Dn7* resistant cultivar 94M370 infested with the US RWA biotype 2; RWA2GS, susceptible Gamtoos line infested by US RWA biotype 2; GR, the resistant cultivar 94M370 control that was un-infested and GS, the susceptible Gamtoos control, un-infested. Abbreviations (B): Tug, the susceptible cultivar Tugela; TuD, the resistant NIL TugelaDN; Tu2, the resistant NIL Tugela *Dn2* and Tu5, the resistant NIL Tugela *Dn5* - all these lines were infested with the South African RWA biotype.

Aphids and plant material

In the 12 slide experiment, the effects of RWA feeding on near isogenic wheat lines (NILs) between different cultivars infested with one RWA biotype (South African RWA biotype), was investigated, while the 18 slide experiment examined the effects of two RWA biotypes (USA biotypes 1 and 2) on resistant and susceptible wheat lines (Botha *et al.* 2010). In the 12 slide experiment, the South African (SA) RWA biotype (Du Toit 1989), obtained from the ARC-Small Grains Institute (Bethlehem, South Africa) and maintained on the susceptible wheat cultivar Scheepers, was used to infest three resistant (TugelaDN, Tugela *Dn2* and Tugela *Dn5*) lines and one susceptible wheat cultivar (Tugela). In the 18 slide experiment, two USA biotypes, RWA1 - the

original USA 1986 RWA introduction (Lapitan *et al.* 2007) and RWA2 - a biotype virulent to *Dn4* resistant winter wheat cultivars (Haley *et al.* 2004), were used to infest Gamtoos (susceptible) and the resistant line, 94M370 (*Dn7*). In short, in both experiments, plants were grown under greenhouse conditions (20 – 25 °C) until the 4 leaf stage, when plant infestations commenced in accordance to Botha *et al.* (1998). Total RNA was extracted with the Qiagen RNeasy Plant Mini Kit (Qiagen, USA), including on column DNA digestion with RNase Free/DNaseI (Qiagen), all in accordance with the manufacturer's instructions. RNA quality and integrity were tested using agarose gels and the Bio-Rad Experion RNA StdSen Chips (Bio-Rad, USA). The South African RNA samples were sent to CPGR (Cape Town, RSA) and the USA samples to the Microarray Core Lab (Aurora, CA) or to the Virginia Bioinformatics Institute Core facility, where additional quality control was performed. RNA labelling, hybridisation, processing and data gathering were all done in accordance to the Affymetrix protocols at these facilities (Botha *et al.* 2010). A full description of the experimental layout can be found in Botha *et al.* (2010).

Data analysis

Statistical analysis was carried out using R 2.6.0 (http://CRAN.R-project.org/, Ihaka & Gentleman 1996) and Bioconductor 2.1 (http://www.bioconductor.org, Gentleman *et al.* 2004). Additional packages, like the affy (Gautier *et al.* 2004) and affyPLM (Bolstad 2007) packages, were obtained from the Bioconductor website (http://www.bioconductor.org/packages/) as needed. Scripts were executed on a Sun Fire V880 (4 × 1.05 GHz SuperSPARC CPU's with 8 GB RAM) running the Solaris 9 operating system, with all R script outputs written to file. The raw data obtained from the various institutions were analyzed as described in *Figure 5.2* using the scripts, codes and files summarized in *Table Appx 5.1* and script source code (Appendix Chapter 5). In short, information in the raw data files (*.CEL* files), obtained from the various institutions were used for the different pre-processing and summarization methods. Phenotypic data, including biological repeats, technical repeats, sample names to be used in analysis, slide numbers and *CEL* filenames, were stated in a tab delimited text file (*Target.txt*, Appendix Chapter 5) for each of the experiments and included in the raw dataset used for each analysis. The raw data was subjected to different quality control checks which included the inspection of the hybridized images, histograms and box plots of $\log_2$(PM) values and the examination of hybridizations. Background correction and normalization were performed using Affymetrix Microarray Suite5 (MAS5.0) (Harr and Schlötterer, 2006), GeneChip

Robust Multichip Average (GCRMA) (Zakharkin *et al.*, 2005), Probe Level Models (PLM) (Bolstad *et al.*, 2003), Robust Multichip Average (RMA) (Irizarry *et al.*, 2003) and Variance Stabilisation (VSN) (Huber *et al.*, 2002). The outputs of these different methods were saved in corresponding *RData* files (*Figure 5.2*).

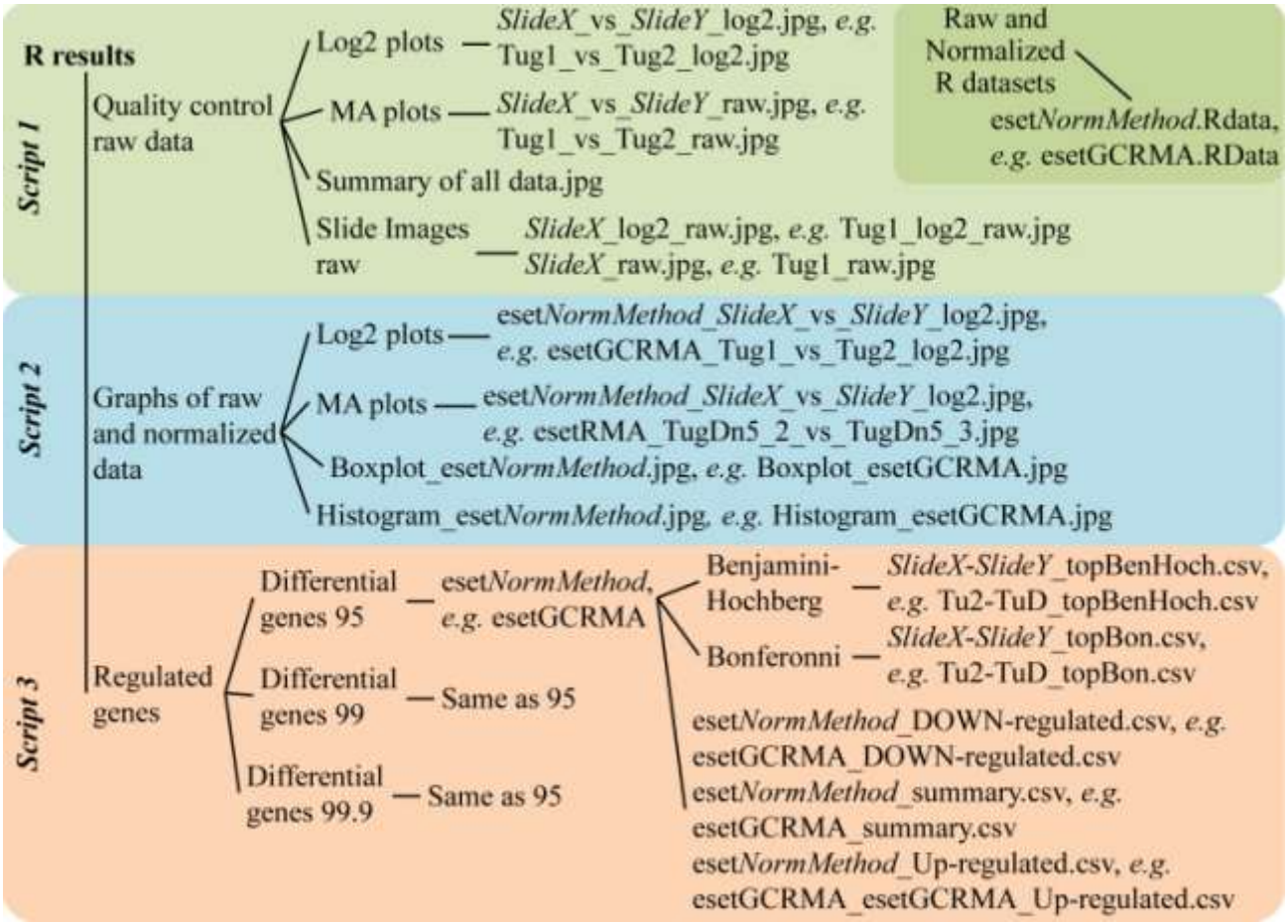


Figure 5.2 A layout of the analysis conducted on each experiment, including some of the output folders and filenames (with examples), as produced by the different R scripts. Files and folders produced during the running of script 1 are represented by a green block, with script 2 and 3 represented by a blue and red block respectively. Variable portions of filenames are given in italics. Details on the input files and the output files/folders are given within the header of each script (*Table Appx 5.1*).

Differentially expressed genes, for datasets of each different background correction and normalization method, were determined for three different *p*-values: 0.05, 0.01 and 0.001. These differentially expressed genes were identified from the different normalized datasets by applying the same moderated *t*-test of the limma (linear models for microarrays, LMM) package and Bayesian analysis (BA) (Smyth & Speed 2003; Smyth 2004). Transcripts were deemed differentially expressed if it had a $\log_2$ fold change larger than 1 ($[\log_2(FC)] >1$) and a *p*-value (or an adjusted p-value) of less-than-or-equal-to 0.05, 0.01 or 0.001 respectively. The comparison-wise false discovery rate (FDR) was controlled with the Benjamini-Hochberg multiple testing adjustment (Benjamini and Hochberg 1995) and the family-wise type I error rate (FWER) by the Bonferroni method (Dudoit & Ge 2004). This produced three datasets before FDR/FWER were determined (95, 99 and 99.9 % confidences) and three for each of the two correction methods.

A MS Excel 2007 macro (*Convert*, Appendix 5) was used to combine all the output spreadsheet files into summary spreadsheets (*Figure 5.3*). The number of times a geneID/probe set was obtained using the different background correction and normalization functions, were calculated in MS Excel 2007. The presence/absence of a regulated GeneID in a slide comparison was represented with a 1/0. Adding these numbers results in the number of times a GeneID/probe set was flagged (for a specific *p*-value) when data was pre-processed differently. The summarized data was used to compile the different graphs using Sigma Plot (Jandel Scientific Inc., USA).

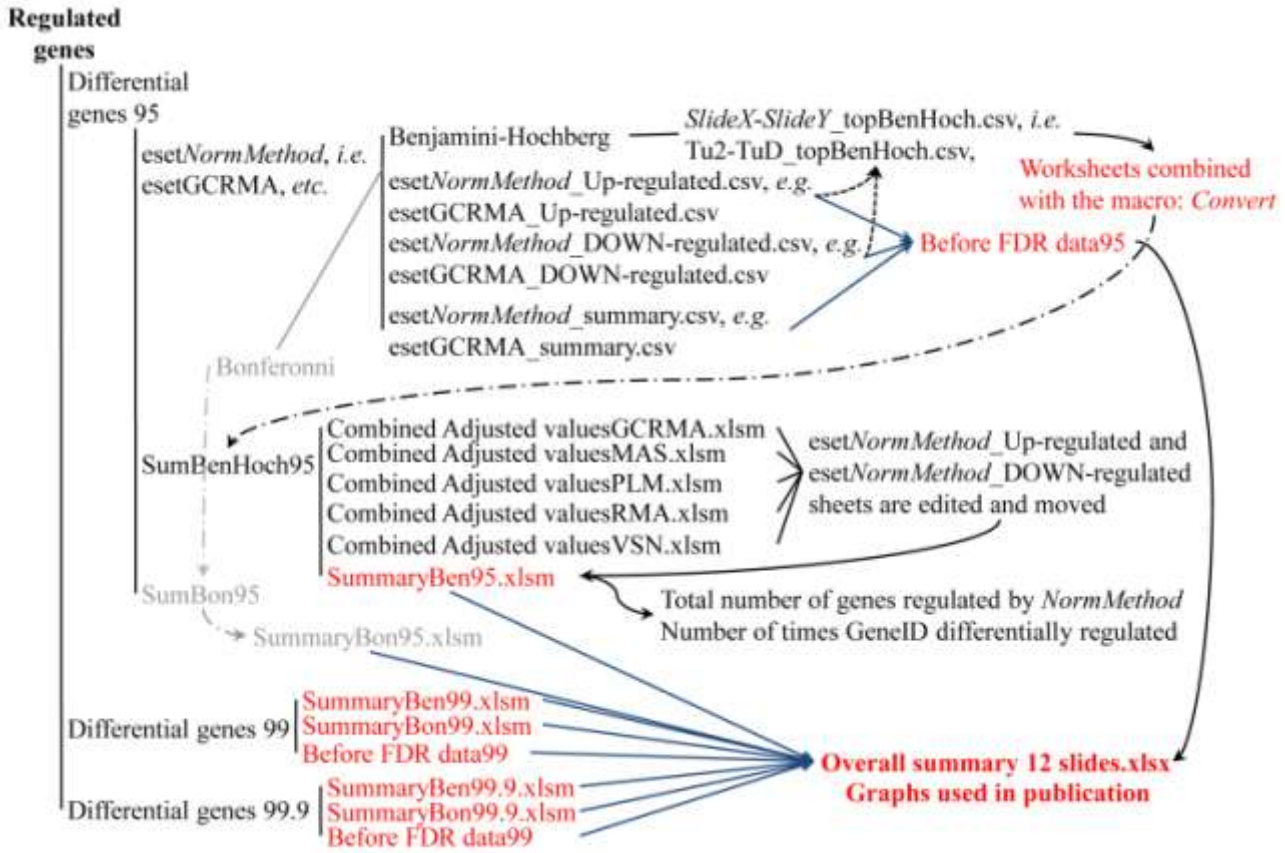


Figure 5.3 The work flow of the data analysis done on the differentially regulated genes, before and after Bonferroni (Bon) and Benjamini-Hochberg (Ben or BenHoch) corrections for 95, 99 and 99.9 % confidences. The same procedures followed for Benjamini-Hochberg analysis were done for Bonferroni. These analyses were repeated for each of the different confidence levels. Files added after the analyses are indicated in red. The summary files are given in bold red.

Results

Quality control of slides, background correction and normalization

Quality control started with an inspection of the original scanned slide images (*Figure Appx 5.1*) to check for any ambiguities on the slides themselves. This was followed by an overall summary figure that contained a histogram, dendogram and box plot of the slides used in the experimental layout (*Figure Appx 5.2*). Other quality control figures produced for the raw data sets, before normalization and background correction, included the $\log_2$ expression comparisons between the pre-normalized slides (*Figure Appx 5.3*) and MA-plots of these slides (*Figure Appx 5.4*). In all

cases where Bioconductor and R modules allowed, additional information like legend names, were incorporated into the figures using the data from the "*Target.txt"* file (Appendix 5). The quality control steps, after normalization and background corrections with GCRMA, MAS5.0, PLM, RMA and VSN, consisted out of summaries of the treatment comparisons using histograms (*Figure Appx 5.5.*), box plots of the $\log_2$(PM) values (*Figure 5.4* and *Figure Appx 5.6*), $\log_2$ expression slide comparisons (*Figure Appx 5.7*) and MA-plot comparisons (*Figure Appx 5.8*). The pre-normalization dataset was also included in these analyses to allow comparisons with the normalized datasets. An example of the results, before and after background correction and normalization with the different methods, are presented in *Figure 5.4* as different box plots of $\log_2$(PM) values of the 12 slide experiment. These box plots, when compared to the raw dataset plot, show the influence of the normalization process by the different methods employed.

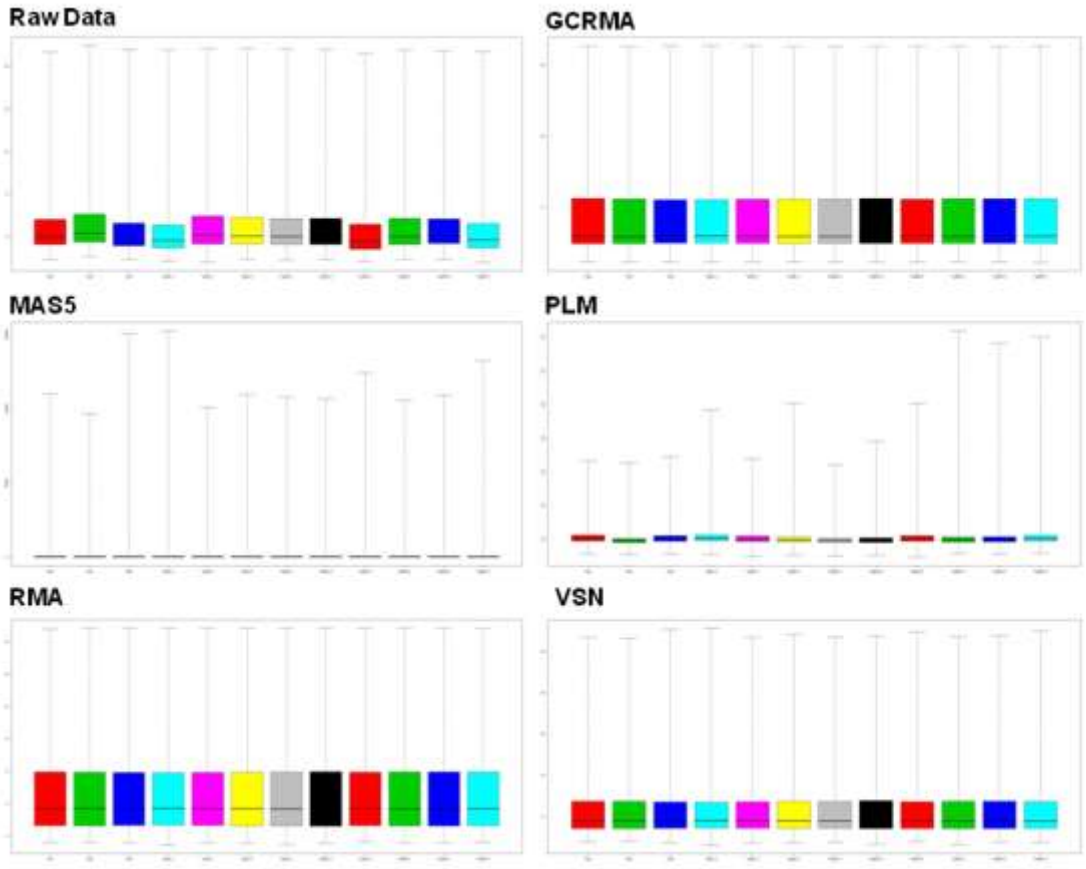


Figure 5.4 Box plots of the $\log_2$ (PM) values, before and after background correction and normalization were done on the 12 slide dataset with the different statistical approaches. All the sample slides occur in the same order in the various graphs in their corresponding triplicate slides (Tugela, TugelaDN, Tugela *Dn2* and Tugela *Dn5*) but that the Y-axis differs in accordance to the method used.

Normalization and background correction methods

The number of times a specific normalization method identified the most or the least number of probe sets, before and after FDR/FWER correction, were compared (*Figure 5.5*). Methods that tend to be less strict in their criteria should more often produce more probe sets than those with stricter criteria. In both the 12 and 18 slide experiments, before FDR/FWER corrections, the PLM method produced the most probe sets across slide comparisons identified as up regulated for all three confidence levels (95, 99 and 99.9 %). The PLM method also produced the most down regulated probe sets more often for the 12 slide experiment slide comparisons, before FDR/FWER correction and across all confidences. The VSN method delivered the most down regulated probe sets for the 18 slide experiment (*Figure 5.5* (*A*) and (*D*)). The MAS5 correction method produced the least number of probe sets more often for both the up and down regulated 12 slide experiment across all the confidence levels and slide comparisons. The VSN method produced more regularly the least number of probe sets for the 18 slide up regulated slide comparisons across the different confidence levels, but different methods contributed for the least number of probe sets for the 18 slide down regulated slide comparisons (*Figure 5.5* (*A*) and (*D*)).

The mean number of probe sets across the difference confidence levels, supported these method percentages (*Figure 5.5*) as producing the largest and smallest number of probe sets for a slide comparison before FDR/FWER correction (*Figure 5.6*). The average probe set number also decreased as the confidence levels increased. However, large differences existed in the number of probe sets identified as differentially regulated for the various slide comparisons within the different experiments. These differences resulted in large errors of their means (*Figure 5.6*). Therefore, significant differences in mean probe set numbers of the different normalization methods are usually only observed between methods producing the mean most and the mean least number of probe sets (*Figure 5.5*). Furthermore, as the number of slides analyzed increased, the average number of probe sets identified as differentially regulated, also increased, from <5 000 in the 12 slide experiment to >10 000 in the 18 slide experiment (*Figure 5*.6), thereby further hampering direct comparisons between the various methods.

In both experiments, the Bonferroni FWER correction and Benjamini-Hochberg FDR correction usually lowered the total mean number of genes identified as differentially regulated (*Figure 5.6*). At the same time these correction methods usually brought the average number of probe sets across different confidence levels, closer together (*Figure 5.6*). In the 12 slide experiment after FWER correction, the RMA method usually produced the most probe sets in the different slide

comparisons (*Figure 5.5* (*B*)). However, for the same experiment, Benjamini-Hochberg FDR correction resulted in the PLM method usually produced the most probe sets under the various normalization conditions (*Figure 5.5* (*C*)). The least number of probe sets after FDR and FWER corrections, for both up and down regulated genes and across all the different confidences, were obtained in 10 of the 12 cases after normalization with the MAS5 method (*Figure 5.5* (*B*) and (*C*)). The VSN method contributed the most genes for the up regulated and the PLM method for the down regulated comparisons, after both FDR and FWER correction and under the various confidence levels for the 18 slide experiment (*Figure 5.5* (*E*) and (*F*)). The Bonferroni FWER corrections for this experiment usually had the MAS5 method producing the least number of probe sets for the different slide comparisons (*Figure 5.5* (*E*)). The least number of probe sets identified as differentially regulated after Benjamini-Hochberg FDR correction was obtained from the MAS5, VSN and the GCRMA methods (*Figure 5.5* (*E*)).

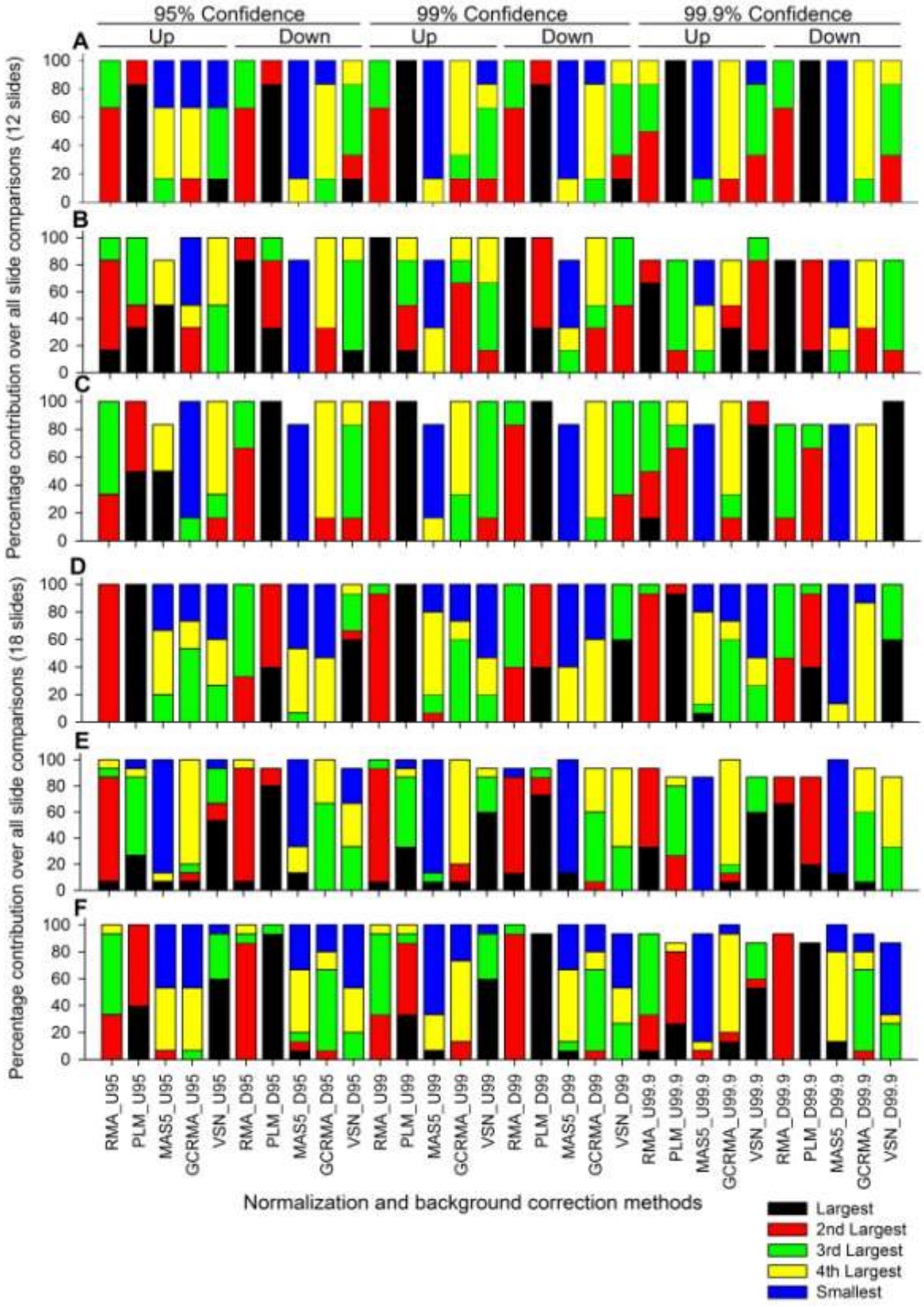

95% Confidence
99% Confidence
99.9% Confidence
Up
Down
A
B
C
D
E
F
Percentage contribution over all slide comparisons (12 slides)
Percentage contribution over all slide comparisons (18 slides)
RMA_U95
PLM_U95
MAS5_U95
GCRMA_U95
VSN_U95
RMA_D95
PLM_D95
MAS5_D95
GCRMA_D95
VSN_D95
RMA_U99
PLM_U99
MAS5_U99
GCRMA_U99
VSN_U99
RMA_D99
PLM_D99
MAS5_D99
GCRMA_D99
VSN_D99
RMA_U99.9
PLM_U99.9
MAS5_U99.9
GCRMA_U99.9
VSN_U99.9
RMA_D99.9
PLM_D99.9
MAS5_D99.9
GCRMA_D99.9
VSN_D99.9
Normalization and background correction methods
Largest
2nd Largest
3rd Largest
4th Largest
Smallest

Figure 5.5 *(previous page)* Normalization method contributions, according to the number of genes differentially up or down regulated, were grouped into the largest, second largest, *etc*. group for each specific slide comparison. The number of times a specific method, expressed as total percentage contribution across all slides, were then calculated and presented here, *i.e.* how many times does a method produce the most, second most, *etc*. number of genes for a specific confidence level, across all the slide comparisons. These values were calculated before FDR/FWER correction ((A), 12 slides and (D), 18 slides) and after FWER/FDR corrections (Bonferroni: (B), 12 slides and (E), 18 slides; Benjamini-Hochberg: (C), 12 slides and (F), 18 slides). The total percentage contribution (y-axis) for that method is less than a 100 % if there are zero genes identified as differentially regulated for one or more slide comparisons. Slide comparisons sometimes delivered the same number of differentially regulated genes for two normalization methods. In these cases, both methods were placed in the same group. Following normalization method identification, "U" indicates up and "D" down regulated comparisons, followed by the confidence level (95, 99 and 99.9 %).

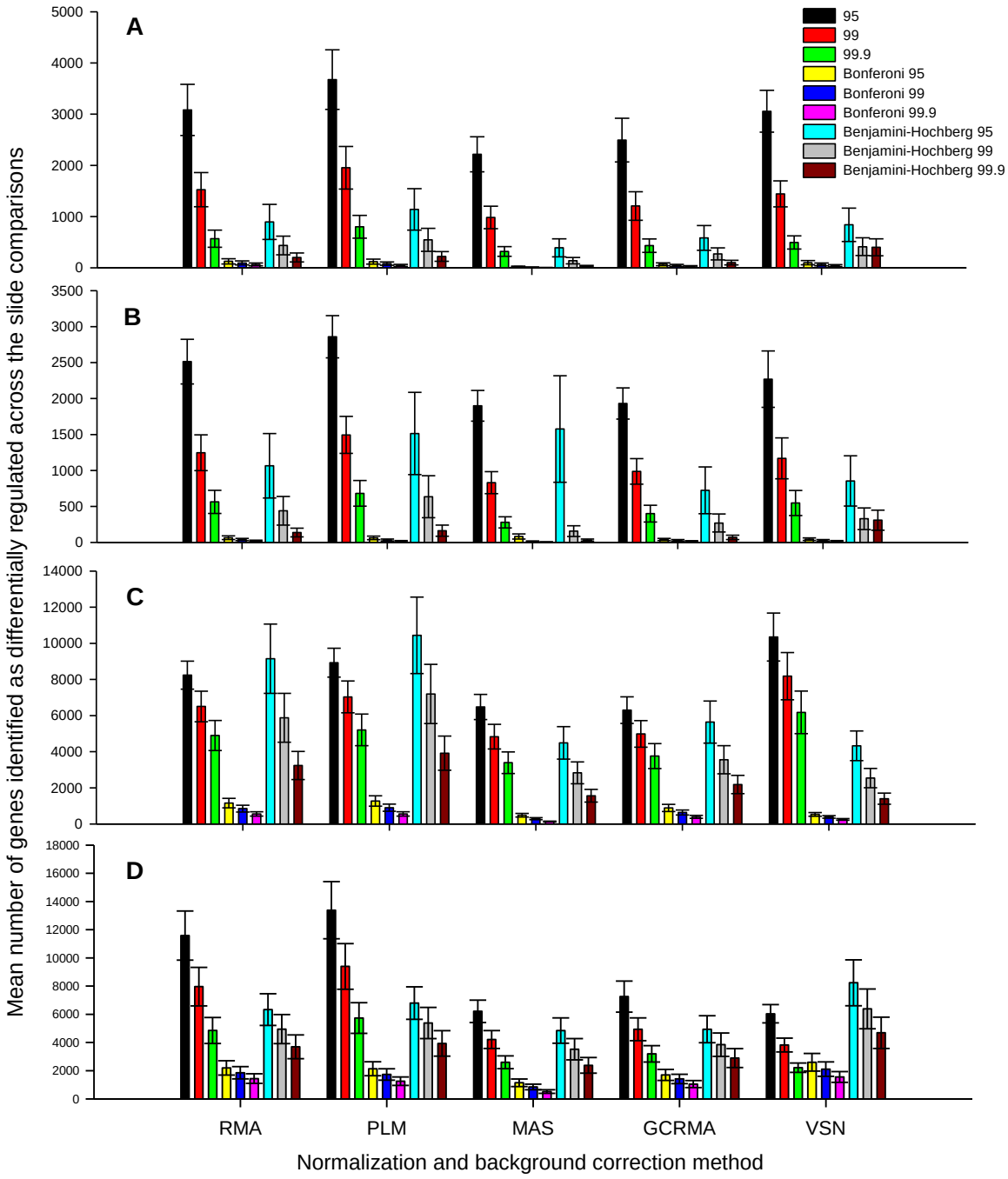


Figure 5.6 The mean number of genes/probe sets found to be differentially regulated as found by the LMM and BA after normalization with the five methods and at 95, 99 and 99.9 % confidence, before/after FDR/FWER corrections. The standard error over all the slide comparisons is indicated by the error bars. (A) and (B) represents the down and up regulated 12 slide experiment respectively, while (C) and (D) represent the down and up regulated genes of the 18 slide experiment, respectively.

Identifying false positives

Differentially expressed genes were identified using linear mixed models (LMM) and Bayesian analysis (BA) for each of the background and normalization methods at 95, 99 and 99.9 % confidence levels and included only those genes that had a $\log_2$ fold change larger than 1. The top differentially regulated probe sets after each normalization method and those unique to a specific slide comparison, for the 12 slide experiment, are given as examples in *Table 5.1* and *Table Appx 5.2*. The FDR/ FWER for these datasets were determined and saved at the set confidence for each slide comparison for a specific background and normalization method. All these datasets were used to calculate the number of times a specific geneID/probe set was differentially regulated after different normalization methods, at different confidences, with/without FWER/FDR correction. Examples of these analyses, *i.e.* with the FDR/FWER corrections for the selected probe sets at each confidence level, are given in *Tables 5.2, 5.3* and *Table Appx 5.3*.

Differentially up- or down regulated genes were determined independently, as well as in combined sets, for both the 12 and 18 slide experiments. These were summarized and used to compute totals of all the percentages of differentially regulated (up, down or combined) GeneIDs/probe sets for the different slide comparisons calculated to be present after 1, 2, 3, 4, or 5 normalization methods, under various confidences (95, 99 and 99.9 %), with/without FDR/FWER corrections. The averages of these total percentages summarize differentially regulated probe sets for all the slide comparisons, thereby allowing for direct comparisons between the methods and experiments (*Figure 5.7*). In all cases, the largest percentages of geneIDs/probe sets were found after normalization by only a single normalization method, regardless of FDR/FWER corrections (*Figure 5.7*). Indeed, an increase in confidence levels and/or with FDR/FWER correction usually increase the percentage probe sets identified as differentially regulated after correction by only a single normalization method, for example a 15 % increase was observed with an increase in confidence (95 vs. 99.9 % confidence, before correction, *Figure 5.7* (C)). Similarly, an average of ca. 30 % at 95 % confidence before FDR/FWER correction can increase to more than 45 % at a 99.9 % confidence with FDR correction (*Figure 5.7* (C)). The number of probe sets deemed differentially regulated after normalization by all 5 methods, as total differentially regulated genes, are < 35 % for the 18 slide experiment, and even smaller for the 12 slide experiment (< 25 %) (*Figure 5.7*).

Table 5.1 Selected geneIDs/probe sets shown to be up regulated (12 slide experiment), either generally or for specific slide comparisons (TuD-Tug, Tu2-Tug, *etc*.). Values under the different normalization methods (VSN, RMA, *etc*.) indicate the number of times the specific probe set was obtained across the six slide comparisons within that specific normalization method, *i.e.* 6 indicates that the probe set was obtained in all six slide comparisons. The slide comparisons indicate the number of times the probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences. *Table Appx 5.2* shows examples of down regulated probe sets.

Probe set ID	***VSN***	***RMA***	***PLM***	***MAS5***	***GCRMA***	***Occ. 95 %***	***TuD - Tug***	***Tu2 - Tug***	***Tu5 - Tug***	***Tu2 - TuD***	***Tu5 - TuD***	***Tu5 - Tu2***	***Occ. 99 %***	***Occ. 99.9 %***
Ta.28669.2.S1_x_at	6	6	6	5	6	**29**	4	5	5	5	5	5	18	16
Ta.9243.1.S1_at	6	6	6	4	6	**28**	4	5	5	4	5	5	21	15
Ta.28669.1.S1_a_at	5	6	6	5	5	**27**	5	5	5	5	5	2	21	18
Ta.10120.1.S1_at	1	1	1	1	1	**5**	5	0	0	0	0	0	3	1
Ta.10120.1.S1_x_at	1	1	1	1	1	**5**	5	0	0	0	0	0	1	0
Ta.10311.1.S1_at	1	1	1	1	1	**5**	5	0	0	0	0	0	4	3
Ta.6995.1.S1_x_at	1	1	1	1	1	**5**	0	5	0	0	0	0	0	0
Ta.1200.1.S1_x_at	1	1	2	2	1	**7**	0	5	0	2	0	0	0	0
Ta.18713.1.S1_s_at	2	1	1	2	1	**7**	0	5	0	2	0	0	1	0
TaAffx.110724.1.S1_at	2	1	1	1	1	**6**	0	0	5	0	0	1	0	0
Ta.12770.1.S1_at	2	2	2	2	2	**10**	0	0	5	0	0	5	2	0
Ta.14005.1.S1_s_at	2	2	2	2	2	**10**	0	0	5	0	0	5	1	0
Ta.10381.1.S1_at	1	1	1	1	1	**5**	0	0	0	5	0	0	1	0
Ta.1055.1.S1_at	1	1	1	1	1	**5**	0	0	0	5	0	0	1	1
Ta.1055.1.S1_x_at	1	1	1	1	1	**5**	0	0	0	5	0	0	2	1
Ta.11414.3.S1_a_at	1	1	1	1	1	**5**	0	0	0	0	5	0	0	0
Ta.6169.1.S1_at	1	1	2	1	1	**6**	0	0	0	0	5	1	1	0
Ta.6626.1.A1_at	1	1	1	2	1	**6**	0	0	0	0	5	1	2	0
Ta.10269.3.S1_at	1	1	1	1	1	**5**	0	0	0	0	0	5	0	0
Ta.10354.2.S1_x_at	1	1	1	1	1	**5**	0	0	0	0	0	5	3	2
Ta.10354.3.S1_x_at	1	1	1	1	1	**5**	0	0	0	0	0	5	2	0

Tug, Tugela; TuD, TugelaDN, Tu2, Tugela Dn2; Tu5, Tugela Dn5

Table 5.2 GeneIDs/probe sets, selected in *Table 5.1*, that were still shown to be differentially regulated after FDR (Benjamini-Hochberg method) or FWER (Bonferroni) correction. Values under the different normalization methods (VSN, RMA, *etc.*) indicate the number of times the specific probe set was obtained across the six slide comparisons (TuD-Tug, Tu2-Tug, *etc.*) within that specific normalization method, *i.e.* 6 indicates that the probe set was obtained in all six slide comparisons. The slide comparisons indicate the number of times the probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences. *Table Appx 5.3* shows examples of differentially up regulated probe sets after FDR/FWER correction.

Probe set ID	***VSN***	***RMA***	***PLM***	***MAS5***	***GCRMA***	***Occ. 95%***	***TuD - Tug***	***Tu2 - Tug***	***Tu5 - Tug***	***Tu2 - TuD***	***Tu5 - TuD***	***Tu5 - Tu2***	***Occ. 99%***	***Occ. 99.9%***
After FDR														
Ta.28669.2.S1_x_at	3	3	5	2	3	**16**	0	4	5	1	5	1	9	6
Ta.9243.1.S1_at	3	3	3	2	2	**13**	0	3	5	0	5	0	8	4
Ta.28669.1.S1_a_at	3	3	4	3	3	**16**	0	5	5	1	5	0	8	4
TaAffx.110724.1.S1_at	1	1	1	0	0	**3**	0	0	3	0	0	0	0	0
Ta.12770.1.S1_at	0	0	1	0	0	**1**	0	0	1	0	0	0	0	0
Ta.14005.1.S1_s_at	0	0	1	0	0	**1**	0	0	0	0	0	1	0	0
Ta.1055.1.S1_at	1	0	0	0	0	**1**	0	0	0	1	0	0	0	0
Ta.1055.1.S1_x_at	1	0	0	0	0	**1**	0	0	0	1	0	0	0	0
Ta.11414.3.S1_a_at	0	0	1	0	0	**1**	0	0	0	0	1	0	0	0
Ta.6169.1.S1_at	0	0	1	0	1	**2**	0	0	0	0	2	0	0	0
Ta.6626.1.A1_at	0	1	1	0	0	**2**	0	0	0	0	2	0	0	0
Ta.10354.2.S1_x_at	1	1	1	0	1	**4**	0	0	0	0	0	4	2	0
Ta.10354.3.S1_x_at	1	1	1	0	0	**3**	0	0	0	0	0	3	1	0
After FWER														
Ta.28669.2.S1_x_at	0	1	1	0	0	**2**	0	0	2	0	0	0	1	0
Ta.9243.1.S1_at	0	1	1	0	0	**2**	0	0	2	0	0	0	1	0
Ta.28669.1.S1_a_at	0	0	1	0	0	**1**	0	0	1	0	0	0	1	0

*Only geneIDs/probe sets of *Table 5.1* were included if still deemed differentially expressed after FDR or FWER correction. Tug, Tugela; TuD, TugelaDN, Tu2, Tugela Dn2; Tu5, Tugela Dn5

Table 5.3 The influences of an increased confidence on the same three geneIDs/probe sets, before and after FDR/FWER correction. Values under the different normalization methods (VSN, RMA, *etc.*) indicate the number of times the specific probe set was obtained across the six slide comparisons (TuD-Tug, Tu2-Tug, *etc.*) within that specific normalization method, *i.e.* 6 indicates that the geneID/probe set was obtained in all six slide comparisons. The slide comparisons indicate the number of times the probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences.

Probe set ID	***Confi-dence***	***VSN***	***RMA***	***PLM***	***MAS5***	***GCRMA***	***Occ.***	***TuD - Tug***	***Tu2 - Tug***	***Tu5 - Tug***	***Tu2 - TuD***	***Tu5 - TuD***	***Tu5 - Tu2***
Before FDR/FWER													
TaAffx.110208.1.S1_at	0.05	6	6	6	6	6	**30**	5	5	5	5	5	5
TaAffx.26346.1.S1_at	0.05	5	6	5	5	6	**27**	5	5	5	5	5	2
Ta.4593.1.A1_at	0.05	6	5	5	5	5	**26**	5	5	5	5	5	1
TaAffx.110208.1.S1_at	0.01	5	6	6	6	6	**29**	5	5	5	4	5	5
TaAffx.26346.1.S1_at	0.01	5	5	5	5	5	**25**	5	5	5	5	5	0
Ta.4593.1.A1_at	0.01	4	4	5	5	4	**22**	5	5	5	2	5	0
TaAffx.110208.1.S1_at	0.001	4	4	6	4	4	**22**	1	5	5	1	5	5
TaAffx.26346.1.S1_at	0.001	3	4	4	4	3	**18**	0	5	5	3	5	0
Ta.4593.1.A1_at	0.001	3	3	3	4	2	**15**	1	5	5	0	4	0
After FDR													
TaAffx.110208.1.S1_at	0.05	3	4	5	3	4	**19**	0	3	5	1	5	5
TaAffx.26346.1.S1_at	0.05	3	3	3	3	2	**14**	0	4	5	0	5	0
Ta.4593.1.A1_at	0.05	3	3	3	3	1	**13**	0	4	5	0	4	0
TaAffx.110208.1.S1_at	0.01	3	3	4	3	3	**16**	0	1	5	0	5	5
TaAffx.26346.1.S1_at	0.01	1	3	2	1	1	**8**	0	1	5	0	2	0
Ta.4593.1.A1_at	0.01	1	1	3	1	1	**7**	0	1	5	0	1	0
TaAffx.110208.1.S1_at	0.001	3	2	3	1	2	**11**	0	0	5	0	4	2
TaAffx.26346.1.S1_at	0.001	1	1	1	0	0	**3**	0	0	3	0	0	0
Ta.4593.1.A1_at	0.001	1	1	1	1	0	**4**	0	0	4	0	0	0

***Table 5.3* cont.**

After FWER													
TaAffx.110208.1.S1_at	0.05	1	1	3	1	1	**7**	0	0	5	0	1	1
TaAffx.26346.1.S1_at	0.05	0	1	0	0	0	**1**	0	0	1	0	0	0
Ta.4593.1.A1_at	0.05	0	0	0	0	0	**0**	0	0	0	0	0	0
TaAffx.110208.1.S1_at	0.01	1	1	2	0	1	**5**	0	0	4	0	1	0
TaAffx.26346.1.S1_at	0.01	0	0	0	0	0	**0**	0	0	0	0	0	0
Ta.4593.1.A1_at	0.01	0	0	0	0	0	**0**	0	0	0	0	0	0
TaAffx.110208.1.S1_at	0.001	0	0	2	0	0	**2**	0	0	1	0	1	0
TaAffx.26346.1.S1_at	0.001	0	0	0	0	0	**0**	0	0	0	0	0	0
Ta.4593.1.A1_at	0.001	0	0	0	0	0	**0**	0	0	0	0	0	0

*For comparison reasons, geneIDs/probe sets no longer deemed differentially expressed after FDR/ FWER correction were included. Tug, Tugela; TuD, TugelaDN, Tu2, Tugela Dn2; Tu5, Tugela Dn5.

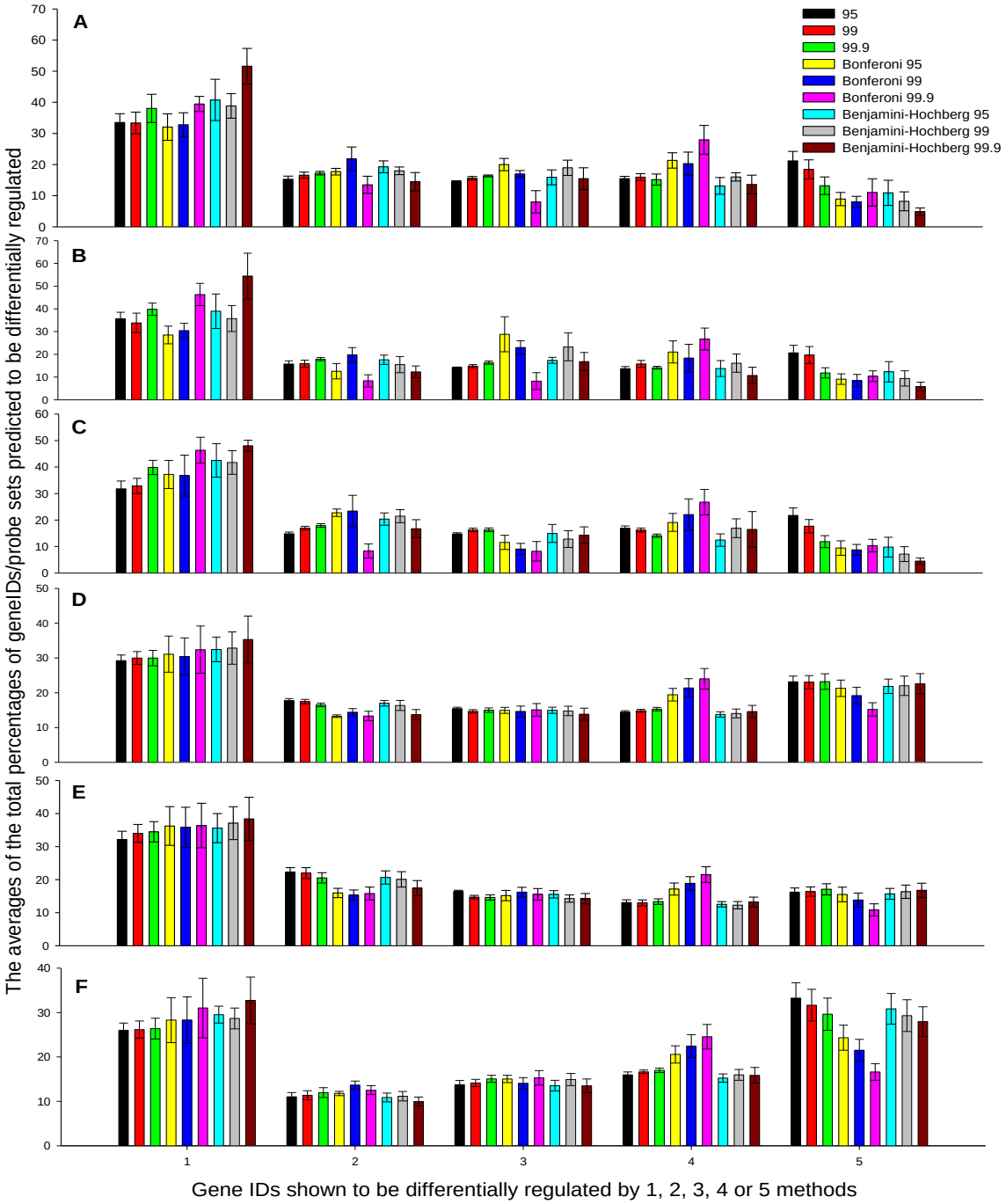


Figure 5.7 The averages of the total percentages of geneIDs/probe sets predicted to be differentially regulated, after normalization at different confidences (95, 99, 99.9 %) and before/after FDR/FWER corrections. The standard errors were calculated on the percentage contribution of each slide comparison to a 1, 2, 3, 4, or 5 method occurrence that was initially used to determine the mean sum (average total) percentage for each specific confidence (95, 99, 99.9 %) and before/after FDR/FWER correction. (A)-(C) represent the 12 slide experiment, with (A) the total differentially regulated genes, *i.e.* the combined down and up regulated percentages, while (B) and (C) represent the down and up regulated genes respectively. Similarly (D)-(F) represent the 18 slide experiment, with (D) the combined, (E) the down and (F) the up regulated differentially expressed gene.

Discussion

Analyzing Affymetrix data can be a daunting prospect, especially with all the numerous normalization and background correction methods available (Smyth & Speed 2003; Smyth 2004; Irizarray *et al.* 2006; Elo *et al.* 2009). The necessity of using these methods is obvious when comparing pre-normalized and normalized datasets, but at the same time the variation in results between the different normalized datasets becomes noticeable (*Figure 5.4* and *Figures Appx 5.2-5.8*). Indeed, normalization procedures are performed to correct for systematic deviations inherent of the technology (Kriel & Russell 2005; Harr & Schlötterer 2006). Increasing the significant cut-offs, *i.e.* confidence levels, or using FDR/FWER corrections methods, can be conceived as suitable countermeasures to any biases introduced by these normalization and background correcting methods. Indeed, an increase in confidence values or in the presence of FDR/FWER correction, the number of times probe sets are categorized as differentially expressed usually does decrease the significant geneIDs/probe sets. However, the same can be observed on a method specific basis (probe sets: Ta.10120.1.S1_at, Ta.10120.1.S1_x_at, Ta.10311.1.S1_at in *Tables 5.1* and *5.2*, also all three accessions of *Table 5.3*). This suggests that an increase in confidence levels or the use of correction methods on a dataset will not neglect the normalization method biases.

The normalization and background correction of Affymetrix slides using different statistical approaches should theoretically produce a large number of similar probe sets, especially if the methods only normalized for the experimental method errors that they were designed for. However, since each method contains inherit biases and statistical assumptions, this is not necessary the case (Kriel & Russell 2005; Shedden *et al.* 2005; Harr & Schlötterer 2006). Indeed, previous studies have shown that normalization methods initially used, have huge implications in detecting differentially regulated genes (Hoffmann *et al.* 2002; Shedden *et al.* 2005), and that these initial correction events have the largest influence on identifying differentially expressed genes (Hoffmann *et al.* 2002; Irizarray *et al.* 2006). This is clearly illustrated in *Figure 5.7*, with the exception for the differentially up regulated part of the 18 slide experiment (*Figure 5.7* (F)), in that more than a third of the geneID/probe set percentages were only associated with a single normalization method. Arguably, this does not necessary indicate that a third of the regulated probe sets are method specific, but that a single normalization method, less strict than the others, could produce more probe sets, a higher probe set percentage and a lower likelihood that another method would have the same probe sets. Similarly, the opposite can also be true. Indeed, there seems to be large differences in the amount of predicted regulated geneIDs/probe sets after normalization with the different

methods, especially at lower confidence levels (95 % No FDR/FWER, *Figure 5.6*). This suggests that an increase in selection criteria stringency, *i.e.* higher confidence levels (99.9 %), fewer geneIDs/probe sets would be unique since highly regulated individuals are preferentially selected. The existence of normalization-method-dependent, differentially regulated, probe set identifications have previously been reported, however only at a single confidence level (Hoffmann *et al.* 2002). Here, the number of probe sets identified as differentially regulated, does decrease to become more comparable when confidence levels increase (99.9 % No FDR/FWER, *Figure 5.6*). However, the averages of the total percentages of geneIDs predicted to be differentially regulated, after normalization at different confidences (95, 99, 99.9 %) and before/after FDR/FWER corrections, occurring in only a single method also increase with this increase in confidence, while those geneIDs/probe sets occurring in all five methods, decrease (*Figure 5.7*). This suggests that a confidence increase does not counteract a normalization method's inherited biases, but does appear to shift geneIDs towards single method identification, *i.e.* increasing the bias.

Normalization method tendencies to produce the most or the least amount of genes for slide comparisons, were investigated for both experiments (*Figure 5.5*). Differences in the number of differentially regulated probe sets have been previously reported for different normalization methods initially used (Hoffmann *et al.* 2002; Shedden *et al.* 2005), as was the case here. In the 12 slide experiment, before FDR/FWER correction, PLM produced the most genes in both up and down regulated slide comparisons at all three confidence levels, with MAS5 the least number of genes (*Figure 5.5*). However, in the 18 slide experiment, PLM only contributed the most genes in the up regulated comparisons, while VSN contributed the largest amount of genes for the down regulated comparisons (*Figure 5.5*). After FDR/FWER correction, however, this switched around, VSN contributes the most genes for the up regulated comparisons, and PLM for the down regulated comparisons. The least number of genes for the up regulated slide comparisons, before FDR/FWER correction, were at each confidence level VSN, while it differed for the down regulated slide comparisons (*Figure 5.5*). FWER correction with Bonferroni of the 18 slide comparisons usually resulted in MAS5 having the smallest number of genes for slide comparisons, while Benjamini-Hochberg FDR corrections having the least number of genes produced by MAS5, VSN and GCRMA (*Figure 5.5*). In the 12 slide experiment, Bonferroni also seemed to lower the number of genes deemed regulated by PLM method, thereby resulting in the RMA method usually producing the most genes in slide comparisons (*Figure 5.5*), but the Benjamini-Hochberg correction did not

(*Figure 5.5*). Again, the MAS5 normalization method usually resulted in the smallest number of gene deemed regulated for 10 out of the 12 overall comparisons (*Figure 5.5*).

False discovery rate (FDR) and FWER correction methods, like Benjamini-Hochberg's and Bonferroni's, tries to lower the number of false positives within predictions of differentially regulated geneIDs/probe sets. From the above results it is clear that normalization methods influence which probe sets will eventually be predicted as differentially expressed (*Figures 5.5* and *5.7*). Furthermore, the probe set eventually identified is dependent on the inherited statistical biases of the normalization method, which may increase as confidence levels increases. These underlying normalization method biases cannot be nullified by FDR/FWER correction methods: Except for Bonferroni at 95 and 99 % confidences, the percentage of probe sets contributed by a single method, increases – especially for Benjamini-Hochberg FDR corrections (*Figure 5.7*). Similarly, the number of probe sets found in all five methods, tend to decrease (*Figure 5.7*). However, Bonferroni FWER correction increased the number of probe sets identified as differentially regulated for the "4 out of 5 normalization methods" class (*Figure 5.7*), though Benjamini-Hochberg's approach didn't when compared with the non-FDR entries. Though the two correction methods tested here usually resulted in a decrease in the number of predicted probe sets (*Figure 5.7*), the same shifting tendency is observed when confidence levels are increased, *i.e.* towards a single geneID-single normalization method, is seen here. The exception, *i.e.* the increased number of probe sets associated with the Bonferroni "4 out of 5 normalization method" class, suggests that this method either decrease false negatives or decrease false positives of a single or multiple methods, thereby increasing the number of probe sets occurring in 4 out of the 5 methods. However, this correction method is very strict and rejects many probe sets, *i.e.* lower geneID numbers (*Figure 5.7*). This suggests that only the highly significant probe sets are retained in each method, thereby increasing the chance of getting the same probe set. A bias in one or more methods probably prevents this FWER correction method from shifting probe sets to the "all 5 methods" class.

New introductions of normalization methods always seem to be better than the last (Irizarray *et al.* 2006). Matters are further complicated by different comparison studies suggesting different normalization method based on different testing approaches; from spike-in data sets (Irizarray *et al.* 2003) to sex linked internal control genes (Galfalvy *et al.* 2003) and co-expressed operon linked control genes (Harr & Schlötterer 2006). Control calibrations/statistics, however, do not consistently produce the same results when using algorithms on experimental datasets (Shedden *et al.* 2005). Indeed, normalization method influence has previous been reported for different study aims, *i.e.*

different methods produced better results for a data set depending on the aim - the identification of differentially expressed genes or co-expressed genes (Harr & Schlötterer 2006; Lim *et al.* 2007). The results presented here suggest that though normalization and background correction methods may produce apparently adequate corrections, they do play a major role in the final geneIDs/probe sets and number of transcripts found to be differentially regulated, as was previously shown (Hoffmann *et al.* 2002; Shedden *et al.* 2005). However, here we further showed that by increasing the confidence levels for selecting differentially regulated probe sets or using FDR/FWER correction methods, the percentage bias of the normalization method initially used may actually increase, while the number of probe sets maybe lowered. How can results be trusted if 20-65 % (*Figure 5.7*) of the geneIDs identified only occurs as the result off one or two normalization methods initially selected? Although probe sets seem to be influenced more by the normalization method initially used, the results highlights the need for using a multiple normalization approach, *i.e.* an approach where the initial normalization and background corrections are done by various methods suitable to the specific experimental design. The various normalized datasets are then carried through the whole process of differentially gene identification, until the resulting probe sets can be compared. This can easily be done using the scripts developed here. If a probe set then occurs in 4 or 5 (or even 3) out of 5 normalized datasets (with/ without FDR/FWER correction), it is more likely to be differentially regulated rather than an introduced normalization method bias. This approach should identify probe sets unique to a specific biological interaction, which is normalization method independent and suitable for marker development.

Acknowledgements

We would like to thank Prof F. Joubert and the Bioinformatics Unit (at the University of Pretoria) for help during the execution of the scripts on their servers, Prof N.L.V. Lapitan (Colorado State University) for collaborating with us on the USA RWA biotypes and Dr. V. Tolmay (Small Grains Institute) for the SA RWA biotype. This research was funded by the Winter Cereal Trust and THRIP (Technology and Human Resources for Industry Programme, South Africa).

References

Affymetrix (2002) Statistical algorithms description document. *Technical report*. Accessed 2009.

Benjamini, Y. and Y. Hochberg (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)* **57**: 289-300.

Bolstad, B. M. (2007) affyPLM: Model Based QC Assessment of Affymetrix GeneChips.

Bolstad, B. M., R. A. Irizarry, M. Astrand and T. P. Speed (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**: 185-193.

Botha, A.-M., M. A. C. Nagel, A. J. van der Westhuizen and F. C. Botha (1998) Chitinase isoenzymes in near-isogenic wheat lines challenged with Russian wheat aphid, exogenous ethylene, and mechanical wounding. *Botanical Bulletin of Academia Sinica* **39**: 99-106.

Botha, A.-M., Z.H. Swanevelder and N.L.V. Lapitan (2010) Transcript profiling of wheat genes expressed during feeding by two different biotypes of *Diuraphis noxia*. *Environmental Entomology: in press.*

Du Toit, F. (1989) Inheritance of resistance in two *Triticum aestivum* lines to Russian wheat aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **82**: 1251-1253.

Dudoit, S. and Y. Ge (2004) Bioconductor's multtest package. *faculty.mssm.edu/gey01/multtest/multtest.pdf*.

Elo, L. L., J. Hiissa, J. Tuimala, A. Kallio, E. Korpelainen and T. Aittokallio (2009) Optimized detection of differential expression in global profiling experiments: case studies in clinical transcriptomic and quantitative proteomic datasets. *Briefings in Bioinformatics* **10**: 547-555.

Galfalvy, H. C., L. Erraji-Benchekroun, P. Smyrniotopoulos, P. Pavlidis, S. P. Ellis, J. J. Mann, E. Sibille and V. Arango (2003) Sex genes for genomic analysis in human brain: internal controls for comparison of probe level data extraction. *BMC Bioinformatics* **4**.

Gautier, L., L. Cope, B. M. Bolstad and R. A. Irizarry (2004) Affy - analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**: 307-315.

Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dutoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. A. Irizarray, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L.

Tierney, J. Y. H. Yang and J. Zhang (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5**: R80.

Haley, S. D., F. B. Peairs, C. B. Walker, J. B. Rudolph and T. L. Randolph (2004) Occurrence of a new Russian wheat aphid biotype in Colorado. *Crop Science* **44**: 1589-1592.

Harr, B. and Schlötterer (2006) Comparison of algorithms for the analysis of Affymetrix microarray data as evaluated by co-expression of genes in known operons. *Nucleic Acids Reseach* **34**: e8.

Hoffmann, R., T. Seidl and M. Dugas (2002) Profound effect of normalization on detection of differentially expressed genes in oligonucleotide microarray analysis. *Genome Biology* **3**: research0033.1-0033.11.

Huber, W., A. von Heydebreck, H. Sültmann, A. Poustka and M. Vingron (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18**: s96-s104.

Ihaka, R. and R. Gentleman (1996) R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**: 299-314.

Irizarray, R. A., B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs and T. P. Speed (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Reseach* **31**: e15.

Irizarray, R. A., Z. Wu and H. A. Jaffee (2006) Comparison of Affymetrix GeneChip expression measures. *Bioinformatics* **22**: 789-794.

Kriel, D. P. and R. R. Russell (2005) There is no silver bullet - a guide to low-level data transforms and normalisation methods for microarray data. *Briefings in Bioinformatics* **6**: 86-97.

Lapitan, N. L. V., Y.-C. Li, J. Peng, F. B. Peairs and A.-M. Botha (2007) Limited nuclear and mitochondrial DNA variation among Russian wheat aphid (*Diuraphis noxia*) biotypes from the United Satates and Africa. *Submitted*.

Lim, W. K., K. Wang, C. Lefebvre and A. Califano (2007) Comparative analysis of microarray normalization procedures: effects on reverse engineering gene networks. *Bioinformatics* **23**: i282-i288.

Shedden, K., W. Chen, R. Kuick, D. Ghosh, J. Macdonald, K. R. Cho, T. J. Giordano, S. B. Gruber, E. R. Fearon, J. M. G. Taylor and S. Hanash (2005) Comparison of seven methods for producing Affymetrix expression scores based on False Discovery Rates in disease profiling data. *BMC Bioinfomatics* **6**: 1-12.

Smyth, G. K. (2004) Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Statistical Applications in Genetics and Molecular Biology* **3**: Article3.

Smyth, G. K. and T. P. Speed (2003) Limma: linear models for microarray data. Bioinformatics and computational biology solutions using R and Bioconductor. R. Gentleman, V. Carey, S. Dudoit, R. Irizarry and W. Huber. New York, Springer**:** 397-420.

Steinhoff, C. and M. Vingron (2006) Normalization and quantification of differential expression in gene expression microarrays. *Briefings in Bioinformatics* **7**: 166-177.

Zakharkin, S. O., K. Kim, T. Mehta, L. Chen, S. Barnes, K. E. Scheirer, R. S. Parrish, D. B. Allison and G. P. Page (2005) Sources of variation in Affymetrix microarray experiments. *BMC Bioinfomatics* **6**: 1-11.

CHAPTER 6

CONCLUSION

The aphid-plant interaction, especially between *Diuraphis noxia* and wheat, has been the subject of numerous studies. In these, the focus was usually on plant defense: either how different wheat cultivars' respond to RWA feeding or how different aphid biotypes infestation profiles differ based on resistant plant differentials. The theory that resistant cultivars forced new RWA biotypes, *a.k.a.* the arms race, is widely accepted. However, aphids' successes as plant sap feeders are directly contributed to their endosymbiont. This would suggest that any selective pressure on the aphid would be on the endosymbiont as well. This in turn implies that any endosymbiont changes beneficial to its aphid host may result in more adaptive RWA biotypes.

The main aim of this study was to look at the role, if any, of the RWA endosymbiont, *Buchnera aphidicola*, on the RWA-wheat interaction. It was established that no secondary symbionts were present in the ten RWA biotypes investigated and that all the aphid biotypes only had their primary endosymbiont, *B. aphidicola*. Previous studies showed that this endosymbiont played a crucial role in aphids' ability to survive on plantsap by increasing essential amino acids that occurs in limited amounts within it. However, other reports on the RWA-*B. aphidicola* interaction suggested the endosymbiont played a decreasing role in RWA fitness in that it produced less essential amino acids. This hypothesis was further supported by studies that showed that the RWA had the ability to alter amino acid composition of susceptible wheat cultivars by increasing those essential amino acids previously supplied by its endosymbiont. However, in these studies, the RWA could not alter the amino acid composition of resistant cultivars. The results presented here suggest that *B. aphidicola* may still play a role in aphid fitness. Firstly, copy numbers of an essential amino acid producing plasmid varied amongst the different *B. aphidicola* originating from the different RWA biotypes. Secondly, though little sequence variation was found amongst these endosymbionts, the single CCC-insert difference did have a role in increasing the expression levels of the subsequent genes when compared to their plasmid copy numbers. Furthermore, the CCC-insert also increased the leader sequence length, thus resulting in more stable 5' UTRs. These findings, together with similar variations identified in this study within other aphid species, suggest a regulatory function for this region. However, they also suggest that *B. aphidicola* may play a role in RWA fitness.

Continuous pressure of low concentrations of essential amino acids in resistant wheat cultivars, selected for aphids that can survive under these limiting conditions. Normally, the RWA would increase these essential amino acids in its diet while feeding. Failing to do so would severely limit or even kill a RWA population. Another way of dealing with this problem is through natural selection of RWA individuals that can upgrade their diet, *i.e.* aphids that have *B. aphidicola* that

supply it with the required amino acids. This can be done by increasing the output of the essential amino acid biosynthesising pathways, either via increased copy numbers or through higher expression levels of these genes. Results in the present study showed that there were differences in an essential amino acid plasmid's copy numbers – to date it is believed to be the only way that the *B. aphidicola* can regulate its gene expression. However, a mutation preceding the leader sequence of a leucine plasmid gene led to differences in transcript levels. This suggests another method through which *B. aphidicola* could alter transcript levels and thus play a role in its host's fitness. Indeed, these changes in plasmid expression in the endosymbiont may potentially result in the development of new biotypes, especially since RWA biotypes are currently distinguished from each other based on the different resistant cultivars they can feed upon. Alternatively, *B. aphidicola* could at the very least allow the RWA populations to survive on resistant wheat cultivars until they've adapted to feeding on them.

The second part of the study looked at the influences of statistical normalization methods on the identification of differentially regulated probe sets, when RWA-plant interactions are investigated with Affymetrix GeneChip technology. The hypothesis of this section stated that there should be a consistent subset of regulated probe sets/geneIDs identified as differentially regulated, regardless of the initial normalization method employed, as long as the rest of the analyses were constant, and that this subset would be normalization-method-independent. Two Affymetrix RWA-wheat interaction datasets, that included the interactions of two South African and two US RWA biotypes with a total of six wheat accessions, were investigated using five different normalization and background correction methods (RMA, GCRMA, MAS5.0, PLM, VSN) at three different confidence levels (95, 99, 99.9 %), in the presence and absence of FDR/FWER correction.

The results illustrated that genes identified as differentially expressed were highly dependent on the specific normalization and background correction method employed. It was also shown that normalization and background correction method dependent biases could not be nullified by increases in confidence levels, but that this actually tends to increase these biases. Furthermore, FDR and FWER detection methods usually add to the normalization and background correction bias, except in normalization approaches that are very strict in their selection criteria. This supports a hypothesis that the genes identified as differentially regulated depend on the inherited statistical biases of the background correction and normalization methods employed, and that both FDR/FWER and increased confidence levels could actually further enhance this problem. A multiple normalization approach on the initial data is proposed to enable the identification of a probe

set subset that is normalization-method-biases-independent. These genes/probe sets should not be influenced by any of the different normalization method biases, and would therefore reflect the true biological differentially regulated genes.

APPENDICES AND SUPPLEMENTARY DATA

APPENDIX – CHAPTER 3

Table Appx 3.1 Primer pairs, targets and PCR conditions used in the different analyses.

Name	***Primer sequence (5' – 3')****	***Target region***	***Annealing Temperature***	***Amplicon length (bp)***
16S DGGE				
PRUN518r[#]	ATTACCGCGGCTGCTGG	*16S rDNA*	58 ºC	460
pA8f-GC[§]	AGAGTTTGATCCTGGCTCAG			
Standards and nested primers: copy numbers				
leuB_5069F	CAAGAGGGCGCGTAACTAAA	*leuB*	60 ºC	411
leuB_5461R	AACCCCCTGCTGGTTCATA			
leuB_5158R	GAGCGCGCTAATTTAAAAGC		60 ºC	109
trpB_3248F	TACTGCAGCAGGACCACATC	*trpB*	59 ºC	375
trpB_3622R	GCATGTTCTGGACCAACAGA			
trpB_3380R	TCCTCCAACACATGCGATTA		60 ºC	133
***Sequencing:* pleuABCD**				
leuC_6856F	GCTATTGTTGTACCAGGTTCAGG	6856-969	TD[d]	1881
repA1_969R	ACACAGAGCATCAGCCATGA			
repA1_894F	TGTTGATAGACCACATTCATCAGA	894-1830	TD[d]	936
CDS1_1855R	CGTAAAAGATGTACAGAAGATATTGC			
CDS1_1521F	TCTCGTTGGTTRATGTTTCC	1521-2501	TD[db]	982
repA2_2501R	TAAACGTGAAACACGGGTGA			

leuA_3289R	GAAGACATTGCATCACCTGCT	1540-3289	TD[d]	1749
CDS1_1521F	TCTCGTTGGTTRATGTTTCC			
leuA_3077F	CGTGATGGTGAACAAGCATT	3077-4145	TD[d]	1068
leuA_4145R	GCCTCTAAATCATAATCAAACACTTG			
leuA_4027F	TTCACGATCAGGAAGAGCTG	4027-6087	TD[d]	2060
leuC_6087R	TGGGAATCACCACACACAAT			
leuB_5069F	CAAGAGGGCGCGTAACTAAA	5088-6087	TD[d]	999
leuC_6087R	TGGGAATCACCACACACAAT			
leuC_5957F	CATCAGGTTCAATGGCAAAA	5957-6918	TD[db]	961
leuC_6918R	CAACCAGGTAAACGCCATTC			
leuC_6856F	GCTATTGTTGTACCAGGTTCAGG	6856-147	TD[db]	1059
spacer_147R	TGAATTAAACTTTTAAATGCATGTTGT			
leuB_5236F	TGGAGAGAAGTGGTTGAAGAGG	1536-6087	TD[d]	834
leuC_6087R	TGGGAATCACCACACACAAT			
leuA_4027F	TTCACGATCAGGAAGAGCTG	4046-5169	TD[d]	1123
leuB_5461R	AACCCCCTGCTGGTTCATA			
leuC_6570F	TTACCTCAAGTTACTTGGGG		TD[db]	1326
spacer_147R	TGAATTAAACTTTTAAATGCATGTTGT	6589-147		
leuA_3876F	TGCCTATACCTGCTAATAAAGC	3897-5158	TD[db]	1261
leuB_5158R	GAGCGCGCTAATTTAAAAGC			
repA2_2354F	ATGAACATCGTGCGTGTGC	2372-3289	TD[db]	917
leuA_3289R	GAAGACATTGCATCACCTGCT			

***Sequencing:* ptrpEG**				
ptrpE_3965F	CCCCTCCTTGATCTCCTACA	7179-1651	TD[φ]	3506
trpE_1674R	CTGAAAATGAGGGAATTCTAAACG			
ptrpE_2260F	TTTGGTGCTTCACCAGAAAG	8115-61	TD[φ]	980
ptrpG_3273R	CTCCAAACCGTCATGATTGA			
ptrpG_3194F	AAGCTTATGGTGGC	14-676	TD[φ]	662
spacer_695R	TCGTTTGGCGACTCATCATA			
spacer_620F	GATCCTGCGCACTCTCAATAG	20-1651	TD[φ]	1631
trpE_1674R	CTGAAAATGAGGGAATTCTAAACG			
trpE_1542F	TCACTCGGTACAACTAACAGCA	1563-2338	TD[d]	775
ptrpE_2357R	ATGTTCCTCCCTCTGGGTCT			

*All primers were designed from Genbank sequences (www.ncbi.nlm.nih.gov) using *Primer3* (Rozen & Skaletsky 2000) and on sequences generated in this study. Total DNA (25 ng) in 25 µL reaction volumes was used for all PCRs, with 0.5 U ExSel High Fidelity DNA polymerase, 1 × reaction buffer with $MgSO_4$ (Southern Cross Biotechnology), 0.4 µM of each primer and 100 µM of each dNTP. GeneAmp 9700 thermocyclers (Applied Biosystems) were used with cycling conditions preceded by an initial denaturation step of 95 ºC for 3 min and followed by a final elongation step of 72 ºC for 10 min. 30 Cycles consisted of 30 sec at 94 ºC, 30 sec at primer annealing temperature and 1 min/1 Kb extension period at 72 ºC. These conditions were standard for all PCR reactions unless otherwise stated. For touchdown PCR: the number of cycles depended on the touchdown range with a cycle increase for each degree change. Normal PCR cycling followed at the lower touchdown temperature with extension times dependent on fragment length (1 min/kb) and a final elongation step at 72 ºC for 10 min. Large fragments were first amplified form the *ptrpEG* plasmid before internal primers were used for sequencing. This was done due to the high homology between the functional and pseudogenes residing on the plasmid. [#] Sequencing primer (Muyzer *et al.* 1993); [§] pA8f-GC contained a 5' 40mer GC-clamp (Fjellbirkeland *et al.* 2001); [d] TD - Touchdown 60-55 ºC (Don *et al.* 1991); [ф] TD - Touchdown 55-50 ºC (Don *et al.* 1991); [φ] TD -Touchdown 60-50 ºC (Don *et al.* 1991).

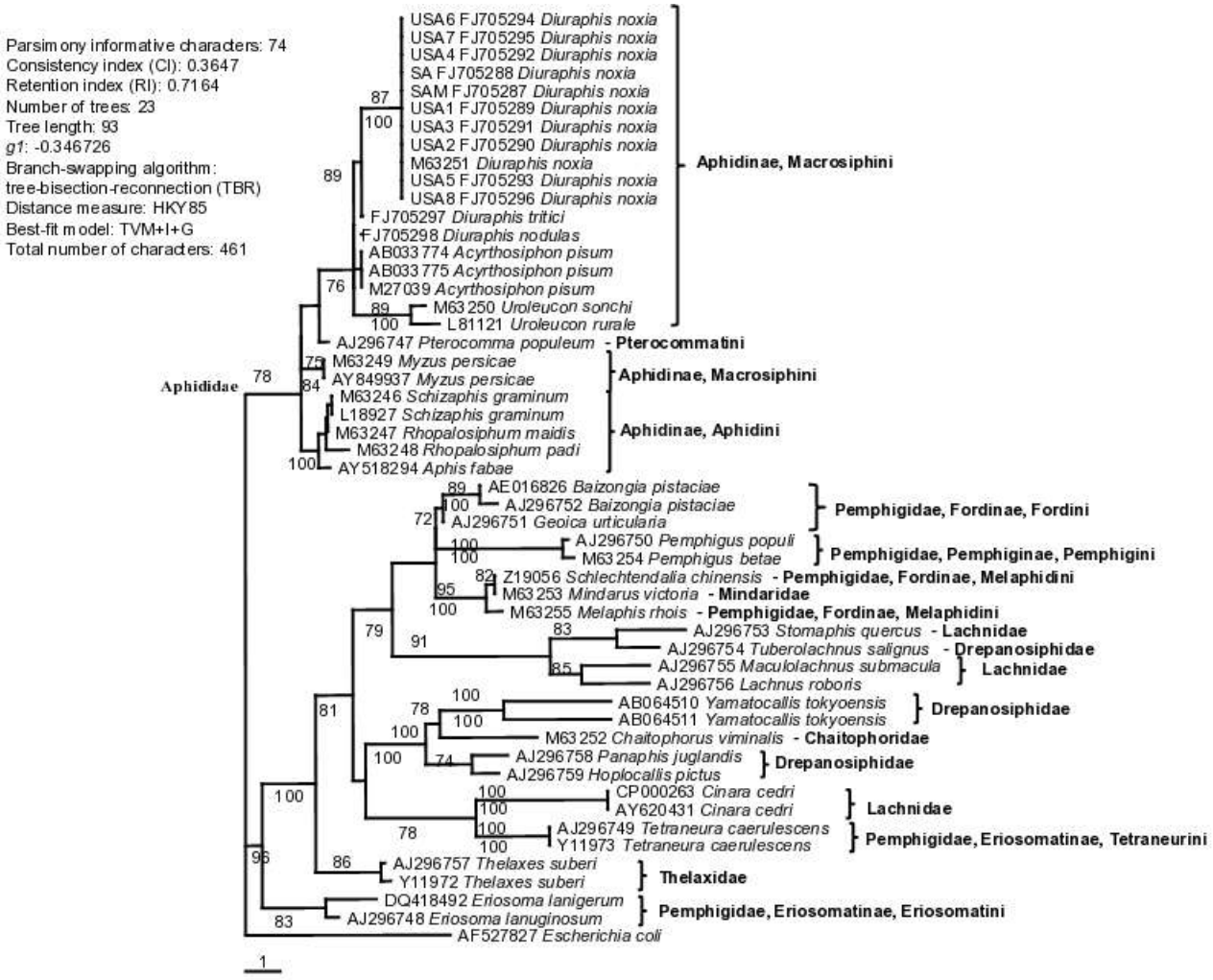


Figure Appx 3.1 Phylogeny of the biotypes to their *16S rDNA* DGGE BLAST results (Genbank). Parsimony topology and branch lengths are shown. Bootstrap percentages from parsimony (70 % and greater, 1000 bootstrap replications) are shown above and Bayesian posterior probability values below branches. The third heuristic tree (out of 23) is presented here. The phylogram had a tree length of 93 and was based on 74 parsimony informative characters with a consistency index of 0.3647 and a retention index of 0.7164. *Escherichia coli* was used as outgroup to root the tree. Branch labels show the Genbank accession, followed by the aphid species name except for the outgroup. Tribes and families for the hosts obtained from Genbank accessions were indicated for the Aphididae. All phylogenetic analyses were done with PAUP 4.0b10 (Swofford 2002) and Bayesian analysis with MrBayes 3.1.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Multiple sequence alignments were done with ClustalW version 2 (Larkin *et al.* 2007) and manually evaluated before further analysis. Ambiguous characters and uninformative nucleotides were

excluded from data prior to analysis and all characters were re-weighted to the consistency index. The best statistical model of DNA evolution for each dataset was determined using modeltest 3.7 (Posada & Crandall 1998) and used in subsequent analysis. Heuristic searches using random sequence additions were performed with the tree-bisection-reconnection (TBR) branch-swapping algorithm and MaxTrees set to auto increase. Phylogenetic signal, *i.e.* consistency index (CI) and retention index (RI), was assessed by evaluating the tree length distributions in each dataset after 100 random generated trees. Only groups with a 70 % or more support were retained in bootstrap analyses over a 1000 replicates. MrBayes utilizes a Metropolis-coupled Markov Chain Monte Carlo search algorithm (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). The general time reversal (GTR) evolutionary model was selected with codon site-specific rate variation. One million generations were run for each dataset, using one cold and 3 heated chains. Bayesian inference posterior probabilities were calculated after the appropriate burn in was determined.

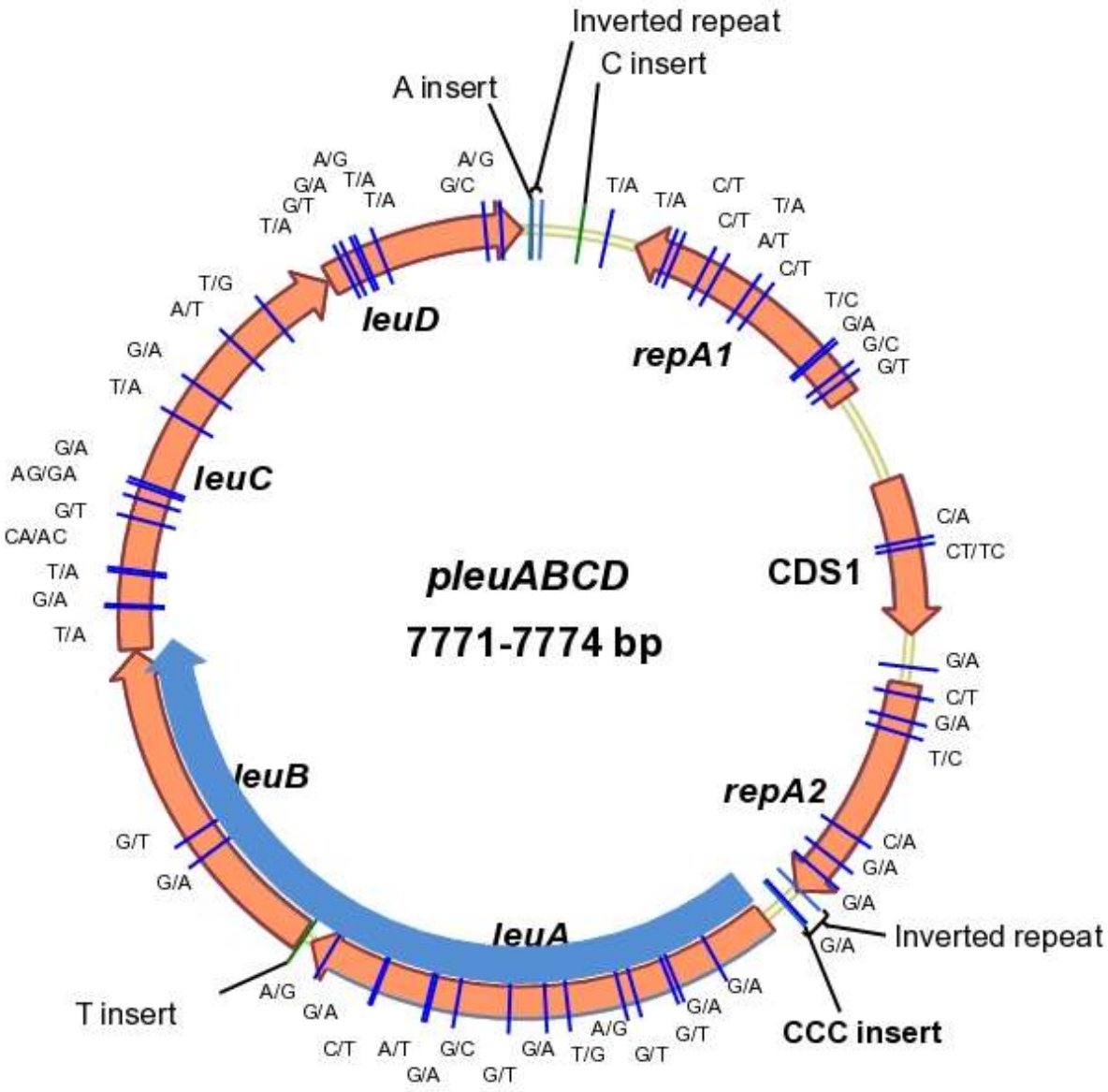


Figure Appx 3.2 Differences between the biotypes (FJ705299-FJ705308) and the Genbank sequences (AF041837, NC001911) for the *pleuABCD* plasmid. The CCC-insert only occurred in the biotypes SA, SAM, USA3 and USA7. A single nucleotide insert, a T-insert, between *leuA* and *leuB* genes, changed *leuB* into the same ORF as the *leuA* gene. The new *leuA-leuB* ORF is indicated in blue. The mutations notation used here indicate the nucleotide of the original Genbank sequences, a slash, followed by the changed nucleotide from sequencing results.

Sequences submitted to Genbank

FJ705277 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnSAM tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaattttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705278 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnSA tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaattttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705279 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnUSA1 tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaattttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705280 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnUSA2 tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaattttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705281 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnUSA3 tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaattttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705282 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnUSA4 tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaattttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat

aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705283 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnUSA5 tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaatttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705284 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnUSA6 tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaatttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705285 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnUSA7 tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaatttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705286 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUHtrpBDnUSA8 tryptophan synthase large subunit (trpB) gene, partial cds.
ctatcctactattgttcgagaatttcaaaaaattattggagaagaaactaaacaacaaatttagaacaagaaacaaaattaccaaatgcaataatcg
catgtgttggaggaggttctaatgcaattggtatttttttcaaattttattaatgataaagaagtaagtttaattggtgttgaacccgggggggaagggtat
aaaaacaggtcaacatggtgcaccattaaaacatggtagaactggtatttttttcggaatgaaatctcatttaatgcaagatcaagaaggtcaaattc
aagaatcttggtctatttctgcaggattagactttccatctgttgg

FJ705287 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnSAM 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705288 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnSA 16S ribosomal RNA gene, partial sequence.

attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705289 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnUSA1 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705290 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnUSA2 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705291 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnUSA3 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705292 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnUSA4 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705293 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnUSA5 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc

tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705294 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnUSA6 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705295 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnUSA7 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705296 *Buchnera aphidicola* (*Diuraphis noxia*) clone BUH16SDnUSA8 16S ribosomal RNA gene, partial sequence.
attgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaacgggtgagtaa
tatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtgggggaccttttggcc
tcatgcttatggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagaggataaccagc
cacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgcagctatgccgc
gtgtatgaagaaggccttagggttgtaaagtactttcagcggggaagaaaaaaataaaactaataatttattcctgacgtacc

FJ705297 *Buchnera aphidicola* (*Diuraphis tritici*) clone BUH16SDtri 16S ribosomal RNA gene, partial sequence.
cctggctcagattgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaac
gggtgagtaatatctggggatctgcccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtggggg
acctttttggcctcatgcttttggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcgacgatctctagctggtctgagag
gataaccagccacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgca
gctatgccgcgtgtatgaagaaggccttagggttgtaaagtactttcagcggggaggaaaaaaataaaactaataatttattacgtgacgt

FJ705298 *Buchnera aphidicola* (*Diuraphis nodulas*) clone BUH16SDnod 16S ribosomal RNA gene, partial sequence.
cctggctcagattgaacgctggcggcaagcctaacacatgcaagtcgagcggcagcgaaagaaagcttgctttcttgtcggcgagcggcaaac
gggtgagtaatatctggggatctacccaaaagagggggataactactagaaatggtagctaataccgcataatgttgaaaaaccaaagtggggg
acctttttggcctcatgcttttggatgaacccagacgagattagcttgttggtagagtaatagcctaccaaggcaacgatctctagctggtctgagag

gataaccagccacactggaactgagacacggtccagactcctacgggaggcagcagtggggaatattgcacaatgggcgaaagcctgatgca
gctatgtcgcgtgtatgaagaaggccttagggttgtaaagtactttcagcggggaggaaaaaaataaaactaata

FJ705299 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(SA), complete sequence.
ttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaattttttttttaaagttttt
ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaatttttttcagaaatatctaatagcataaaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgtttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttctttgctttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttattcataacaaaatctatatgaagtatt
cattgtactaaaaatttttatatataaaatatatactttacttttatttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaattttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaattttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttat
agtgagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaatt
acaaatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatca
aaaaaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaa
ttcacatttttttagctacttcaacacttcatatggaatctaaattaagaaagaattttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcatta
cgttatactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtga
aaactattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctat
aatttctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatggg
atgggagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataa
agaaatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggt

attcatcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgat
caggaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcaga
taaaaaaggacaagtgtttgattatgatttagaagctttagcatttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttagtgtgcaa
tctatttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctattt
atcaagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagata
ttttagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatg
gaaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatatt
gctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaattttacaggttttaagagaacatttttcattatttataaaaacaaa
agaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattagg
atctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaat
ttaagaccatctaatttatataaggaattaaatttttttatcacccttacgtaatgatattgtaaaacatggtttaatatattatgtgttagagaattaactgg
aggaatttattttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattg
ctcatttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctattttatggagagaagtggttg
aagaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgt
gttctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggttt
atatgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagt
atgaatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattattta
aaaacaaatgagatgggtgatgttattgctgatttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgt
acattctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgt
acagttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaac
aaatgcaaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacag
ggtatgactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagta
gaacatgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaagga
tgttatcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaa
cgaatgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttatt
ctcctcatggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatc
aaatattttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaa
agatttagctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgct
cgaatagaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaa
gacaagctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggt
tatcagaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcag
cattagctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgt
agttgctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttca
tgattggcgttatcttgattctaaccaattgaaaaaaaataataaatttattttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatt
ttggttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagtt
ttaataataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttacta
actaataaagttattatcaatcaaaaaactttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaatt
ctttaaaagaaataaatgtatatgaaaataatatcccatttttttctttgaatagaaaagaatttaaatctcattaa

FJ705300 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(SAM), complete sequence.
ttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaaactttaaaaaaaaaatttttttttaaagttttt

ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaatttttttcagaaatatctaatagcataaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgtttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttctttgctttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt
cattgtactaaaaatttttatatataaaatatatactttacttttatttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaattttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaatttttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttat
agtgagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaatt
acaaatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatca
aaaaaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaa
ttcacatttttttagctacttcaacacttcatatggaatctaaattaagaaagaattttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcatta
cgttatactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtga
aaactattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctat
aatttctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatggg
atgggagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataa
agaaatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggt
attcatcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgat
caggaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcaga
taaaaaaggacaagtgtttgattatgatttagaagctttagcatttttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttagtgtgcaa
tctatttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctattt
atcaagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagata
ttttagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatg

gaaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatatt
gctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaattttacaggttttaagagaacatttttcattatttataaaaacaaa
agaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattagg
atctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaat
ttaagaccatctaatttatataaggaattaaatttttttatcacccttacgtaatgatattgtaaaacatggtttaatatattatgtgttagagaattaactgg
aggaatttattttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattg
ctcatttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctattttatggagagaagtggttg
aagaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgt
gttctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggttt
atatgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagt
atgaatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattattta
aaaacaaatgagatgggtgatgttattgctgatttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgt
acattctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgt
acagttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaac
aaatgcaaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacag
ggtatgactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagta
gaacatgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaagga
tgttatcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaa
cgaatgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttatt
ctcctcatggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatc
aaatattttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaa
agatttagctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgct
cgaatagaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaa
gacaagctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggt
tatcagaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcag
cattagctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgt
agttgctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttca
tgattggcgttatcttgattctaaccaattgaaaaaaaataataaatttattttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatt
ttggttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagtt
ttaataataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttacta
actaataaagttattatcaatcaaaaaactttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaatt
ctttaaaagaaataaatgtatatgaaaataatatcccatttttttctttgaatagaaaagaatttaaatctcattaa

FJ705301 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(USA1), complete sequence.

ttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaattttttttttaaagttttt
ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaatttttttcagaaatatctaatagcataaaaaataatggagttaattta

atcattttaggcatataattacctaatattttatcccatatttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgtttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttcttttgctttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt
cattgtactaaaaatttttatatataaaatatatactttacttttatttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaattttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttagaaccaatgggtttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaatttttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttatagt
gagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaattaca
aatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatcaaaa
aaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaattca
catttttttagctacttcaacacttcatatggaatctaaattaagaaagaatttttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcattacgtt
atactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtgaaaac
tattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctataattt
ctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatgggatgg
gagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataaaga
aatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggtattc
atcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgatcag
gaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcagataa
aaaaggacaagtgtttgattatgatttagaagctttagcatttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttagtgtgcaatct
atttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctatttat
caagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagatattt
tagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatgg
aaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatattg
ctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaattttacaggttttaagagaacattttcattatttataaaaacaaaa
gaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattaggat
ctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaattt
aagaccatctaatttatataaggaattaaatttttatcacccttacgtaatgatattgtaaaacatggttttaatatattatgtgttagagaattaactggag

gaatttattttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattgctc
atttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctatttttatggagagaagtggttgaa
gaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgtgtt
ctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggtttata
tgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagtatg
aatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattatttaaaa
acaaatgagatgggtgatgttattgctgatttttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgtacat
tctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgtacag
ttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaacaaatg
caaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacagggtat
gactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagtagaac
atgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaaggatgtta
tcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaacgaa
tgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttattctcct
catggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatcaaata
ttttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaaagattt
agctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgctcgaat
agaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaagacaa
gctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggttatca
gaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcagcatta
gctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgtagttg
ctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttcatgatt
ggcgttatcttgattctaaccaattgaaaaaaaataataaatttatttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatttgg
ttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagttttaat
aataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttactaactaa
taaagttattatcaatcaaaaaacttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaattctttaa
aagaaataaatgtatatgaaaataatatcccatttttctttgaatagaaaagaatttaaatctcattaa

FJ705302 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(USA2), complete sequence.

ttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaatttttttttaaagttttt
tttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaattttttttcagaaatatctaatagcataaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgttttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttcttttgctttttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt

cattgtactaaaaatttttatatataaaatatatactttacttttatttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaattttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaatttttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttatagt
gagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaattaca
aatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatcaaaa
aaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaattca
catttttttagctacttcaacacttcatatggaatctaaattaagaaagaattttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcattacgtt
atactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtgaaaac
tattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctataattt
ctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatgggatgg
gagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataaaga
aatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggtattc
atcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgatcag
gaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcagataa
aaaaggacaagtgtttgattatgatttagaagctttagcatttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttagtgtgcaatct
atttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctatttat
caagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagatattt
tagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatgg
aaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatattg
ctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaatttacaggttttaagagaacatttttcattatttataaaaacaaaa
gaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattaggat
ctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaattt
aagaccatctaatttatataaggaattaaatttttatcacccttacgtaatgatattgtaaaacatggttttaatatattatgtgttagagaattaactggag
gaatttattttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattgctc
atttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctattttatggagagaagtggttgaa
gaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgtgtt
ctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggtttata
tgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagtatg

aatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattatttaaaa
acaaatgagatgggtgatgttattgctgatttttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgtacat
tctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgtacag
ttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaacaaatg
caaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacagggtat
gactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagtagaac
atgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaaggatgtta
tcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaacgaa
tgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttattctcct
catggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatcaaata
ttttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaaagattt
agctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgctcgaat
agaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaagacaa
gctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggttatca
gaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcagcatta
gctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgtagttg
ctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttcatgatt
ggcgttatcttgattctaaccaattgaaaaaaaataataaatttattttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaattttgg
ttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagttttaat
aataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttactaactaa
taaagttattatcaatcaaaaaactttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaattctttaa
aagaaataaatgtatatgaaaataatatcccatttttctttgaatagaaaagaatttaaatctcattaa

FJ705303 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(USA3), complete sequence.

ttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaattttttttttaaagttttt
ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcattttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaatttttttcagaaatatctaatagcataaaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgtttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttctttgctttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt
cattgtactaaaaattttttatatataaaatatatactttacttttattttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaattttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt

aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaatttttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttat
agtgagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaatt
acaaatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatca
aaaaaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaa
ttcacatttttttagctacttcaacacttcatatggaatctaaattaagaaagaattttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcatta
cgttatactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtga
aaactattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctat
aatttctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatggg
atgggagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataa
agaaatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggt
attcatcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgat
caggaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcaga
taaaaaaggacaagtgtttgattatgatttagaagctttagcatttttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttttagtgtgcaa
tctatttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctattt
atcaagcattaaataaaattataaatttttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagata
ttttagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatg
gaaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttttatgcataaacaatatcatatt
gctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaatttacaggttttaagagaacatttttcattatttataaaaacaaa
agaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgatttttattagg
atctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaat
ttaagaccatctaatttatataaggaattaaatttttttatcacccttacgtaatgatattgtaaaacatggttttaatatattatgtgttagagaattaactgg
aggaatttattttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattg
ctcatttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctatttttatggagagaagtggttg
aagaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgt
gttctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggttt
atatgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagt
atgaatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattattta
aaaacaaatgagatgggtgatgttattgctgatttttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgt
acattctgaaaaaaatggtttatctatttttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgt
acagttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaac
aaatgcaaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacag

ggtatgactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagta
gaacatgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaagga
tgttatcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaa
cgaatgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttatt
ctcctcatggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatc
aaatattttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaa
agatttagctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgct
cgaatagaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaa
gacaagctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggt
tatcagaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcag
cattagctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgt
agttgctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttca
tgattggcgttatcttgattctaaccaattgaaaaaaaataataaatttatttttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatt
ttggttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagtt
ttaataataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttacta
actaataaagttattatcaatcaaaaaactttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaatt
ctttaaaagaaataaatgtatatgaaaataatatcccatttttttctttgaatagaaaagaatttaaatctcattaa

FJ705304 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(USA4), complete sequence.

ttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaatttttttttaaagttttt
ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaatttttttcagaaatatctaatagcataaaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgttttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttcttttgctttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt
cattgtactaaaaattttatatataaaatatatactttacttttattttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttaaaaaagtttaataaaaaattaatttaggga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaattttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata

gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaaattttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttatagt
gagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaattaca
aatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatcaaaa
aaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaattca
catttttttagctacttcaacacttcatatggaatctaaattaagaaagaattttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcattacgtt
atactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtgaaaac
tattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctataattt
ctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatgggatgg
gagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataaaga
aatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggtattc
atcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgatcag
gaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcagataa
aaaaggacaagtgtttgattatgatttagaagctttagcatttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttagtgtgcaatct
atttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctatttat
caagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagatattt
tagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatgg
aaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatattg
ctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaattttacaggttttaagagaacatttttcattatttataaaaacaaaa
gaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattaggat
ctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaattt
aagaccatctaatttatataaggaattaaatttttttatcacccttacgtaatgatattgtaaaacatggttttaatatattatgtgttagagaattaactggag
gaatttatttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattgctc
atttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctattttatggagagaagtggttgaa
gaggtttctaaggaatatcctgatgttatttttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgtgtt
ctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggtttata
tgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagtatg
aatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattatttaaaa
acaaatgagatgggtgatgttattgctgatttttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgtacat
tctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgtacag
ttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaacaaatg
caaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacagggtat
gactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagtagaac
atgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaaggatgtta
tcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaacgaa
tgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttattctcct
catggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatcaaata

ttttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaaagattt
agctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgctcgaat
agaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaagacaa
gctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggttatca
gaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcagcatta
gctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgtagttg
ctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttcatgatt
ggcgttatcttgattctaaccaattgaaaaaaaataataaatttatttttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatttttgg
ttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagttttaat
aataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttactaactaa
taaagttattatcaatcaaaaaactttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaattctttaa
aagaaataaatgtatatgaaaataatatcccatttttttcttttgaatagaaaagaatttaaatctcattaa

FJ705305 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(USA5), complete sequence.

ttttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaattttttttttaaagttttt
ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaattttttttcagaaatatctaatagcataaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgtttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttctttgctttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt
cattgtactaaaaatttttatatataaaatatatactttacttttattttgttttatgaataaatataatttgtaagataaaaacattcttatcataaatttttaattatt
ttcaatataacttattaaatttttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaatttttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaattttaaatgctttaattaattattatacagcaaatgaatt

aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttatagt
gagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaattaca
aatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatcaaaa
aaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaattca
catttttttagctacttcaacacttcatatggaatctaaattaagaaagaattttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcattacgtt
atactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtgaaaac
tattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctataattt
ctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatgggatgg
gagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataaaga
aatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggtattc
atcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgatcag
gaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcagataa
aaaaggacaagtgtttgattatgatttagaagctttagcatttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttagtgtgcaatct
atttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctatttat
caagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagatattt
tagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatgg
aaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatattg
ctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaatttttacaggttttaagagaacatttttcattatttataaaaacaaaa
gaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattaggat
ctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaatttt
aagaccatctaatttatataaggaattaaatttttttatcacccttacgtaatgatattgtaaaacatggttttaatatattatgtgttagagaattaactggag
gaatttatttttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattgctc
atttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctattttatggagagaagtggttgaa
gaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgtgtt
ctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggtttata
tgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagtatg
aatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattatttaaaa
acaaatgagatgggtgatgttattgctgatttttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgtacat
tctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgtacag
ttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaacaaatg
caaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacagggtat
gactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagtagaac
atgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaaggatgtta
tcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaacgaa
tgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttattctcct
catggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatcaaata
ttttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaaagattt
agctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgctcgaat
agaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaagacaa
gctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggttatca
gaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcagcatta

gctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgtagttg
ctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttcatgatt
ggcgttatcttgattctaaccaattgaaaaaaaataataaatttatttttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatttttgg
ttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagttttaat
aataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttactaactaa
taaagttattatcaatcaaaaaacttttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaattctttaa
aagaaataaatgtatatgaaaataatatcccatttttttcttttgaatagaaaagaatttaaatctcattaa

FJ705306 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(USA6), complete sequence.

ttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaattttttttttaaagttttt
ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaatttttttttcagaaatatctaatagcataaaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgtttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttctttttgctttttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt
cattgtactaaaaatttttatatataaaatatatactttacttttattttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaattttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaattttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttatagt
gagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaattaca
aatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatcaaaa
aaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaattca

catttttttagctacttcaacacttcatatggaatctaaattaagaaagaattttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcattacgtt
atactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtgaaaac
tattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctataattt
ctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatgggatgg
gagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataaaga
aatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggtattc
atcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgatcag
gaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcagataa
aaaaggacaagtgtttgattatgatttagaagctttagcatttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttagtgtgcaatct
atttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctatttat
caagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagatattt
tagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatgg
aaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatattg
ctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaattttacaggttttaagagaacatttttcattatttataaaaacaaaa
gaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattaggat
ctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaatttt
aagaccatctaatttatataaggaattaaatttttttatcacccttacgtaatgatattgtaaaacatggttttaatatattatgtgttagagaattaactggag
gaatttattttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattgctc
atttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctattttatggagagaagtggttgaa
gaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgtgtt
ctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggtttata
tgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagtatg
aatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattatttaaaa
acaaatgagatgggtgatgttattgctgatttttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgtacat
tctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgtacag
ttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaacaaatg
caaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacagggtat
gactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagtagaac
atgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaaggatgtta
tcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaacgaa
tgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttattctcct
catggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatcaaata
ttttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaaagattt
agctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgctcgaat
agaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaagacaa
gctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggttatca
gaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcagcatta
gctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgtagttg
ctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttcatgatt
ggcgttatcttgattctaaccaattgaaaaaaaataataaatttatttttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatttttgg
ttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagttttaat
aataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttactaactaa

taaagttattatcaatcaaaaaactttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaattctttaa
aagaaataaatgtatatgaaaataatatcccatttttttcttttgaatagaaaagaatttaaatctcattaa

FJ705307 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(USA7), complete sequence.

ttttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaattttttttttaaagttttt
ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaattttttttcagaaatatctaatagcataaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgtttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttctttgctttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt
cattgtactaaaaattttatatataaaatatatactttacttttattttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaatttttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaatttttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtaccccccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttat
agtgagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaatt
acaaatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatca
aaaaaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaa
ttcacatttttttagctacttcaacacttcatatggaatctaaattaagaaagaatttttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcatta
cgttatactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtga
aaactattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctat
aatttctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatgggg
atgggagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataa

agaaatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggt
attcatcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgat
caggaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcaga
taaaaaaggacaagtgtttgattatgatttagaagctttagcatttttttagtaaaaaacaagaaaatgcagaatatttttatttaaaatttttttagtgtgcaa
tctatttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctattt
atcaagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagata
ttttagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatg
gaaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatatt
gctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaattttacaggttttaagagaacatttttcattatttataaaaacaaa
agaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattagg
atctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaat
ttaagaccatctaatttatataaggaattaaatttttttatcacccttacgtaatgatattgtaaaacatggttttaatatattatgtgttagagaattaactgg
aggaatttattttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattg
ctcatttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctattttatggagagaagtggttg
aagaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgt
gttctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggttt
atatgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagt
atgaatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattattta
aaaacaaatgagatgggtgatgttattgctgatttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgt
acattctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgt
acagttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaac
aaatgcaaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacag
ggtatgactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatctttggtattggcacatcagaagta
gaacatgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaagga
tgttatcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaa
cgaatgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttatt
ctcctcatggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatc
aaatattttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaa
agatttagctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgct
cgaatagaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaa
gacaagctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggt
tatcagaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcag
cattagctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgt
agttgctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttca
tgattggcgttatcttgattctaaccaattgaaaaaaaataataaatttattttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatt
ttggttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagtt
ttaataataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttacta
actaataaagttattatcaatcaaaaaactttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaatt
ctttaaaagaaataaatgtatatgaaaataatatcccatttttttcttttgaatagaaaagaatttaaatctcattaa

FJ705308 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pLeu-Dn(USA8), complete sequence.

ttttttgttttaaaaaaaaaatagcgaaagccttaattttcatcaaggctttcgctatttttattatatgtttatttatctttttatataagattaaaacactgtta
ttttgcatttatattatagaaagcatatatcaaaaaaaacaacatgcatttaaaagtttaattccaaaaaaaaactttaaaaaaaaattttttttttaaagttttt
ttttgaattaaactaaataatataattaatgataaatttttaaaaaaaaaattgaaataagtaaattttagaaaaagaaaatcgtacgcaaaaaaaaacg
tattattttagaaaaacattttatgtttaatttttcaatacatgttcgtagctatttttcgtaaataatgataactaatatttacttgtttctttagaccactagga
cctagtttagttaactcactcagagtgtacttagatactaatgcacgaagaatagtttgtctggcttctttttcatctaatgctattagacgttgtgcatttttt
ctttttttataaaaagcatgtcttgacattctgtatttaaaaatacttttcatttgaatatctttagatcgtcttttagcatctagcatagttataggatttagac
ctttactaataagatttttatttatccatcccaattgttgttgtttagcattcattaatttttttcagaaatatctaatagcataaaaaataatggagttaattta
atcattttaggcatataattacctaatattttatcccatatttttcacagttaacaaaacccataggttccataaaatttgtgattaatcgtgatgcacgagt
aatagatttatttccagattgagatattgttgatagaccacattcatcagataattgttcaactgaagcttgaactaattcagaagaaatattaaaatgat
aaagcattgctaaaaccatagctctcatggctgatgctctgtgtttatttaattgtctaaatctttttaagtgcaatcctgtcctaatgtttttgcttgtaaaa
aataatttaattcacatcttgctacgtctatttctttgctttttcatggcatagcaaataaaaaatggtcgacgtttattatttttaggaggtttaaaaattg
gattgggattatgtatgtaacattttcttgatttcatagatgatacttaaaatttaaatgataaaaaattgatttttttattcataacaaaatctatatgaagtatt
cattgtactaaaaatttttatatataaaatatatactttacttttatttgttttatgaataaatataatttgtaagataaaaacattcttatcataaattttaattatt
ttcaatataacttattaaatttttaacatattttatgttatataaaaaaaattaatgcattttatttaggatttttaaaaaagtttaataaaaaattaatttagga
attataaaaatggaaagaattattgaaaaagctatatatgcatctcgttggttgatgtttcctgtttatgttggtttatcatttggttttatattattaacattga
aatttttttcaacaaattgtatttattatcccagatattttagctatgtcagaatccggtttggtattagttgtattatcattaattgatattgctttagtaggtgg
acttttagtaatggttatgttttctggatatgagaattttatttctaaaatggatattcaagataatgaaaaaagattaggttggatgggtactatggatgt
aaactctataaaaaataaagtagcttcatcaatagttgcaatatcttctgtacatcttttacgtctttttatggaagctgaaaaaatattagatgataaaatt
atgttatgtgttataatccatcttacttttgtattatctgcatttggtatggcatacattgataaaatgagtaaaaaaaaacatgttcttcactaatacaaag
agaaaaataatttaattagattggttaaaacaatttgtaatgtaatcacattaatgttatattcattttaaaaatcaaaattcatttttcgataaaagaattagt
ttttctaaaaatgtttttatatacaaatatattttacttaatcaagtgtatttagtgcctagaaaaaattatatatataatcctaaaccagtttttaatccaccta
aaaataagagaaaaatatctacatttatttgttatgctatgaaaaaagcatccaaaatagatgttgctagaagcaatttaaattataccctattacctata
gatcctaagactggtaatattttacctcgatttagaagattaaatgaacatcgtgcgtgtgctatgagagctatagttctggctatgttatattattttgat
attcattcgaatttagttgaagcttctattgaaaagttagcagatgaatgtggtctttctacgttttcagattctggtaataaatctatcacccgtgtttcac
gtttaataaatgatttttagaaccaatgggttttgttagatgtaaaaaaatcaaaagaaaattcgttagcaattatatacctaaaaaaatatttttgacac
caatgtttttatgttgtttaacatatcacaatcaaagataaatagttatttatttaagtcaaagaagatgtctcaaaatttaaaaattacagaaaaaaaaat
atttatttctttttcagatattaaggtcatgtcacaattagatgaaaaatctattagaaaaaaaattttaaatgctttaattaattattatacagcaaatgaatt
aacgaaaataggccctaagggcctaaagaaacgaatagatattgaatataataatttatgtaaattgtttaaaaaaattaaaaaataaatttatacaat
aaaaatttcataagacatgggggtacaaaacattgtacccccatgtcttatgaaatttttattgtataaatttatttttttaataaaacatctataatttatagt
gagaaaatttttatgagttctaaagttattatttttgataccaccttacgtgatggtgaacaagcattacaagcaagtttaagcgttaaagaaaaattaca
aatagcattgtccctagaaaaatgtggcatagatattattgaagtaggatttcctatttcatcacctggagattttaaatcagttcagaccatatcaaaa
aaaattaaaaatagtaaaatatgtagtttagcaagatgtgttgaaaaagatatagaggtagcaggtgatgcaatgtcttcatctgatttttttcgaattca
cattttttagctacttcaacacttcatatggaatctaaattaagaaagaattttaatgaaattatagatatgtctgtttcatcagtaaaaaaagcattacgtt
atactgatgatgttgaattttcttgtgaagatgcaagtagaaccacaatggataatttatgtcgaattgtagagaaattgataaaatctggtgtgaaaac
tattaatatccctgatacagtaggttatgctatacctaatgaattatctaatattataaaaaatttatttgaacgagtaccaaatattcataaatctataattt
ctgttcattgtcataatgatttaggaatggcagtaggtaattcgatatctgctatacaagcaggtgctagacagattgaaggtactattaatgggatgg
gagaaagagctggaaatacagcattagaagaaattattatggctataaaagttagagaagatattttaggtgtttcaacaaacatagtacataaaga
aatttatcgtactagtcaaattatcagtcaaatttgtaatatgcctatacctgctaataaagcaattgtaggtagtaatgcattttcacattcttctggtattc
atcaagatggtgtattaaaaaatagaaaaaattatgaaattatggaacctaatactattggtgtaaaagaagtaaaacttaatttaacttcacgatcag
gaagagctgcagtaaaatactatatggataaaatgggttataaagatcatgattatgatatagatgaactttactctgcatttttaaaattagcagataa
aaaaggacaagtgtttgattatgatttagaagctttagcatttttagtaaaaaacaagaaaatgcagaatattttatttaaaatttttagtgtgcaatct
atttctaatggcttatctactgcttcagtgaaattaaaatgtggtaaaaaagtttatacagaatcttctactactagtaatggaccggtagatgctatttat

caagcattaaataaaattataaattttccaataacattacaaaaatttcaacttgtagcaaaaggaaagggtaaagatgcattaggtcaagtagatattt
tagtgaaatatgaaaatcgtcaatttcatggaataggtttagctacggatattattgaatcgtcagctcaagctatgattgatgttttaaataatatatgg
aaatctcaacaagtgaataaaaagctaaaaaatttaaagaaatattaaaattataatatttattattgaggttaatttttatgcataaacaatatcatattg
ctgtattacctggagatggaataggtcctgaagttatgcaagaagcatataaaattttacaggttttaagagaacatttttcattatttataaaaacaaaa
gaattcgatattggagggatagctattgataatcatggtatagcattacctaaaaaaacactaataggatgtgaaaattctgatgcgattttattaggat
ctatcggagggaaaaaatgggatacattgcctataaatgaacgtcctgaaagagcttcactattacccctaagaaaacatttcaatttttttgtaattt
aagaccatctaatttatataaggaattaaatttttttatcacccttacgtaatgatattgtaaaacatggttttaatatattatgtgttagagaattaactggag
gaatttattttggaaaaccaagagggcgcgtaactaaaaaaaatctaatgtatgcttttgatacagaaatttattataaatttgaaattgttcgtattgctc
atttagcttttaaattagcgcgctctagaaagcataagttatgttctatagataaagctaatgttctcgaaagttctattttatggagagaagtggttgaa
gaggtttctaaggaatatcctgatgttattttatctcatttatatattgacaatgtttgtatgcaaattattaaagatccaaatcaatttgatgtacttttgtgtt
ctaatctttttggggatatcatttcagatgaatgtgccatgattacaggttctataggaatgttaccttctgctagtttaaatgaaaaaaactttggtttata
tgaaccagcaggggttcagctcctgatattcagggaaaaaatattgctaatcctatagctcagattctttcgctttctatgttaattagatacagtatg
aatttaaataaaatagcaaataaaattgataacgctgttattaatgttttaaaaaaaggttataaaactatggatatatctaaagatcaaaattatttaaaa
acaaatgagatgggtgatgttattgctgatttttttaaaaagagataaataaaaatgaataaaacattatatgaaaaaatatatgattcacatgttgtacat
tctgaaaaaaatggtttatctattttatatgtagatttacatttattgcatgaagttacatctcctcaagcttttgaatcactgcgcaataagaatcgtacag
ttagacaacctaaaaaaacatttgctacaatggatcataatgtttcaacagaaagtaaagatattaatgcatcaggttcaatggcaaaaaaacaaatg
caaacattaataaaaaattgtaaagagtttcatatatcattatatgatttaaatcaccctaatcaaggtatagttcatgtaatcagtcctgaacagggtat
gactttacctggcatggttattgtgtgtggtgattcccatacttctactcatggagcatttggtgcattatcttttggtattggcacatcagaagtagaac
atgttcttgctacacaaacattaagacaacagcgttttaaaaatatgaaaatagaagttataggaaaaataggaaaatttattacagctaaggatgtta
tcttatatattatagggaagataggatcatctgctggaactggatatataattgagttttgtggaaacgttattaaaaagatgagtatggaagaacgaa
tgacagtttgtaatatggcaattgaactcggtgcaaaatctggattaatagcaccagacgaaactacttatttatatttaaaaaataaaacttattctcct
catggtcaaaattggcaaaaagcaatagagtactggaaaactcttaaaacagatcataatgcaatttttgataaagtatttactattgatatatcaaata
ttttacctcaagttacttggggtacaaatcctgatcaagttattggaattaatgaaaagatacctgattttacttctttccagaatattgtgaagaaagattt
agctaaatctgcatgtaaatatatggatttaaaaccaggtacttatttaacagatattacgattgataaagtttttattggatcttgcacaaatgctcgaat
agaagatttaagatctgcttcaaagatattaaaacataataagatttcaaaaaatgttaaagctattgttgtaccaggttcaggtttagtaaaaagacaa
gctgaaagtgaaggtttagataaaatttttattgaatctggttttgaatggcgtttacctggttgttctatgtgcttaggtatgaataatgataggttatca
gaaaatgaacgttgtgcttctactagtaatagaaattttgaaggtcgtcagggtagaaatggtcgcacacatttagttagtcctattatggcagcatta
gctgctttatacgggaaattttctaatcctaataaattaaattagataacatgagaatgattaagaaaatgcttaaatttattgaacatactggtgtagttg
ctcccttggaaatttctaatgtagatacagatactataatacctaaacaatttttaaagggtatagataaaaaggggttaggtaaatttttatttcatgatt
ggcgttatcttgattctaaccaattgaaaaaaaataataaatttatttttaaataaaaaaatttatgaaaatgctagcattttattaactaaaaaaaatttgg
ttgtggttcatcaagagagcatgcagtttggtctttattagattatggttttaaggtaatagtggcatctagttttagtgatatttttataataatagttttaat
aataaattgcttttaattacattagacgaaaaaaaaattgatagtatttttgatattgtaaaaaaaaatttaggtataaacatatctataaatttactaactaa
taaagttattatcaatcaaaaaactttttttttcaaattagatgaatttcgtcgtgtatctttcctaaatgatattgatcatatcgatttaactatgaattctttaa
aagaaataaatgtatatgaaaataatatcccatttttttcttttgaatagaaaagaatttaaatctcattaa

FJ705309 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnSAM anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.

attgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgcta
gatatcattcactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagt
gtgtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaa
ccgcaatatattaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgt

tataataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataattttattgttaaaagatatgata
tattgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatg
gcatatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagca
cttatacccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatct
caagttatacacagaaattgtggataacttttttaaaaatatttttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttatttttaccat
gaatataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaat
aaaaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaaagaaatatatgtatataatatttattgttattgttattt
ttttttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttttctgaaaattatcaaaaatatcgtatgacc
atattgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatc
aaaatcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagca
gaaattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcatt
atctaaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctca
ttttcagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacaga
aagcaatgttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgtttttattt
agcagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaa
aaagagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatga
gtgacgaacaatattgttccataataaaaaaattacaaatttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttc
taatcctttatctgcctatcaaaaattaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaa
agttctttaaaatatgatgacaca

FJ705310 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnSA anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.

attgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgcta
gatatcattcactagtggcatgtaaagatattcctaataatttgattattaattcttttttttaaaaaaaccattatgtctatacgaaataataaagatcgagt
gtgtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaa
ccgcaatatattaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgt
tataataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataattttattgttaaaagatatgata
tattgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatg
gcatatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagca
cttatacccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatct
caagttatacacagaaattgtggataacttttttaaaaatatttttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttatttttaccat
gaatataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaat
aaaaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaaagaaatatatgtatataatatttattgttattgttattt
ttttttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttttctgaaaattatcaaaaatatcgtatgacc
atattgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatc
aaaatcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagca
gaaattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcatt
atctaaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctca
ttttcagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacaga
aagcaatgttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgtttttattt
agcagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaa

aaagagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatga
gtgacgaacaatattgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttc
taatcctttatctgcctatcaaaaattaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaa
agttctttaaaatatgatgacacaacaagacaagtagaattatatccgattgctgg

FJ705311 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnUSA1 anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.
gattgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgct
agatatcattcactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagt
gtgtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaa
ccgcaatatattaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgt
tataataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataattttattgttaaaagatatgata
tattgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatg
gcatatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagca
cttatacccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatct
caagttatacacagaaattgtggataacttttttaaaaatatttttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttatttttaccat
gaatataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaat
aaaaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaaagaaatatatgtatataatatttattgttattgttattt
ttttttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttttctgaaaattatcaaaaatatcgtatgacc
atattgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatc
aaaatcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagca
gaaattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcatt
atctaaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctca
ttttcagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacaga
aagcaatgttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgttttttattt
agcagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaa
aaagagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatga
gtgacgaacaatattgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttc
taatcctttatctgcctatcaaaaattaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaa
agttctttaaaatatgatgacacaacaagacaagtag

FJ705312 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnUSA2 anthranilate synthase component II (ptrpG) pseudogene, partial sequence; anthranilate synthase component I (trpE) gene, complete cds; and anthranilate synthase component II (trpG) gene, partial cds.
gattgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgct
agatatcattcactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagt
gtgtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaa
ccgcaatatattaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgt
tataataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataattttattgttaaaagatatgata
tattgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatg

gcatatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagca
cttatacccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatct
caagttatacacagaaattgtggataacttttttaaaaatatttttttattataaaaaaaaaaatgtaacactatgttttaattatatttattctttattttaccat
gaatataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaat
aaaaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaagaaatatatgtatataatatttattgttattgttattt
ttttttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttctgaaaattatcaaaaatatcgtatgacc
atattgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatc
aaaatcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagca
gaaattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcatt
atctaaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctca
ttttcagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacaga
aagcaatgtttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgtttttattt
agcagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaa
aaagagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatga
gtgacgaacaatattgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttc
taatcctttatctgcctatcaaaaattaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaa
agttctttaaaatatgatgacacaacaagacaagtagaattatatccgattgctggcactagacccagagggaggaacatggatggcacattaaat
ctagatttggacagtcgaatcgaacttgaaatgagaaccaatcataaggaactcgccgaacatttaatgttagtagatttagctcgtaacgatttagc
acggatctgtgaaccgggatcaagatatgtctcggatttggttagagtcgacaaatatcctcatgtcatgcatttagtctctagagtcgtaggaacgtt
aaaaccagagttagatgcattacatgcttatgcagcttgcatgaatatgggtacattaactggtgctcctaaaattcgcgctatggaattgattgctga
atatgaaatggaacagaggggcagttacggtggtgccataggatattttactgatttaggaaatttagacacatgcattacaatacgttcggcttatgt
agaggacaatattgcaaccattcaatcaggatcaggtattgtttataattccatacccgaagacgaagtcaaagaaggtatcaataaagcgaaacc
tgtaataaatgctatccaacacgctcatcatttagtttaaggaacattgtacaatatggcaaatatcttattattagataatgtagattcttttacttataatc
ttgtagaacaactgaggaatcaaaacaatcaagttttaatttatcgtaatacggtagatattgaggtcatttttgacgctcttaagaaattaaaaaaccc
cattttaatgttatctccagggccgagtatcccccaatatgctggatgtatgttacccttgataaaaaaagtgaaaggttacctacctataataggtattt
gtttaggtcatcaggcgatagtagaagcttatggtggc

FJ705313 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnUSA3 anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.

gattgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgct
agatatcattcactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagt
gtgtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaa
ccgcaatatattaaatacagttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgt
tataataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataatttttattgttaaaagatatgata
tattgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatg
gcatatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagca
cttatacccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatct
caagttatacacagaaattgtggataacttttttaaaaatatttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttattttaccat
gaatataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaat
aaaaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaagaaatatatgtatataatatttattgttattgttattt
ttttttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttctgaaaattatcaaaaatatcgtatgacc

atattgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatc
aaaatcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagca
gaaattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcatt
atctaaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctca
ttttcagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacaga
aagcaatgttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgtttttattt
agcagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaa
aaagagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatga
gtgacgaacaatattgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttc
taatcctttatctgcctatcaaaaattaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaa
agttctttaaaatatgatgacacaacaagacaagtagaa

FJ705314 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnUSA4 anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.
gattgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgct
agatatcattcactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagt
gtgtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaa
ccgcaatatattaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgt
tataataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataatttttattgttaaaagatatgata
tattgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatg
gcatatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagca
cttataccccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatct
caagttatacacagaaattgtggataacttttttaaaaatatttttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttatttttaccat
gaatataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaat
aaaaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaaagaaatatatgtatataatatttattgttattgttattt
ttttttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttttctgaaaattatcaaaaatatcgtatgacc
atattgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatc
aaaatcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagca
gaaattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcatt
atctaaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctca
ttttcagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacaga
aagcaatgttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgtttttattt
agcagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaa
aaagagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatga
gtgacgaacaatattgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttc
taatcctttatctgcctatcaaaaattaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaa
agttctttaaaatatgatgacacaacaagacaagtagaattatatccgattgctgg

FJ705315 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnUSA5 anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.

tgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgctag
atatcattcactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagtgt
gtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaacc
gcaatatattaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgtta
taataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataattttattgttaaaagatatgatata
ttgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatggc
atatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagcactt
ataccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatctca
agttatacacagaaattgtggataacttttttaaaaatatttttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttatttttaccatga
atataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaataa
aaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaagaaatatatgtatataatatttattgttattgttatttttttt
tttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttctgaaaattatcaaaaatatcgtatgaccata
ttgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatcaaa
atcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagcaga
aattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcattatc
taaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctcatttt
cagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacagaaa
gcaatgttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgtttttatttag
cagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaaaa
agagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatgagtg
acgaacaatattgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttctaat
cctttatctgcctatcaaaaattaaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaaagtt
ctttaaaatatgatgacacaacaagacaagtagaat

FJ705316 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnUSA6 anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.

tgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgctag
atatcattcactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagtgt
gtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaacc
gcaatatattaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgtta
taataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataattttattgttaaaagatatgatata
ttgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatggc
atatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagcactt
ataccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatctca
agttatacacagaaattgtggataacttttttaaaaatatttttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttatttttaccatga
atataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaataa
aaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaagaaatatatgtatataatatttattgttattgttatttttttt
tttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttctgaaaattatcaaaaatatcgtatgaccata
ttgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatcaaa
atcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagcaga
aattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcattatc

taaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctcatttt
cagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacagaaa
gcaatgttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgttttttatttag
cagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaaaa
agagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatgagtg
acgaacaatattgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttctaat
cctttatctgcctatcaaaaattaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaaagtt
ctttaaaatatgatgacacaacaagacaagtagaa

FJ705317 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnUSA7 anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.
gcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgctagatatcatt
cactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagtgtgtggtttt
cagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaaccgcaatat
attaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgttataataact
atatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataattttattgttaaaagatatgatatattgaaaac
aaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatggcatatatcta
ggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataatttagatcatcattcatataaaatagaagtgcagcacttatacccccc
atataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatctcaagttataca
cagaaattgtggataacttttttaaaaatatttttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttattttaccatgaatataatac
tgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaataaaaaatatta
aagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaaagaaatatatgtatataatatttattgttattgttatttttttttggtttat
aaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttttctgaaaattatcaaaaatatcgtatgaccatattgctttaat
tttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatcaaaatcattcaa
aaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagcagaaattaacaa
aaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcattatctaaaaatgg
tgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctcattttcagaaaaa
tttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacagaaagcaatgttttt
tggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgttttttatttagcagaaacat
tactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaaaaagagggag
agaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatgagtgacgaacaat
attgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagtt

FJ705318 *Buchnera aphidicola* (*Diuraphis noxia*) plasmid pTrpEG clone BUHpTrpEGDnUSA8 anthranilate synthase component II (ptrpG) pseudogene, partial sequence; and anthranilate synthase component I (trpE) gene, partial cds.
tgaatatgcaggtgagatattccatggaaaagcctctttaatcaatcatgacggtttggagatgtttaaagatatccctcagccattacctgttgctag
atatcattcactagtggcatgtaaagatattcctaataatttgattattaattctttttttaaaaaaaccattatgtctatacgaaataataaagatcgagtgt
gtggttttcagtttcatcctgaatctattttaactacatgcggtgatcaaatattaaaacaaataatttattgggcttcgttgcaatatgttatgaagtaacc
gcaatatattaaatacagtttttattatattgtgatatggttttaattcagcaatttttttaacacccctccgtttgcatgtttgtatatatttctatcaaattgtta
taataactatatatttatctgtatatgttgtgttacttatcagtcatcatttattcttatcatgcatattcataataaaataattttattgttaaaagatatgatata

ttgaaaacaaatcaattcacatgtgctcattttattagttgatatttgcatatatcatgatcctgcgcactctcaatagagtggtgatataaagacatggc
atatatctaggtaatatgaaagaaaatagattatatgatgagtcgccaaacgattaataattttagatcatcattcatataaaatagaagtgcagcactt
ataccccatataaaaatttttgttatagactgtttgtcaaatattgttatcatcataatgatacatcaaacaacataaatacaaaatgaattttatatctca
agttatacacagaaattgtggataacttttttaaaaatatttttttttattataaaaaaaaaaaatgtaacactatgttttaattatatttattctttatttttaccatga
atataatactgttattaacatgtttataagtctggggataagtttggtataatatgcaatgcccctccccacaatataaaaaattagaaaatttgaaataa
aaaatattaaagacaattcaataatctcaaaagataaaatccttaaaagatttaaaaaaaaaaagaaatatatgtatataatatttattgttattgttatttttttt
tttggtttataaaaaaattattagtatcattatatttaacatagtgtgttatttacttatatatattgtatttttttctgaaaattatcaaaaatatcgtatgaccata
ttgctttaattttgtaaattagttattatttcttctttataaaaatattcttttgacccccctaagagatttgaattatatggaaaaaaagccatacgagatcaaa
atcattcaaaaaaaagcaaagtatcatcctgacccaacaatagtatttaatcatatttgtggatctcaaaaacaaacattgttactagaaacagcaga
aattaacaaaaaaaatgatctagaaagtattatgatcatcgatgctgcgctacgaatttcttctgaaagaaatcactcggtacaactaacagcattatc
taaaaatggtgaaaatattttatcaattttaaaaagcaatctgaaacaaaaagttcaaatgttcatacaagatacatctattcgtttagaattccctcatttt
cagaaaaatttagatgaagacaaaaaaatttttttcattgtctatatttgacacttttagatttattatgaaatttttcaaaaatcgcaataaagtacagaaa
gcaatgttttttggtggactattctcctatgatttaatttctaattttgagttattaccaaaattaaaaaaaacacaaaaatgccctcacttttgttttttatttag
cagaaacattactgattgtagatcatcaaaaaaaaacatgtttaattcaaaatagtttatttacgaaaaattcccacgaacagatgagagtagaaaaa
agagggagagaaatacaaaaaaaacttgaagcatctttaaactctattcctgtaaggcaagaagtaaaaaatagtatgttaactgcaaatatgagtg
acgaacaatattgttccataataaaaaaattacaaattttaattcgaaaaggtgagatttttcaagttgtaccatctcgaaaatttttttttaccctgttctaat
cctttatctgcctatcaaaaattaaaaaaaagcaatcctagtccttatatgttttttatgcaagataaagattttaccttatttggtgcttcaccagaaagtt
ctttaaaatatgatgacacaacaagacaagtagaattatatccg

APPENDIX – CHAPTER 4

Sequences submitted to Genbank

GU145279 5'UTR from mRNA, leader sequence for leuA from *B. aphidicola* (*D. noxia*, SA biotype) plasmid pLeu-Dn(SA), primary endosymbiont.
1tataaatttatttttttaataaaacatctataatttatagtgagaaaatttttatgagttctaaagttattatttttgataccaccttacg

GU145280 5'UTR from mRNA, leader sequence for leuA from *B. aphidicola* (*D. noxia*, USA biotype) plasmid pLeu-Dn(USA), primary endosymbiont.
tttatttttttaataaaacatctataatttatagtgagaaaatttttatgagttctaaagttattatttttgataccaccttacg

GU145281 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, primary endosymbiont of *Diuraphis mexicana*, partial.
tcaatcacccgagcttcacgtttaataaatgatttttttagaaccaatgggttttattagatgtaaaaaaataaaaagtaaatctgttagtaattatatacct
aaaaaaatatttttaacaccaatgtttttttatgttatttaacatatcacaatcaaaaataaatagttatttatctaagtcaatgaaaatatctcaaaaattaaa
gatttcagaaaaaaaaatatttatttctttttcagatattaaggtaatgtcaaagttagatgaaaaatctgttagaaaaaaaatttttaaatgttttaattaatt
attatacagcaagtgaactgacaaaaataggccctaagggcctaaagaaacgaatagatgttgaatataataatttatgtaaattgtttaaaaaaatta
aaaaataaatttatacaataaaaatttcataagacatgaaggtacaaatttttgtaccttccatgtcttatgaaatttttattgtataaatttatttttttaaaaaa
aacattcaaaatctatagtgagaaaaatttttatgagttctaaagttattatctttgataccaccctacgtgatggtgaacaggc

GU145282 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, primary endosymbiont of *Diuraphis tritici*, partial.
Atgcggcttatctacttttttccattctggtaataaatctatcacccgtgcttcacgtttaataaatgatttttttagaaccgatggggtttgttaaatgcaaa
aaaataaacagtgaatctgttagtaattatatacctaaaaaaatatttttaacaccaatgttctttatgttgtttaacataccacaatcaaaaataaatagtt
atttatctaagtcaatgaacatgtctcaaaagttaaaaataacagaaaaaaaaatgtttatttctttttcagatattaaggtaatgtcaaaattagatgaaa
aatctgttagaaaaaaaatttttaaatgctttaattaattattatacagcaagtgaactaacaaaaataggccctaagggcctaaaacaaagaatagat
gttgaatataataatttatgtaaattatttaaaaaaattaaaaaataagttgatacaataaaaatttcataagacatgtaggtacacaattttgtacctacat
gtcttatgaaatttttttagtgtatccttatttttttttaaaaacaacttttttataatctttaaaaaggggagaaatttttttatgagttctaaagttattatctttgata
ccaccttacgtgatggtgaacaagcattacaagcaagtttgagcgttaaagaaaaattacaaatcgcattatcttctagaaaaatgtgggatagatat
ttattgagtagcttttcttatttcatcacctgg

GU145283 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, endosymbiont of *Macrosiphum rosae* (MF), partial.
aaattggtccgaaaggtcttaaaaaaaagatagatattgaatataataatttatgtaatttgcataaaaaaaaatcataaaataaatttatacaacaaaa
tttcataagacatcatatacaaaaaacatccttgtatatgatgtcttatgaaattttgttgtataaatttatttttattggtataaaacattctatttcctaatcga
gaaaaattttatgaattcgaaagttattatttttgatacaactttgcgggatggtgaccaagcattaggat

GU145284 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, endosymbiont of *Myzus persicae*, partial.
cgaaaattggtccgaaaggtcttaaaaaaaagatagatattgaatatcataatttatgtaaattatacaaaagaaaataataaaataaatttatacacaa
aaatttcataagacgtaaaaatgcaatacatgtgtttttaatgtcttatgaaatttttgtgtataaatttatttttatttaataaatcattttatttattgtcgagaa
aatttttatgaactctaaagttattatttttgatacaa

GU145285 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, endosymbiont of *Uroleucon sonchi* (ACAM), partial.
gaattaactaaaattggtcctcaaggtctgaaaaaaaagatagatatcgaatataataatttatgtaagttaaacaaacataaataacatcatattatac
aataaaaatttcataagacatcgttgtacaaaaatcatttttttgtacaacgatgtcttatgaaatttttattgtataatacgatgttatttaacaaacattcatt
ttatcggaataattttatgaattctaaagttattatttttgatacaactctacgtgatggtgatcaagcattaggatac

GU145286 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, endosymbiont of *Hyalopterus pruni* (MF), partial.
gaattaacgaaaattggtccgaagggtcttaaaaaaaagatagatattgaatatggtaaattatgtaaattatataaaaaaaaaatattagataaattata
caataaaaatttcataaagcagaaagatgcatttttaaaaatgcatctttctgttttatgaaatttttattgtataatttatctaatttaaacaaactatttttttaa
aattagcgagaaaaaatttatgaattcaaaaattattatttttgatacgactctacgagatggggaccaagcattaggatctagta

GU145287 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, endosymbiont of *Hyalopterus pruni* (MF), partial.
gaattaacaaaaataggtcctcaaggtcttaaaaaaaagatagatattgaatatggtaaattatgtaaattatataaaaaaaaaatattagataaattata
caataaaaatttcataaaacagaaagatgcatttttaaaaatgcatctttctgttttatgagatttttattgtataatttatctaatttaaacaaactatttttttaa
aattagcgagaaaaaatttatgaattcaaaaattattatttttgatacgactctacgggatggagaccaagcattaggatacagaatgg

GU145288 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, endosymbiont of *Brevicoryne brassicae* (MF), partial.
aatagatattgaatataataatttatgtaaattatttaacaaaattaaaaaataaatttacacaataaaaatttcataagacatcaaggtacaaaattttgta
ccttgatgtcttatgaaatttttattgtgtaaatttatttttattcaaaaacatttttatttttttaaaagatgagaaaaatttttatgaattctaaagttattatttttt

GU145289 repA, leuA and inverted repeat region of *Buchnera aphidicola* plasmid pLeuABCD, endosymbiont of *Brevicoryne brassicae* (MF), partial.
gaattaacgaaaataggtcctaaagggcttaaaaaacaaatagatattgaatataataatttatgtaaattatttaacaaaattaaaaaataaatttaca
caataaaaatttcataagacatcaaggtacaaaattttgtaccttgatgtcttatgaaatttttattgtgtaaatttatttttattcaaaaacatttttatttttttaa
aagatgagaaaaatttttatgaattctaaagttattatttttgataccaccttacgcgatggagatcaagcattaggatac

APPENDIX – CHAPTER 5

Table Appx 5.1 Scripts, codes and files used and/or produced during the data analyses of the different Affymetrix experiments.

Script, code or filename	***Description***
Script 1 Data in, normalization & quality check	Read all the **.Cel* files and *Target.txt* files in and normalizes with RMA, PLM, GCRMA, VSN and MAS5.0. Provides various figures to investigate the quality of the slides, *i.e.* raw data (*Figure 5.2*).
Script 2 Quality control analysis	Uses the **.RData* files generated with script 1 and produce graphs for the different normalized datasets (*Figure 5.2*).
Script 3 Getting differentially expressed genes	Identifying differentially expressed genes after normalization with RMA, GCRMA, VSN, PLM and MAS5, using LIMMA and BA packages for 95, 99 and 99.9 % confidence. The FDR is determined using Bonferonni and Benjamini-Hochberg FDR methods and written to excel files and different folders (*Figure 5.2*).
Target.txt	A summary, tab delimited file that contains sample and slide names, filenames, replicates, *etc.* Data in this file is used for the figure legends.
Convert	*MS Excel 2007* macro that copies all the worksheets, *i.e.* **.csv* files, of one or more workbooks in the same directory as the workbook in which the macro is run, to that workbook. Note that the worksheets must be named differently.

Note: Script 1, 2 and 3, Target and Convert follows this table via the links above.

Script 1 Data in, normalization & quality check

```
#Name:    Script 1 Data in, normalization & quality check.R
#Author   Z.H. Swanevelder, Department of Genetics, University of Pretoria, RSA
#Ref:     Please cite this script as: Z.H. Swanevelder (2010). Aphid-plant
          interactions and the possible role of an endosymbiont in aphid biotype
          development. Ph.D. Thesis, University of Pretoria.
#Aim:     Read the *.Cel files in and normalizes with RMA, GCRMA, PLM, VSN and
          MAS5.0
#Usage:   Run script in directory containing the *.CEL & Target.txt files
#Notes:   1. Target.txt is a file with phenotypic data like: Names (sample
            names), CEL Filenames, etc. in tab delimited format. Labels in the
            Target.txt file are used in figures (see example attached at the end
            of these scripts)
#         2. Running the script twice in the same directory may give a warning
            that the directories exist - it can be ignored HOWEVER files WILL BE
            overwritten
#         3. Output RData files are saved into a newly made directory called "Raw
            and Normalized R datasets"
#         4. Figures of the raw data is saved into the directory "/R
            results/Quality control raw data/" and a summary of the raw data is
            given in the file "Summary of all data.jpg". Subfolders: "Slide
            Images raw/"; "Log2 plots/" & "MA plots/" contains raw data slide
            comparisons and slide images


#1.LOAD LIBRARIES
print("Loading libraries....")
library(affy)
library(gcrma)
library(limma)
library(affyPLM)
library(marray)
library(annaffy)
library(geneplotter)
library(made4)
library(multtest)
library(vsn)
library(graphics)
library(IDPmisc)
print("Libraries Loaded")

#2.READ DATA FROM .CEL FILES INTO VARIABLES
base<-getwd()
DATA<-paste(base,"/Raw and Normalized R datasets/",sep=""); DATA
dir.create(DATA,showWarnings=TRUE,recursive=FALSE)
if (file.exists("Target.txt"))
     {
     print("Target.txt exist. Reading targets and other info in")
     pd<-read.AnnotatedDataFrame("Target.txt",header=TRUE,row.names=1)
     data<-ReadAffy(filenames=pData(pd)$FileName)
     data@phenoData<-pd
     save(data,file=paste(DATA,"Rawdata.RData",sep=""))
 } else {
     print("Target.txt DOESN'T exist. NO extra phenoData will be added!!")
     data<-ReadAffy()
```

```
     save(data,file=paste(DATA,"Rawdata.RData",sep=""))
     }

setwd("Raw and Normalized R datasets")
print("Running Background and Normalization scripts....")
esetRMA<-rma(data); print("RMA DONE...")
save(esetRMA,file="esetRMA.RData"); print("RMA FILE DONE...")
esetVSN<-vsnrma(data); print("VSN DONE...")
save(esetVSN,file="esetVSN.RData"); print("VSN FILE DONE...")
esetGCRMA<-gcrma(data); print("GCRMA DONE...")
save(esetGCRMA,file="esetGCRMA.RData"); print("GCRMA FILE DONE...")
esetPLM<-fitPLM(data); print("PLM DONE...")
save(esetPLM,file="esetPLM.RData"); print("PLM FILE DONE...")
esetMAS<-mas5(data); print("MAS5 DONE...")
save(esetMAS,file="esetMAS.RData"); print("MAS5 FILE DONE...")
setwd(base)

#SETTING GRAPHICS LEGEND
print("Setting graphics")
fil<-(2:(length(data)+1))
lgnd<-(is.null(data$Name))
if (lgnd)
     {
     print("Target.txt doesn't exist. Default samplenames!")
     legend<-(sampleNames(data))
   } else {
     print("Setting the legend according Target.txt Name")
     legend<-(data$Name)
     }
print("Legend set...done")

#DIRECTORIES
base<-getwd()
RESULTS<-paste(base,"/R results/Quality control raw data/",sep=""); RESULTS
dir.create(RESULTS,showWarnings=TRUE,recursive=TRUE)
setwd(base)
print("Directory Results rawdata analysis made")

#ASSESSMENT OF RAW DATA
print("ASSESSMENT OF RAW DATA: IMAGE ANALYSIS")

#PLOTTING RAW IMAGES OF ARRAYS
print("Single slide comparisons")
RAWIMAG<-paste(RESULTS,"Slide Images raw/",sep=""); RAWIMAG
dir.create(RAWIMAG, showWarnings = TRUE, recursive = FALSE)
number<-length(data)
lgnd<-(is.null(data$Name))
if (lgnd)
    {
    print("No target.txt file!")
    for (x in seq(1,number,1))
       {
       filename<-paste(RAWIMAG,sampleNames(data[,x]),"_raw.jpg",sep="")
       filename2<-paste(RAWIMAG,sampleNames(data[,x]),"_log2_raw.jpg",sep="")
       jpeg(filename)
       par(mfrow=c(1,1))
       image(data[,x],transfo=I)
       dev.off()
```

```
      jpeg(filename2)
      par(mfrow=c(1,1))
      image(data[,x],transfo=log2)
      dev.off()
      }
      print("Raw Images DONE")
  } else {
      print("Name from target.txt file is used!")
      for (x in seq(1,number,1))
        {
        filename<-paste(RAWIMAG,data$Name[x],"_raw.jpg",sep="")
        filename2<-paste(RAWIMAG,data$Name[x],"_log2_raw.jpg",sep="")
        jpeg(filename)
        par(mfrow=c(1,1))
        image(data[,x])
        dev.off()
        jpeg(filename2)
        par(mfrow=c(1,1))
        image(data[,x],transfo=log2)
        dev.off()
        print(paste(data$Name[x],"...done",sep=""))
        }
      print("Raw Images DONE")
      }

#EXPRESSION SET OF LOG2 FOR SLIDE COMPARISONS
print("Log2 plot comparing replicates in an treatment: RAW data")
LOGP<-paste(RESULTS,"Log2 plots/",sep=""); LOGP
dir.create(LOGP, showWarnings = TRUE, recursive = FALSE)
exprs(data) <- log2(exprs(data))
number<-length(data)
lgnd<-!(is.null(data$Name))
if (lgnd)
    {
    for (x in seq(1,number,3))
      {
      y<-(x+1)
      z<-(x+2)
      filename<-paste(LOGP,data$Name[x],"_vs_",data$Name[y],"_log2.jpg",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      xlable<-paste("Log2 expression on array:",data$Name[x],sep=" ")
      ylable<-paste("Log2 expression on array:",data$Name[y],sep=" ")
      plot(2^exprs(data)[,x],2^exprs(data)[,y],log="xy",xlab=xlable,ylab=ylable,
      cex=.25,pch=16,yaxt="n")
      lines(2^seq(-2,16),2^seq(-1,17),col="red")
      lines(2^seq(-1,17),2^seq(-2,16),col="red")
      dev.off()
      filename<-paste(LOGP,data$Name[x],"_vs_",data$Name[z],"_log2.jpg",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      xlable<-paste("Log2 expression on array:",data$Name[x],sep=" ")
      ylable<-paste("Log2 expression on array:",data$Name[z],sep=" ")
      plot(2^exprs(data)[,x],2^exprs(data)[,z],log="xy",xlab=xlable,ylab=ylable,
      cex=.25,pch=16,yaxt="n")
      lines(2^seq(-2,16),2^seq(-1,17),col="red")
      lines(2^seq(-1,17),2^seq(-2,16),col="red")
      dev.off()
```

```
      filename<-paste(LOGP,data$Name[y],"_vs_",data$Name[z],"_log2.jpg",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      xlable<-paste("Log2 expression on array:",data$Name[y],sep=" ")
      ylable<-paste("Log2 expression on array:",data$Name[z],sep=" ")
      plot(2^exprs(data)[,y],2^exprs(data)[,z],log="xy",xlab=xlable,ylab=ylable,
      cex=.25,pch=16,yaxt="n")
      lines(2^seq(-2,16),2^seq(-1,17),col="red")
      lines(2^seq(-1,17),2^seq(-2,16),col="red")
      dev.off()
      }
    }

#MA PLOT COMPARISONS
print("MA plot comparing replicates in an treatment ")
MAPLOTS<-paste(RESULTS,"MA plots/",sep="")
MAPLOTS
dir.create(MAPLOTS, showWarnings = TRUE, recursive = FALSE)
number<-length(data)
lgnd<-(is.null(data$Name))
if (lgnd)
    {
    print("No target.txt file!")
    for (x in seq(1,number,3))
      {
      y<-(x+1)
      z<-(x+2)
      filename<-
      paste(MAPLOTS,sampleNames(data[,x]),"_vs_",sampleNames(data[,y]),"_raw.jpg
      ",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      MAplot(data,which=c(x,y),ref=x)
      dev.off()
      print(filename)
      filename<-
      paste(MAPLOTS,sampleNames(data[,x]),"_vs_",sampleNames(data[,z]),"_raw.jpg
      ",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      MAplot(data,which=c(x,z),ref=x)
      dev.off()
      print(filename)
      filename<-
      paste(MAPLOTS,sampleNames(data[,y]),"_vs_",sampleNames(data[,z]),"_raw.jpg
      ",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      MAplot(data,which=c(y,z),ref=y)
      dev.off()
      print(filename)
      }
    print("MA comparisons DONE");
   }else {
    for (x in seq(1,number,3))
      {
      y<-(x+1)
      z<-(x+2)
```

```
      filename<-
paste(MAPLOTS,data$Name[x],"_vs_",data$Name[y],"_raw.jpg",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      MAplot(data,which=c(x,y),ref=x,identify=TRUE)
      dev.off()
      filename<-
paste(MAPLOTS,data$Name[x],"_vs_",data$Name[z],"_raw.jpg",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      MAplot(data,which=c(x,z),ref=x,identify=TRUE)
      dev.off()
      filename<-
paste(MAPLOTS,data$Name[y],"_vs_",data$Name[z],"_raw.jpg",sep="")
      jpeg(filename)
      par(mfrow=c(1,1))
      MAplot(data,which=c(y,z),ref=y,identify=TRUE)
      dev.off()
      }
    print("MA & Hexbin comparisons DONE")
   }

#SUMMARY OF RAW DATA
filename<-paste(RESULTS,"Summary of all data.jpg",sep="")
jpeg(filename)
par(mfrow=c(1,1))
overview(data)
dev.off()

print("Analysis done! Hit enter to return to shell")
```

Script 2 Quality control analysis

```
#Name:    Script 3 Quality control analysis.R
#Author:  Z.H. Swanevelder, Department of Genetics, University of Pretoria, RSA
#Ref:     Please cite this script as: Z.H. Swanevelder (2010). Aphid-plant
          interactions and the possible role of an endosymbiont in aphid biotype
          development. Ph.D. Thesis, University of Pretoria.
#Aim:     Generating graphs for normalized datasets
#Usage:   Run script after and in the same directory as script 1. It uses the
          normalized data in the directory: "Raw and Normalized R datasets"
#Notes:   1. Labels in the Target.txt file are used in figures
#         2. Running the script twice in the same directory may give a warning
             that the directories exist - it can be ignored HOWEVER files WILL BE
             overwritten
#         3. Output: Files are saved into the directory "/R results/" under the
             sub-directories: Graphs of raw and normalized data/"; "Graphs of raw
             and normalized data/Log2 plots/" & "Graphs of raw and normalized
             data/MA plots/"

#FUNCTION: COMPARE REPLICATES WITH EACH OTHER
compare<-function(slidenum_x,slidenum_y,data_obj,direct,method)
  {
  head<-
  paste(pData(data_obj)$Name[slidenum_x],"_vs_",pData(data_obj)$Name[slidenum_y]
  ,sep="")
```

```
  print(head)
  filename<-paste(direct,"MA plots/",method,"_",head,".jpg",sep="")
  jpeg(filename,width=1000,height=1200,pointsize=12)
  par(mfrow=c(1,1))
  MAplot(data_obj,which=c(slidenum_x,slidenum_y),ref=slidenum_x)
  dev.off()
  filename<-paste(direct,"Log2 plots/",method,"_",head,"_log2.jpg",sep="")
  jpeg(filename)
  par(mfrow=c(1,1))
  xlable<-paste("Log2 expression ",pData(data_obj)$Name[slidenum_x],sep=" ")
  ylable<-paste("Log2 expression ",pData(data_obj)$Name[slidenum_y],sep=" ")
  plot(2^exprs(data_obj)[,slidenum_x],2^exprs(data_obj)[,slidenum_y],log="xy",xl
  ab=xlable,ylab=ylable,cex=.25,pch=16,yaxt="n")
  lines(2^seq(-2,16),2^seq(-1,17),col="red")
  lines(2^seq(-1,17),2^seq(-2,16),col="red")
  dev.off()
  }

comparePLM<-function(slidenum_x,slidenum_y,data_obj,direct,method)
  {
  head<-
  paste(pData(data_obj)$Name[slidenum_x],"_vs_",pData(data_obj)$Name[slidenum_y]
  ,sep="")
  print(head)
  filename<-paste(direct,"MA plots/",method,"_",head,".jpg",sep="")
  jpeg(filename,width=1000,height=1200,pointsize=12)
  par(mfrow=c(1,1))
  MAplot(data_obj,which=c(slidenum_x,slidenum_y),ref=slidenum_x)
  dev.off()
  }

#LOAD LIBRARIES
print("Loading libraries....")
library(affy)
library(gcrma)
library(limma)
library(affyPLM)
library(marray)
library(annaffy)
library(geneplotter)
library(made4)
library(multtest)
library(IDPmisc)
library(vsn)
library(graphics)
print("Libraries loaded")

#LOADING R OBJECT FILES
base<-getwd()
setwd("Raw and Normalized R datasets")
print("Loading ...")
load("Rawdata.RData"); print("Raw data in ...")
load("esetRMA.RData"); print("RMA data in ...")
load("esetGCRMA.RData"); print("GCRMA data in ...")
load("esetMAS.RData"); print("MAS data in ...")
load("esetVSN.RData"); print("VSN data in ...")
load("esetPLM.RData"); print("PLM data in ...")
setwd(base)
```

```
#READING OBJECTS INTO LIST
x<-1
normList<-
list("rawdata"=x,"esetRMA"=x,"esetVSN"=x,"esetGCRMA"=x,"esetMAS"=x,"esetPLM"=x)
if(is.object(data)) {normList$rawdata<-data} else {print("RAW DATA NOT LOADED")}
if(is.object(esetRMA)) {normList$esetRMA<-esetRMA} else {print("RMA DATA NOT
LOADED")}
if(is.object(esetGCRMA)) {normList$esetGCRMA<-esetGCRMA} else {print("GCRMA DATA
NOT LOADED")}
if(is.object(esetMAS)) {normList$esetMAS<-esetMAS} else {print("MAS DATA NOT
LOADED")}
if(is.object(esetVSN)) {normList$esetVSN<-esetVSN} else {print("VSN DATA NOT
LOADED")}
if(is.object(esetPLM)) {normList$esetPLM<-esetPLM} else {print("PLM DATA NOT
LOADED")}

#MAKING THE BASE DIRECTORIES
base<-getwd()
RESULTS<-paste(base,"/R results/",sep=""); RESULTS
dir.create(RESULTS,showWarnings=TRUE,recursive=FALSE)
IMAG<-paste(RESULTS,"Graphs of raw and normalized data/",sep=""); IMAG
dir.create(IMAG, showWarnings = TRUE, recursive = FALSE)
LOGG<-paste(RESULTS,"Graphs of raw and normalized data/Log2 plots/",sep="");
IMAG
dir.create(LOGG, showWarnings = TRUE, recursive = FALSE)
LOGG<-paste(RESULTS,"Graphs of raw and normalized data/MA plots/",sep=""); IMAG
dir.create(LOGG, showWarnings = TRUE, recursive = FALSE)

#GENERATING PLOTS
number2<-length(normList)
print("Number of objects to be analized:"); number2
name<-names(normList)
for (d in seq(1,number2,1))
  {
  if (is.object(normList[[d]]))
      {
      print("Analysing")
      print(name[d])
      nar<-normList[[name[d]]]
      number<-length(pData(nar)$Name)
      fil<-(2:(length(pData(nar)$Name)+1))
      legend<-pData(nar)$Name

      #BOXPLOTS
      print("Drawing the boxplot")
      filename<-paste(IMAG,"Boxplot_",name[d],".jpg",sep="")
      jpeg(filename,width=2000,height=1000,pointsize=16)
      par(mfrow=c(1,1))
      boxplot(nar,col=fil,names=legend)
      dev.off()
      print("Boxplot done!")

      if (d != 6)
          {
          #MA PLOT COMPARISONS IN AN TREATMENT
          print("Doing MA-Plots ...")
          for (x in seq(1,number,3))
```

```
            {
            y<-(x+1)
            z<-(x+2)
            Compared<-compare(x,y,nar,IMAG,name[d])
            Compared<-compare(x,z,nar,IMAG,name[d])
            Compared<-compare(y,z,nar,IMAG,name[d])
            }

          #HISTOGRAM OF DENSITY VS LOG INTENSITIES
          print("Drawing histogram")
          filename<-paste(IMAG,"Histogram_",name[d],".jpg",sep="")
          jpeg(filename,width=800,height=1000,pointsize=16)
          par(mfrow=c(1,1))
          print(nar)
          hist(nar,col=fil,names=pData(nar)$Name)
          legend("topright",legend,inset=.02,col=fil,lty=fil,bg="white")
          dev.off()
          print("Histogram done")
        } else {
          print("PLM being done")
          print("Doing MA-Plots ...")
          for (x in seq(1,number,3))
            {
            y<-(x+1)
            z<-(x+2)
            Compared<-comparePLM(x,y,nar,IMAG,name[d])
            Compared<-comparePLM(x,z,nar,IMAG,name[d])
            Compared<-comparePLM(y,z,nar,IMAG,name[d])
            }
          }
      }
  }
print("Analysis done! Hit enter to return to shell")
```

Script 3 Getting differentially expressed genes

```
#Name:    Script 3 Getting differentially expressed genes.R
#Author:  Z.H. Swanevelder, Department of Genetics, University of Pretoria, RSA
#Ref:     Please cite this script as: Z.H. Swanevelder (2010). Aphid-plant
          interactions and the possible role of an endosymbiont in aphid biotype
          development. Ph.D. Thesis, University of Pretoria.
#Aim:     Identifying differentially expressed genes after normalization with
          RMA, GCRMA, VSN, PLM and MAS5, using LIMMA and BA for 95, 99 and 99.9%
          confidence. The FDR is determined using Bonferroni and Benjamini-
          Hochberg FDR methods.
#Usage:   Run script in the same directory and after script 1 & 2.
#Notes:   1. Labels in the Target.txt file are used in figures
#         2. Running the script twice in the same directory may give a warning
            that the directories exist - it can be ignored HOWEVER files WILL BE
            overwritten
#         3. Output: Generating *.csv excel files with differentially expressed
            genes in the directory "/R results/Regulated genes/" & the sub-
            directories: "Differential genes 95/"; "Differential genes 99/" &
            "Differential genes 99.9/". The latter 3 directories are again
            divided into the different methods used for the normalization which
```

```
#           contains the files before FDR corrections, as well as the
            directories for the FDR correction methods.
#         4. Please change the 'model.matrix' layout, 'colnames' and 'contrast
            matrix' to suite your dataset

#LOAD LIBRARIES
print("Loading libraries....")
library(affy)
library(gcrma)
library(limma)
library(affyPLM)
library(marray)
library(annaffy)
library(vsn)
print("Libraries Loaded")

#LOAD FILES
base<-getwd()
setwd("Raw and Normalized R datasets")
print("Loading ...")
load("esetRMA.RData"); print("RMA data in ...")
load("esetGCRMA.RData"); print("GCRMA data in ...")
load("esetMAS.RData"); print("MAS data in ...")
load("esetVSN.RData"); print("VSN data in ...")
load("esetPLM.RData"); print("PLM data in ...")
setwd(base)

#WRITING NORMALIZED DATASETS TO A LIST FOR EASY ACCESS
x<-1
normList<-list("esetPLM"=x,"esetRMA"=x,"esetGCRMA"=x,"esetMAS"=x,"esetVSN"=x)
if(is.object(esetRMA))  {normList$esetRMA<-esetRMA} else {print("RMA DATA NOT
LOADED")}
if(is.object(esetGCRMA)){normList$esetGCRMA<-esetGCRMA} else {print("GCRMA DATA
NOT LOADED")}
if(is.object(esetMAS))  {normList$esetMAS<-esetMAS} else {print("MAS DATA NOT
LOADED")}
if(is.object(esetVSN))  {normList$esetVSN<-esetVSN} else {print("VSN DATA NOT
LOADED")}
if(is.object(esetPLM))  {normList$esetPLM<-esetPLM} else {print("PLM DATA NOT
LOADED")}

#CREATING OUTPUT DIRECTORIES
counter<-1
pvalueList<-c(99.9, 99, 95)
for (c in seq(1,3,1))
  {
  base<-getwd()
  RESULT<-paste(base,"/R results/Regulated genes/",sep="")
  dir.create(RESULT, showWarnings = TRUE, recursive = FALSE)
  DIFF<-paste(RESULT,"Differential genes ",pvalueList[c],"/",sep="")
  dir.create(DIFF, showWarnings = TRUE, recursive = FALSE)
  setwd(base)

  #RUNNING NORMALIZED DATA THROUGH LIMMA AND EBAYES
  number2<-length(normList)
  for (d in seq(1,number2,1))
      {
      if (is.object(normList[[d]]))
```

```
{
pval<-((100-pvalueList[c])/100)
pval
nam<-names(normList)
print(nam[d])
RESULT<-paste(DIFF,nam[d],"/",sep="")
dir.create(RESULT, showWarnings = TRUE, recursive = FALSE)
BON<-paste(RESULT,"Bonferonni/",sep="")
dir.create(BON, showWarnings = TRUE, recursive = FALSE)
BENHOCH<-paste(RESULT," Benjamini-Hochberg/",sep="")
dir.create(BENHOCH, showWarnings = TRUE, recursive = FALSE)
nar<-normList[[nam[d]]]
design<-model.matrix(~-1+factor(c(1,1,1,2,2,2,3,3,3,4,4,4)))
colnames(design)<-c("Tug","TuD","Tu2","Tu5")
print("Doing LIMMA")
fit<-lmFit(nar,design)
contrast.matrix<-makeContrasts(TuD-Tug,Tu2-Tug,Tu5-Tug,Tu2-TuD,Tu5-
TuD,Tu5-Tu2,levels=design)
fit2<-contrasts.fit(fit,contrast.matrix)
print("To eBayes")
fit2<-eBayes(fit2)
results<-decideTests(fit2,method="nestedF",p.value=pval,lfc=1)
if (nam[d] == "esetPLM")
    {
    summary(fit2$F.p.value)
 } else {
    index<-grep("AFFX",featureNames(nar))
    summary(fit2$F.p.value[index])
    }
results<-classifyTestsF(fit2,p.value=pval)
res<-summary(results)
filename<-paste(RESULT,nam[d],"_summary.csv",sep="")
write.table(res,file=filename,sep=",",row.names=FALSE,quote=FALSE)
options(digits=3)
UpTotal<-0
DownTotal<-0
compar<-dim(results)
compar<-compar[2]
for (comp in seq(1,compar,1))
   {
   UpReg<-results[results[,comp]==1,]
   DownReg<-results[results[,comp]==-1,]
   UpTotal<-rbind(UpTotal,UpReg)
   DownTotal<-rbind(DownTotal,DownReg)
   }
UpNDown<-UpTotal
filename1<-paste(RESULT,nam[d],"_UP-regulated.csv",sep="")
for (UD in seq(1,2,1))
   {
   UpNDownLength<-dim(UpNDown)
   UpNDownLength<-(UpNDownLength[1])
   UpNDown<-UpNDown[2:UpNDownLength,]
   write.table(UpNDown,file=filename1,sep=",",col.names=NA)
   UpNDown<-DownTotal
   filename1<-paste(RESULT,nam[d],"_DOWN-regulated.csv",sep="")
   }

#Getting the p-values with Toptable
```

```
          comparisons<-c("TuD-Tug","Tu2-Tug","Tu5-Tug","Tu2-TuD","Tu5-TuD","Tu5-
Tu2")
          numberComp<-length(comparisons)
          for (UD in seq(1,numberComp,1))
             {
             name<-comparisons[UD]
             topBon<-
topTable(fit2,coef=UD,number=30000,sort.by="P",adjust="holm")
             #STEP-DOWN BONFERRONI
             filenameB<-paste(BON,name,"_topBon.csv",sep="")
             write.table(topBon,file=filenameB,sep=",",col.names=NA)
             topBenHoch<-
topTable(fit2,coef=UD,number=30000,sort.by="P",adjust="fdr")
             #STEP-DOWN BENJAMINI-HOCHBERG
             filenameB<-paste(BENHOCH,name,"_topBenHoch.csv",sep="")
             write.table(topBenHoch,file=filenameB,sep=",",col.names=NA)
             }
          }
      }
   }
print("Analysis done! Hit enter to return to shell")
```

Target.txt

SlideNumber	FileName	Target	TechRep	BioRep	Name	
1	TUG 1.CEL	1	0	1	Tug1	
2	TUG 2.CEL	1	0	2	Tug2	
3	TUG 3.CEL	1	0	3	Tug3	
4	TUG DN1_1.CEL	2	0	1		TugDn_1
5	TUG DN1_2.CEL	2	0	2		TugDn_2
6	TUG DN1_3.CEL	2	0	3		TugDn_3
7	TUG DN2-2.CEL	3	1	1		TugDn2_1
8	TUG DN2-3.CEL	3	1	1		TugDn2_2
9	TUG DN2-4.CEL	3	0	2		TugDn2_3
10	TUG DN5-1.CEL	4	0	1		TugDn5_1
11	TUG DN5-2.CEL	4	0	2		TugDn5_2
12	TUG DN5-3.CEL	4	0	3		TugDn5_3

Convert

```
Option Explicit
Sub Convert()

Dim ws As Worksheet
Dim myDir As String
Dim fn As String
Dim Dirr As String
myDir = ThisWorkbook.Path & "\"
Dirr = (myDir & "*.csv")
fn = Dir(Dirr)

If fn = "" Then
 MsgBox "no file"
Exit Sub
End If
```

```
Do While fn <> ""
 With Workbooks.Open(myDir & fn)
  For Each ws In .Worksheets
   ActiveSheet.Copy Before:=ThisWorkbook.Sheets(1)
  Next
 Close False
End With
 fn = Dir
Loop

End Sub
```

Data visualization before normalization and background correcting

The data in this subsection, produced by Script 1 (*Table Appx 5.1*), can be found in the folder: */R results/Quality control raw data/* and its sub folders.

Figure Appx 5.1 **- 12 Slides**

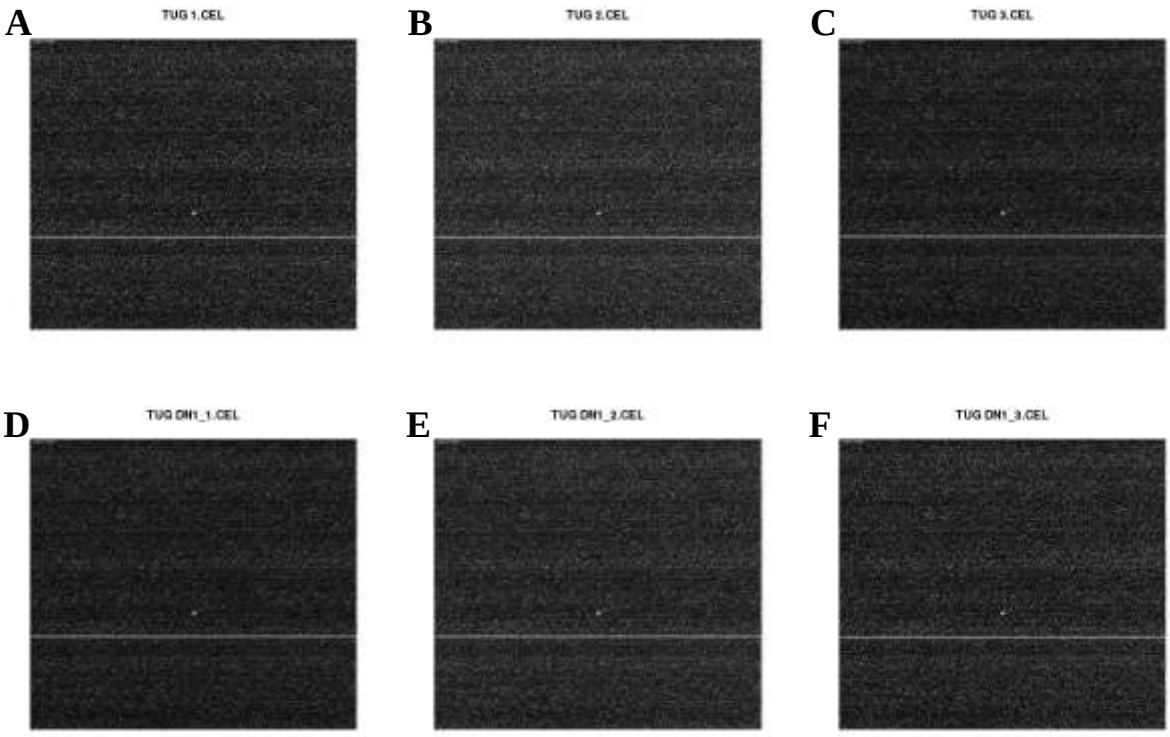
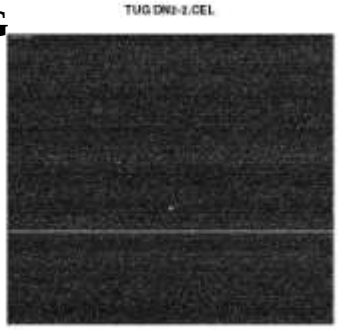
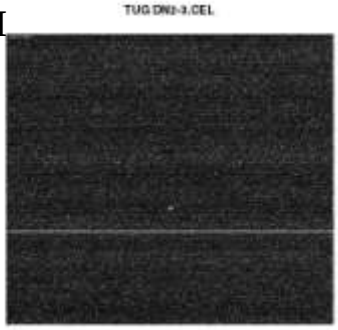


Figure Appx 5.1 **- 12 Slides cont.**

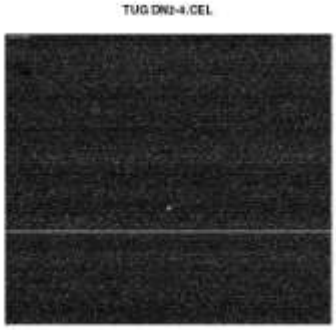
G



H

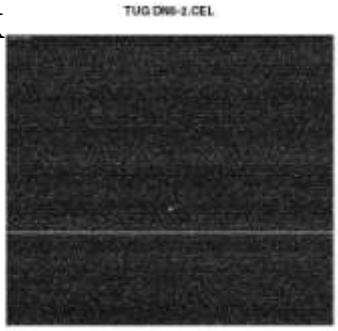


I



J

K



L

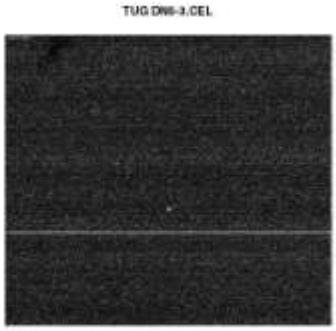
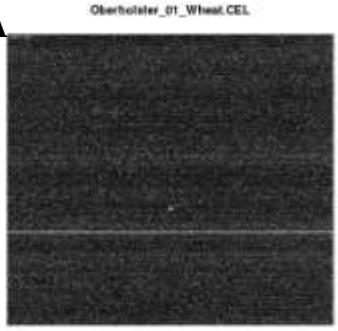
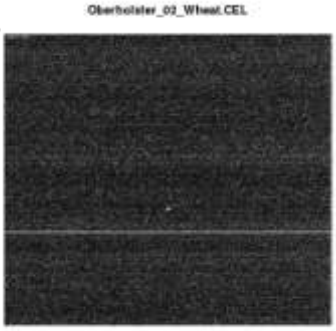


Figure Appx 5.1 **- 18 Slides**

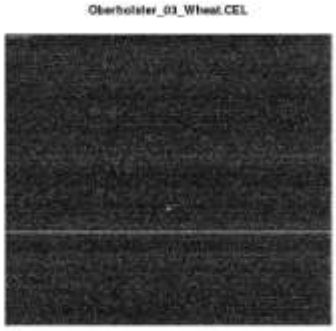
A



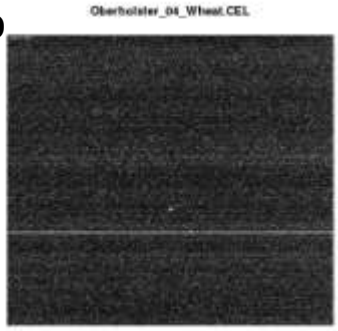
B



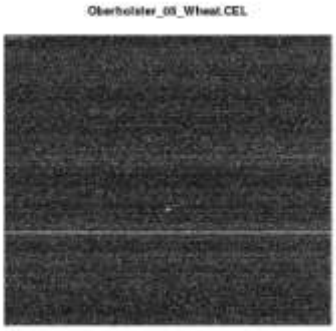
C



D



E



F

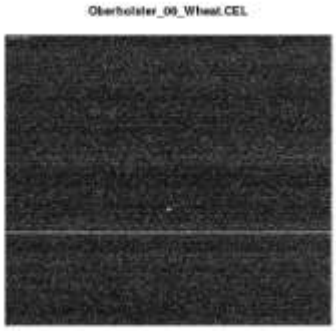


Figure Appx 5.1 **- 18 Slides cont.**

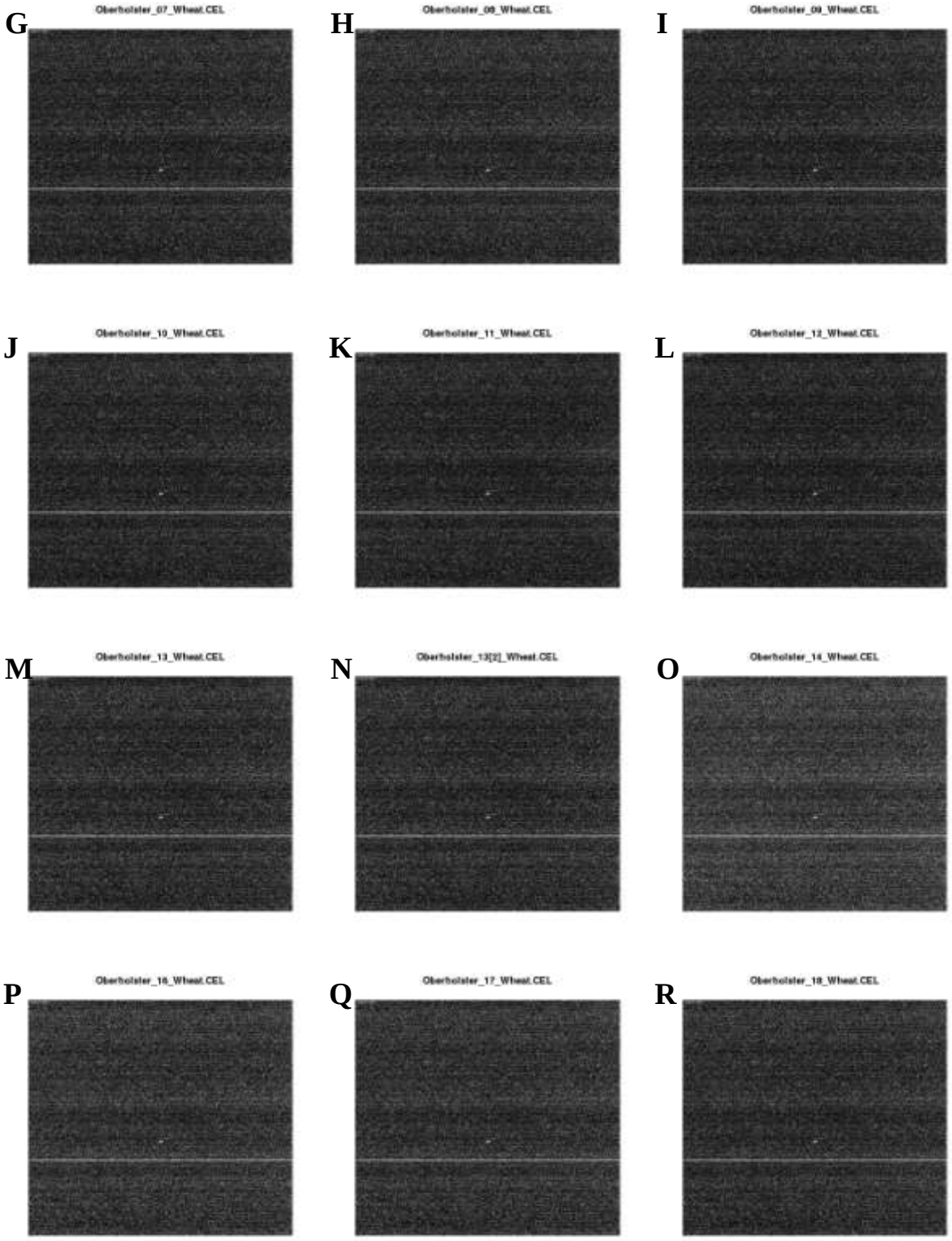


Figure Appx 5.1 Images of the scanned Affymetrix slides after hybridization with the different treatments. The images were produced from the *CEL* files. The 12 slides images were the originally produced *CEL* images, while the 18 slides images are the $\log_2$ derivatives of the original 18 slides

images. Both types were produced by script 1 and saved in */R results/Quality control raw data/Slide Images raw/*. The filenames for the 12 slides are: (A) Tug1_raw.jpg, (B) Tug2_raw.jpg, (C) Tug3_-raw.jpg, (D) TugDn_1_raw.jpg, (E) TugDn_2_raw.jpg, (F) TugDn_3_raw.jpg, (G) TugDn2_1_-raw.jpg, (H) TugDn2_2_raw.jpg, (I) TugDn2_3_raw.jpg, (J) TugDn5_1_raw.jpg, (K) TugDn5_2_-raw.jpg, and (L) TugDn5_3_raw.jpg.

The log_2 derivatives filenames for the 18 slides are: (A) RWA1_GS_3_log2_raw.jpg, (B) RWA1_GS_2_log2_raw.jpg, (C) RWA1_GS_1_log2_raw.jpg, (D) RWA2_GS_1_log2_raw.jpg, (E) RWA2_GS_2_log2_raw.jpg, (F) RWA2_GS_3_log2_raw.jpg, (G) RWA1_GR_1_log2_raw.jpg, (H) RWA1_GR_2_log2_raw.jpg, (I) RWA1_GR_3_log2_raw.jpg, (J) RWA2_GR_1_log2_raw.jpg, (K) RWA2_GR_2_log2_raw.jpg, (L) RWA2_GR_3_log2_raw.jpg, (M) Gam_S_1_log2_raw.jpg, (N) Gam_S_2_log2_raw.jpg, (O) Gam_S_3_log2_raw.jpg, (P) Gam_R_1_log2_raw.jpg, (Q) Gam_R_2_log2_raw.jpg, (R) Gam_R_3_log2_raw.jpg.

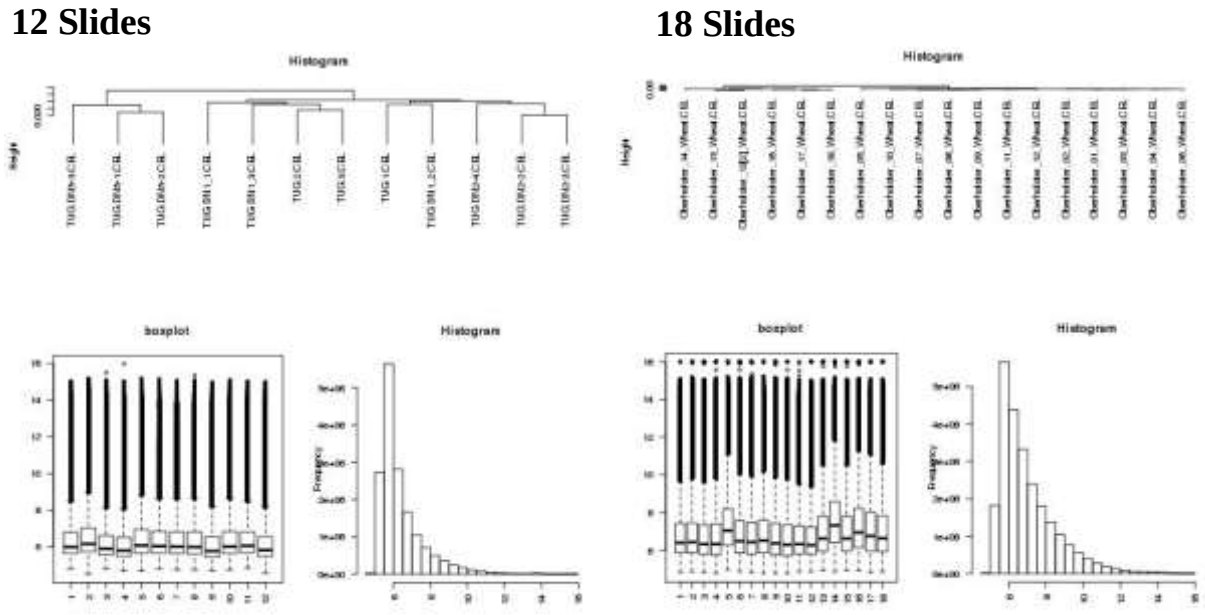


Figure Appx 5.2 Summary figures, including histograms and boxplots, of all the slide data before normalization for the 12 slide and 18 slide experiment. These figures were produced with script 1 and saved as *Summary of all data.jpg.*

Figure Appx 5.3 **- 12 Slides**

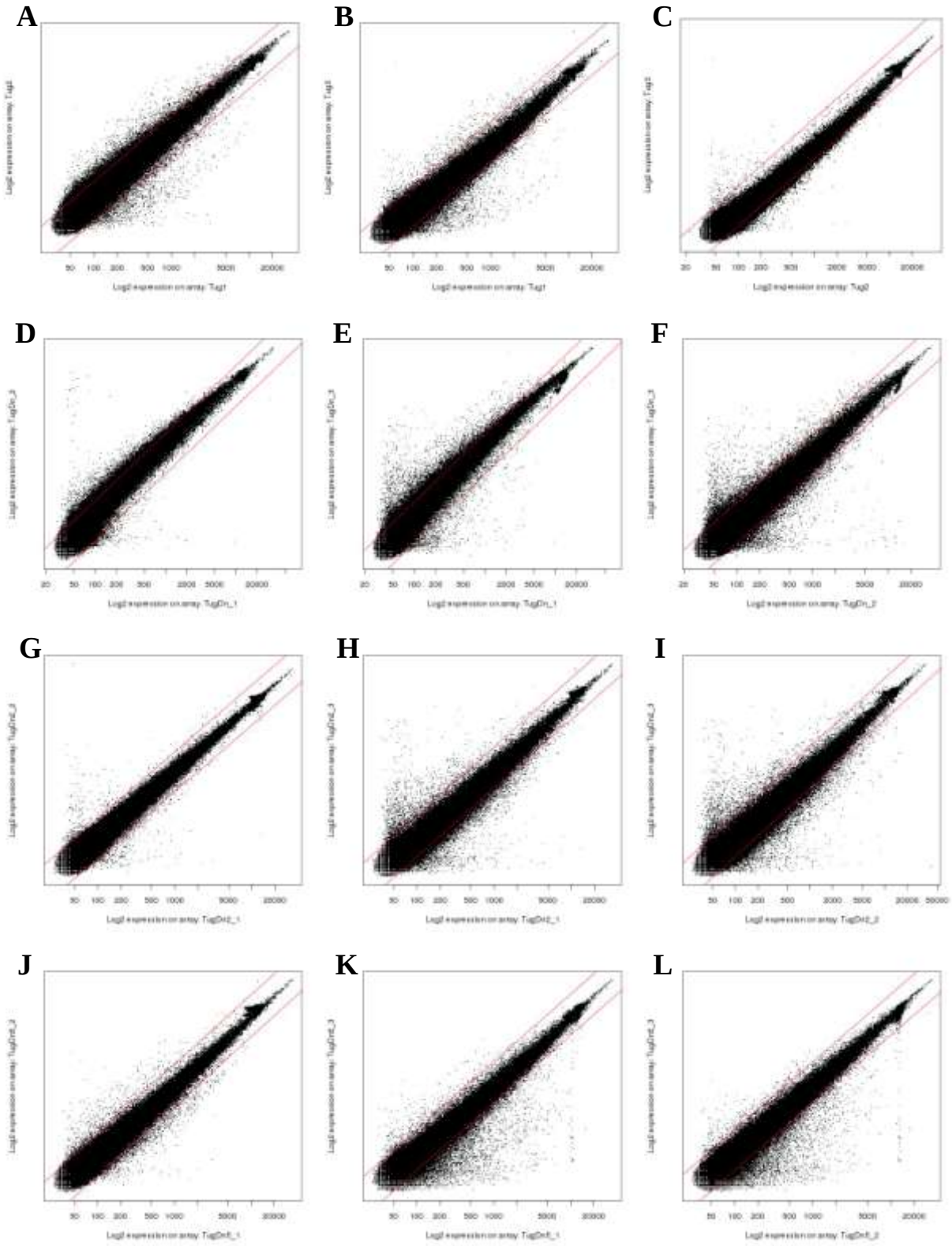


***Figure Appx 5.3* - 18 Slides**

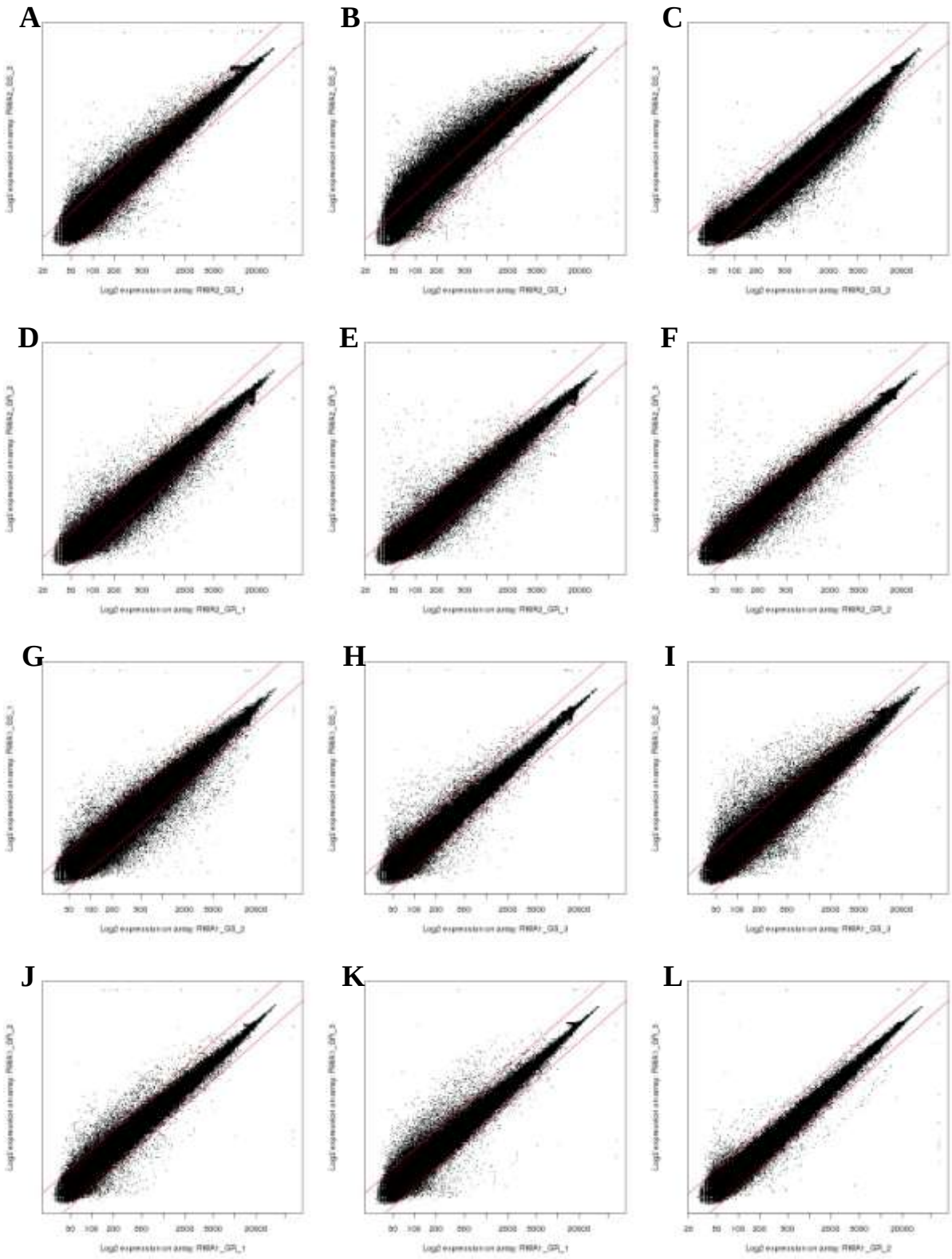


***Figure Appx 5.3* - 18 Slides cont.**

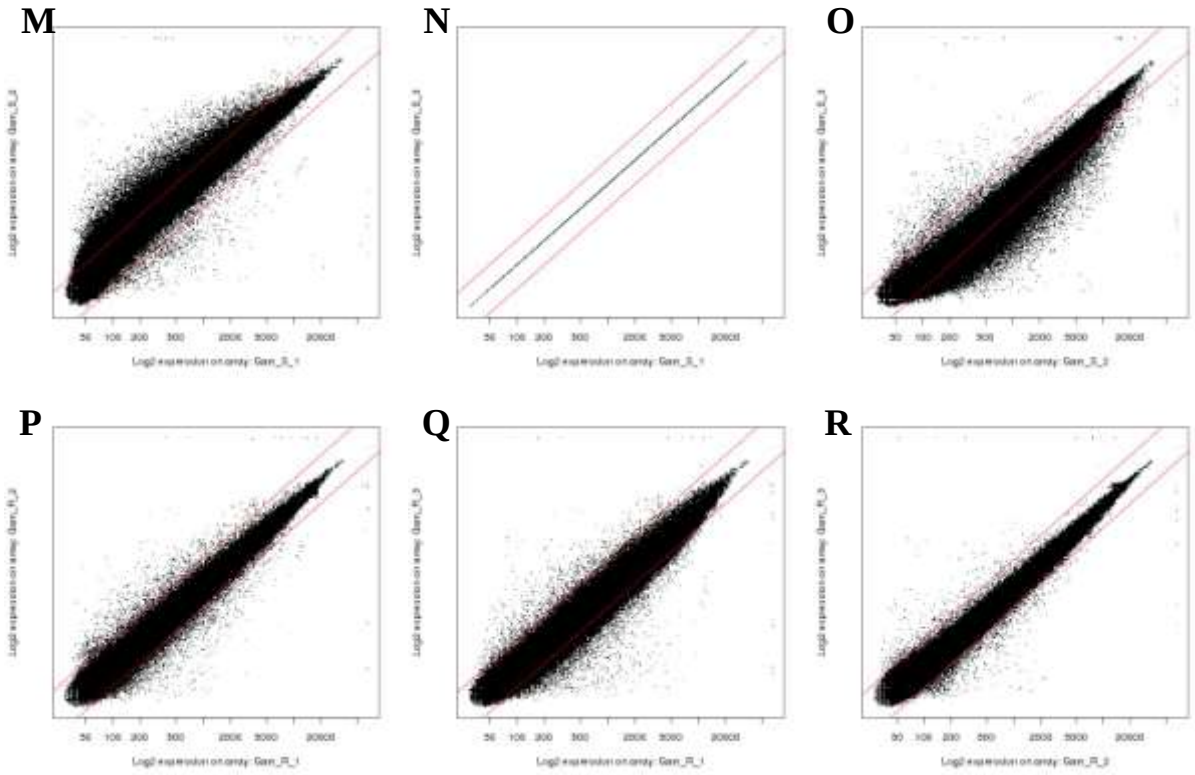


Figure Appx 5.3 The $\log_2$ expression comparisons of the raw, pre-normalized data within treatments. The 12 slides files are: (A) Tug1_vs_Tug2_log2.jpg, (B) Tug1_vs_Tug3_log2.jpg, (C) Tug2_vs_Tug3_log2.jpg, (D) TugDn_1_vs_TugDn_2_log2.jpg, (E) TugDn_1_vs_TugDn_3_log2.-jpg, (F) TugDn_2_vs_TugDn_3_log2.jpg, (G) TugDn2_1_vs_TugDn2_2_log2.jpg, (H) TugDn2_-1_vs_TugDn2_3_log2.jpg, (I) TugDn2_2_vs_TugDn2_3_log2.jpg, (J) TugDn5_1_vs_TugDn5_2_-log2.jpg, (K) TugDn5_1_vs_TugDn5_3_log2.jpg, (L) TugDn5_2_vs_TugDn5_3_log2.jpg.

The 18 slides files are: (A) RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (B) RWA2_GS_1_-vs_RWA2_GS_3_log2.jpg, (C) RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (D) RWA2_GR_1_vs_-RWA2_GR_2_log2.jpg, (E) RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (F) RWA2_GR_2_-vs_RWA2_GR_3_log2.jpg, (G) RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (H) RWA1_GS_3_-vs_RWA1_GS_1_log2.jpg, (I) RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (J) RWA1_GR_1_-vs_RWA1_GR_2_log2.jpg, (K) RWA1_GR_1_vs_RWA1_GR_3_log2.jpg, (L) RWA1_GR_2_-vs_RWA1_GR_3_log2.jpg, (M) Gam_S_1_vs_Gam_S_2_log2.jpg, (N) Gam_S_1_vs_Gam_S_3_-log2.jpg, (O) Gam_S_2_vs_Gam_S_3_log2.jpg, (P) Gam_R_1_vs_Gam_R_2_log2.jpg, (Q) Gam_-R_1_vs_Gam_R_3_log2.jpg, (R) Gam_R_2_vs_Gam_R_3_log2.jpg.

Figure Appx 5.4 **- 12 Slides**

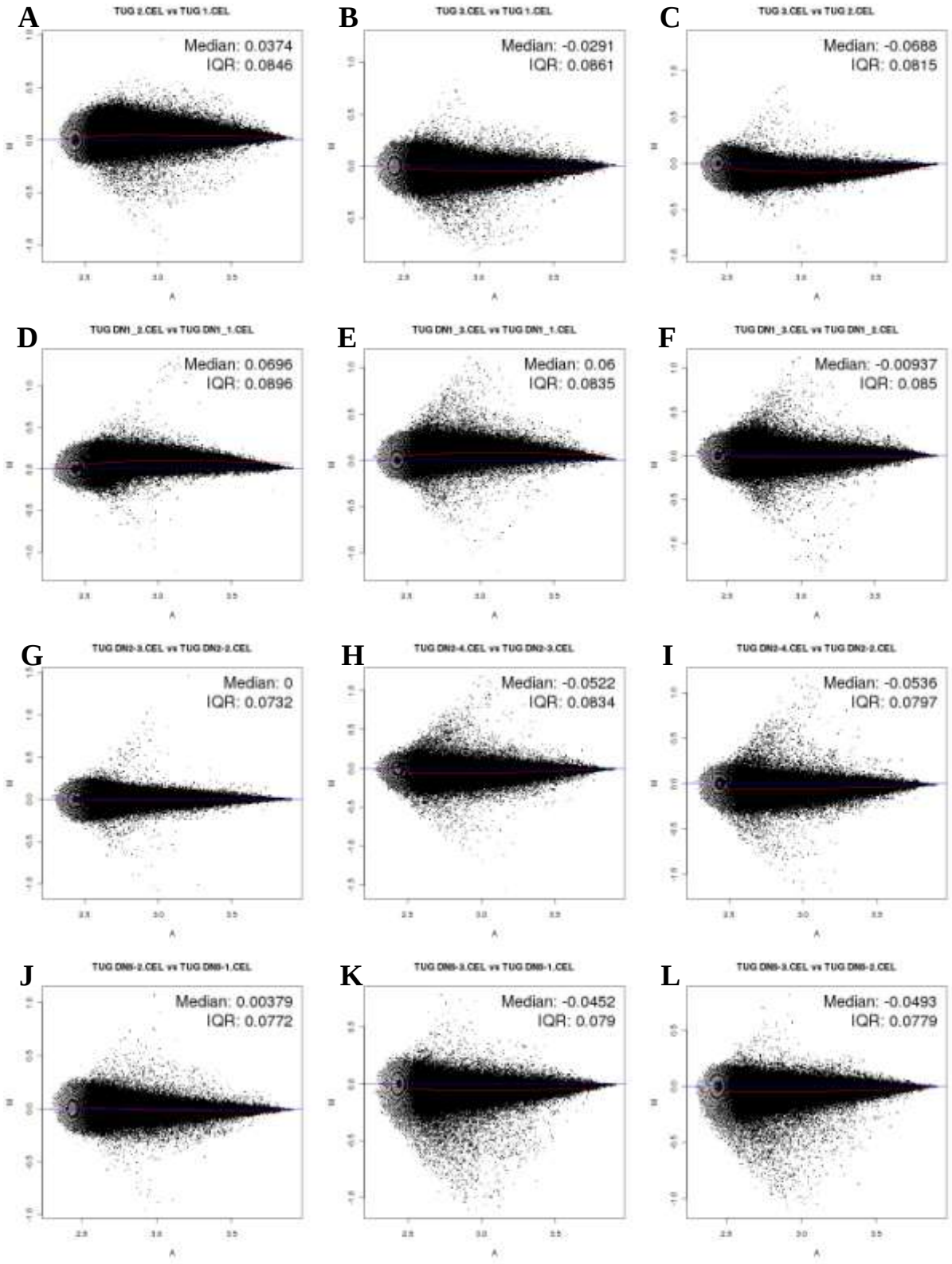


Figure Appx 5.4 - 18 Slides

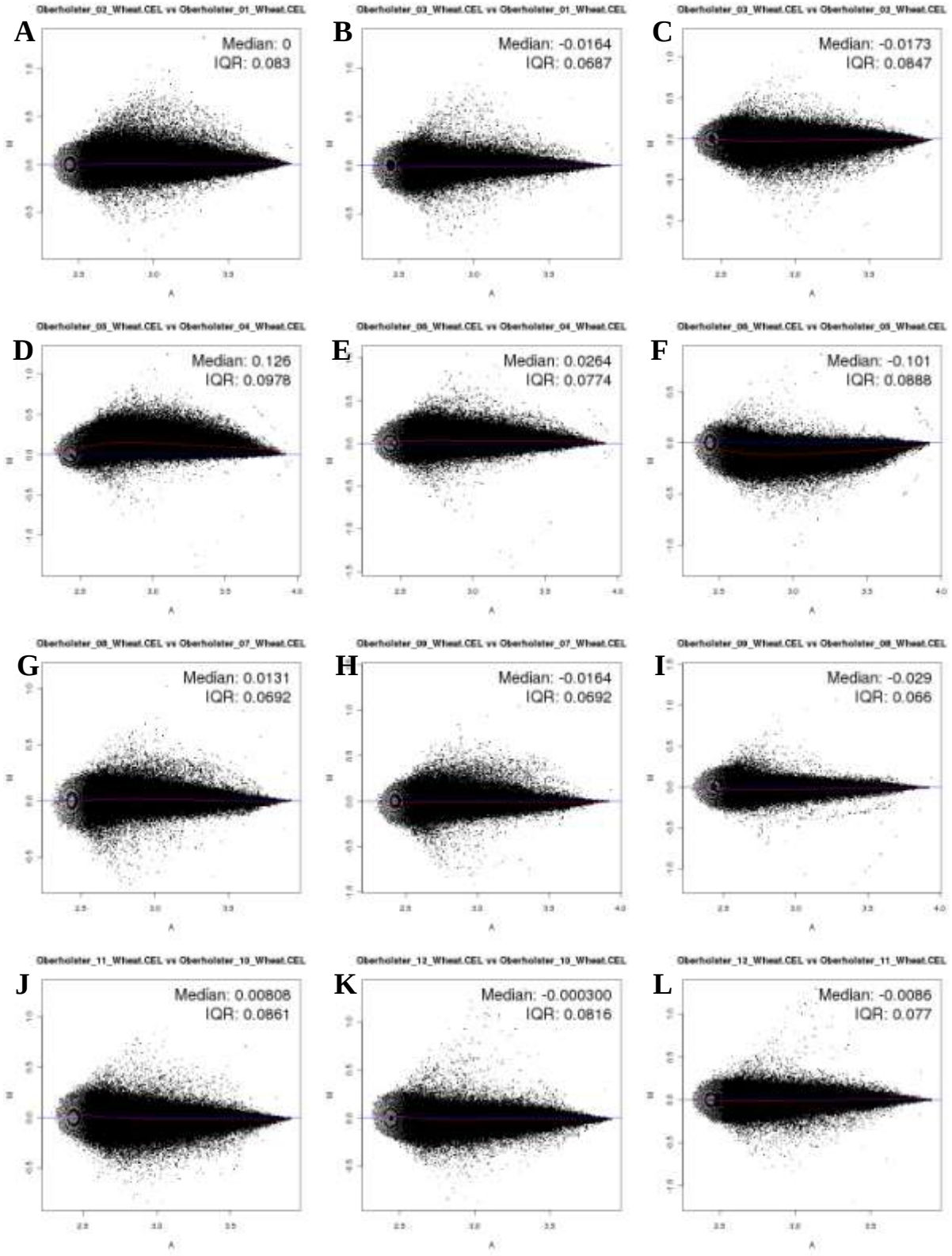


***Figure Appx 5.4* - 18 Slides cont.**

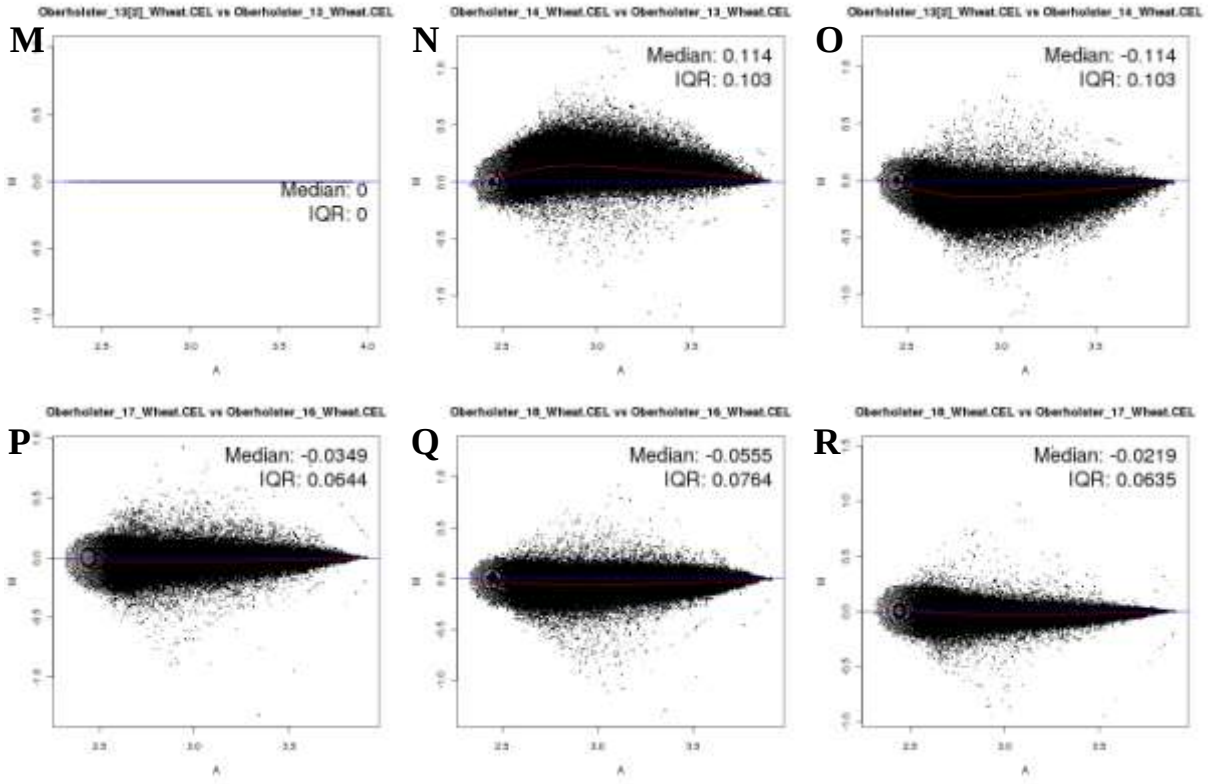


Figure Appx 5.4 MA-plots of the raw, pre-normalized slide comparisons within treatments. The 12 slides files are: (A) Tug1_vs_Tug2_raw.jpg, (B) Tug1_vs_Tug3_raw.jpg, (C) Tug2_vs_Tug3_-raw.jpg, (D) TugDn_1_vs_TugDn_2_raw.jpg, (E) TugDn_1_vs_TugDn_3_raw.jpg, (F) TugDn_2_-vs_TugDn_3_raw.jpg, (G) TugDn2_1_vs_TugDn2_2_raw.jpg, (H) TugDn2_1_vs_TugDn2_3_-raw.jpg, (I) TugDn2_2_vs_TugDn2_3_raw.jpg, (J) TugDn5_1_vs_TugDn5_2_raw.jpg, (K) TugDn5_1_vs_TugDn5_3_raw.jpg, (L) TugDn5_2_vs_TugDn5_ 3_raw.jpg.

The 18 slides files are: (A) RWA2_GS_1_vs_RWA2_GS_2_raw.jpg, (B) RWA2_GS_1_vs_-RWA2_GS_3_raw.jpg, (C) RWA2_GS_2_vs_RWA2_GS_3_raw.jpg, (D) RWA2_GR_1_vs_-RWA2_GR_2_raw.jpg, (E) RWA2_GR_1_vs_RWA2_GR_3_raw.jpg, (F) RWA2_GR_2_vs_-RWA2_GR_3_raw.jpg, (G) RWA1_GR_1_vs_RWA1_GR_2_raw.jpg, (H) RWA1_GR_1_vs_-RWA1_GR_3_raw.jpg, (I) RWA1_GR_2_vs_RWA1_GR_3_raw.jpg, (J) RWA1_GS_2_vs_-RWA1_GS_1_raw.jpg, (K) RWA1_GS_3_vs_RWA1_GS_1_raw.jpg, (L) RWA1_GS_3_vs_-RWA1_GS_2_raw.jpg, (M) Gam_S_1_vs_Gam_S_2_raw.jpg, (N) Gam_S_1_vs_Gam_S_3_-raw.jpg, (O) Gam_S_2_vs_Gam_S_3_raw.jpg, (P) Gam_R_1_vs_Gam_R_2_raw.jpg, (Q) Gam_R_-1_vs_Gam_R_3_raw.jpg, (R) Gam_R_2_vs_Gam_ R_3_raw.jpg.

Data visualization after background correcting and normalization

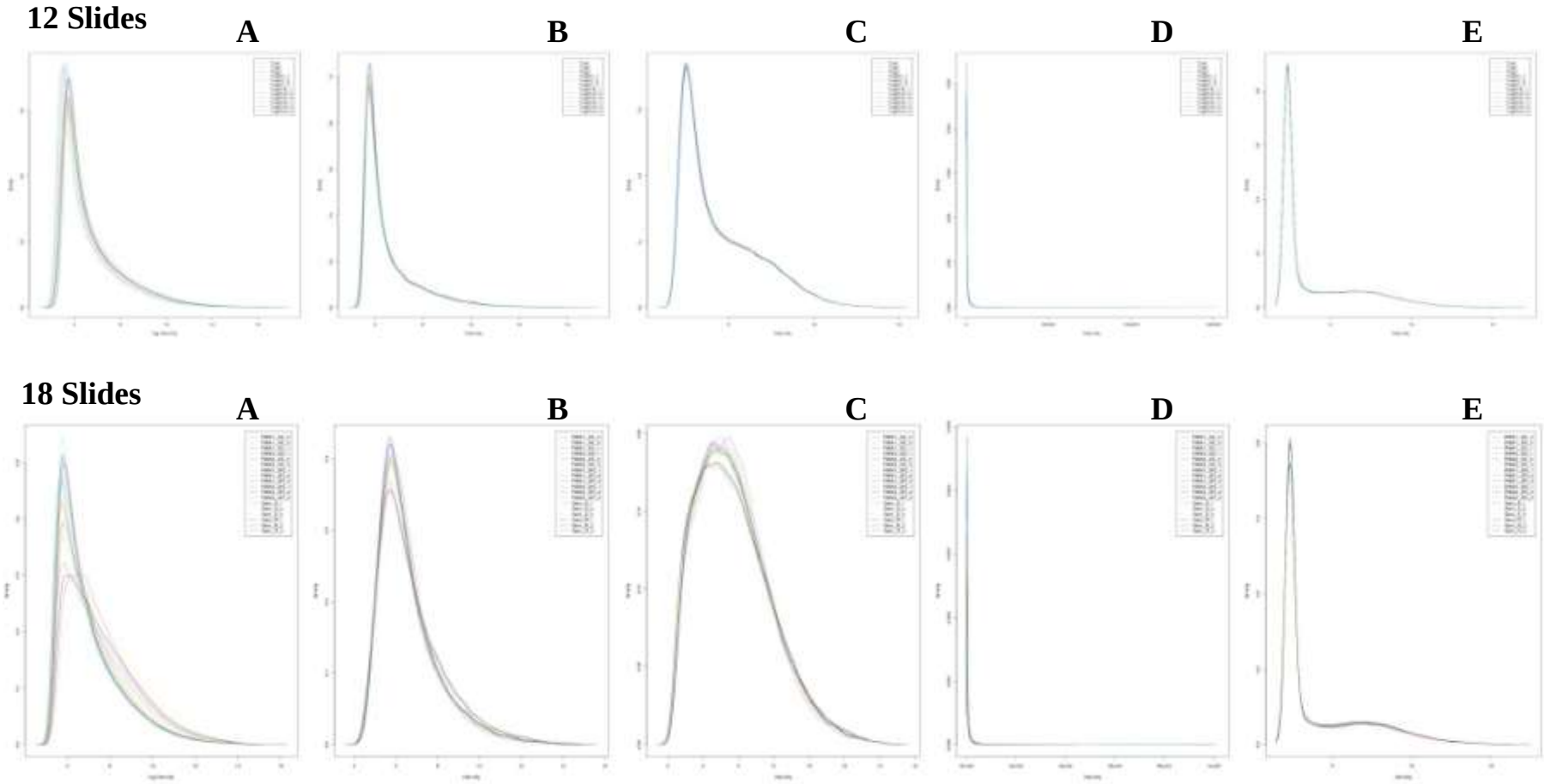


Figure Appx 5.5 Histograms of the slides before ((A) Histogram_rawdata.jpg) and after normalization ((B) Histogram_esetVSN.jpg, (C) Histogram_esetRMA.jpg, (D) Histogram_esetMAS.jpg, (E) Histogram_eset GCRMA.jpg) for the 12 Slide and 18 Slide experiments.

18 Slides

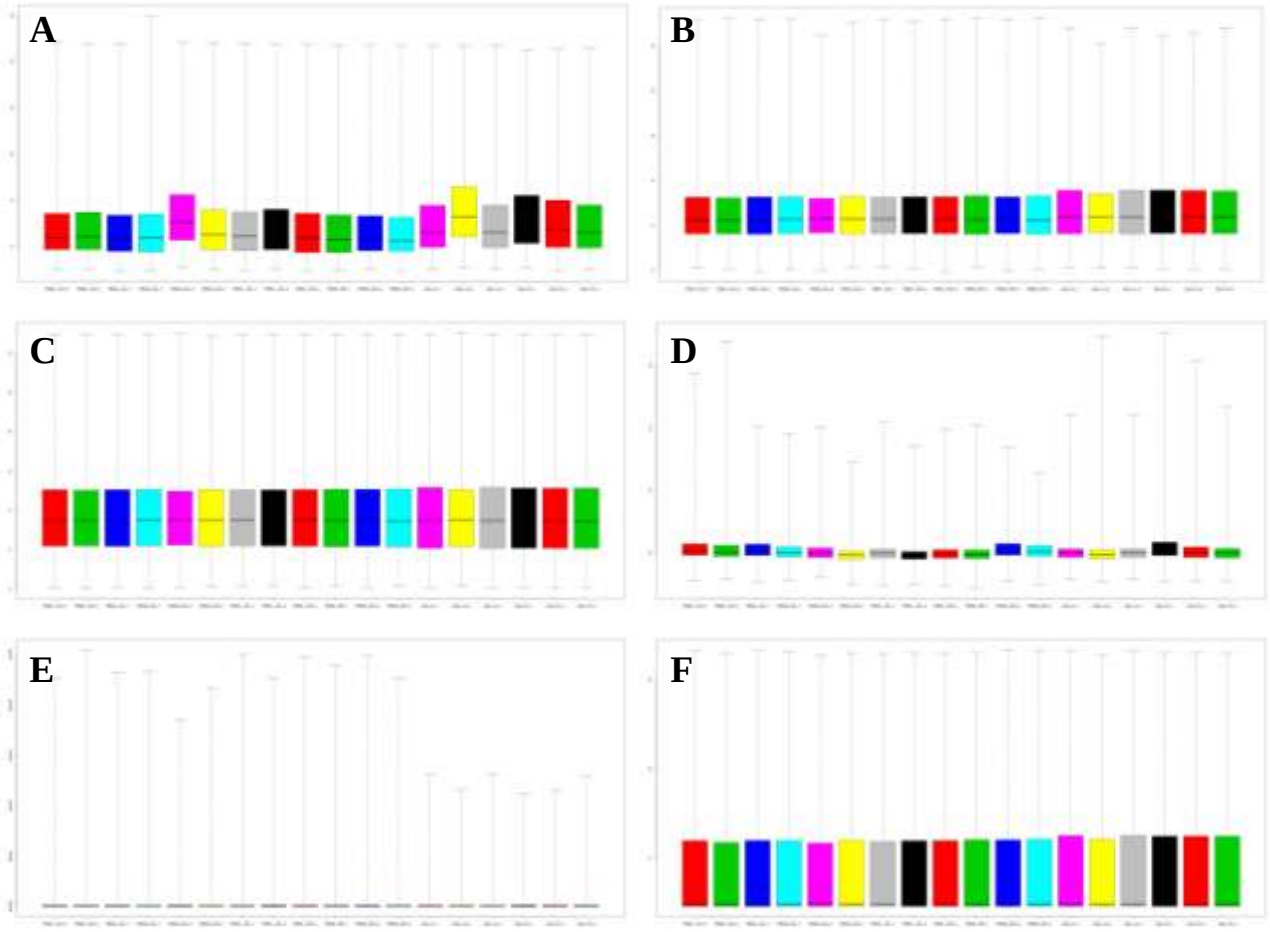


Figure Appx 5.6 Box plots of the $\log_2$(PM) values, before ((A) Boxplot_rawdata.jpg) and after normalization ((B) Boxplot_esetVSN.jpg, (C) Boxplot_esetRMA.jpg, (D) Boxplot_esetPLM.jpg, (E) Boxplot_esetMAS.jpg, (F) Boxplot_esetGCRMA.jpg) for the 12 Slide (see *Figure 5.4*) and the 18 Slide experiment using different statistical approaches. All the sample slides occur in the same order in the various graphs in their corresponding triplicate slides, and are: RWA1_GS_3, RWA1_GS_2, RWA1_GS_1, RWA2_GS_1, RWA2_GS_2, RWA2_GS_3, RWA1_GR_1, RWA1_GR_2, RWA1_GR_3, RWA2_GR_1, RWA2_GR_2, RWA2_GR_3, Gam_S_1, Gam_S_2, Gam_S_3, Gam_R_1, Gam_R_2, Gam_R_3. The Y-axis differs in accordance to the method used.

***Figure Appx 5.7* - 12 Slides**

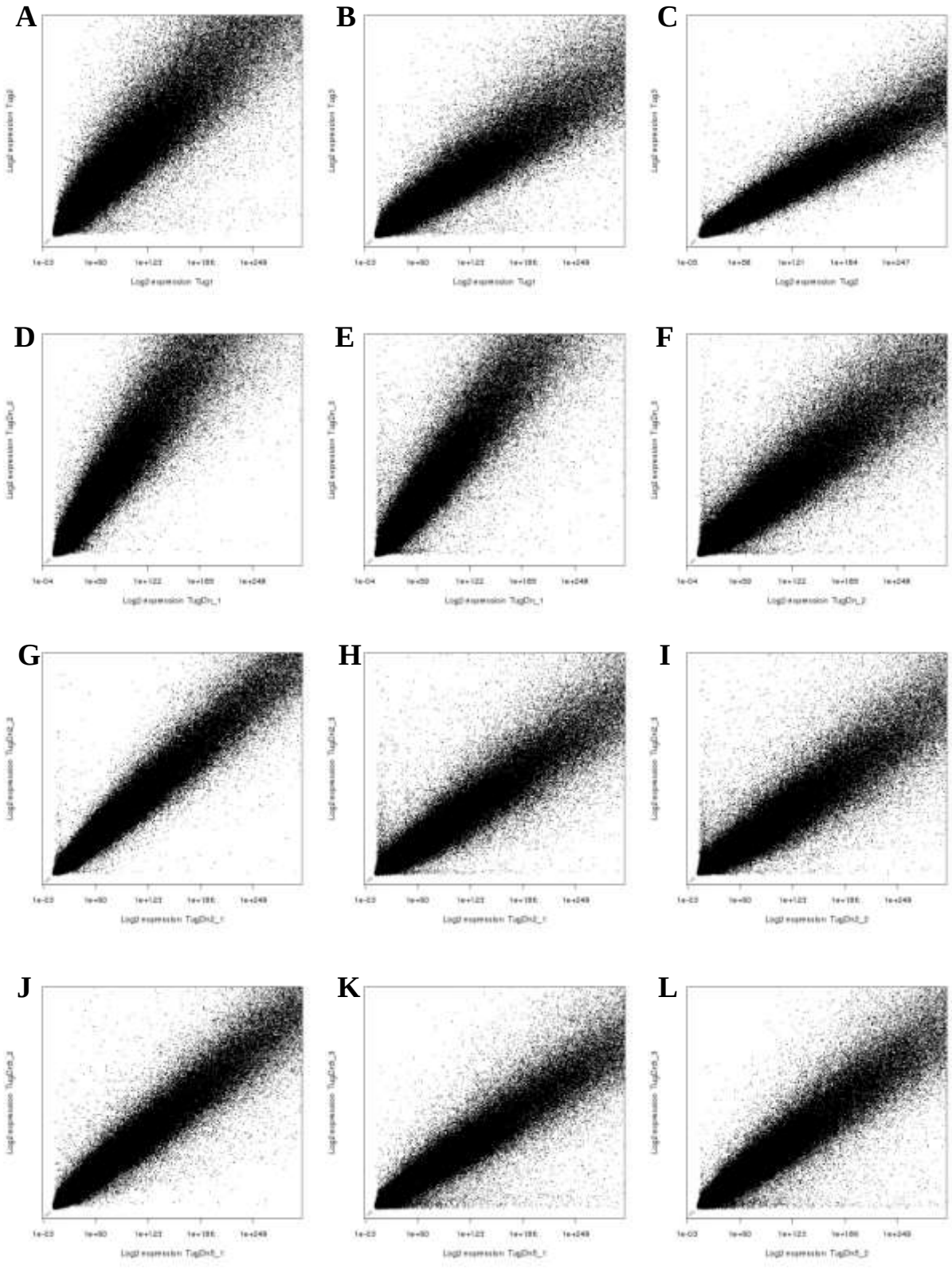


***Figure Appx 5.7* - 12 Slides cont.**

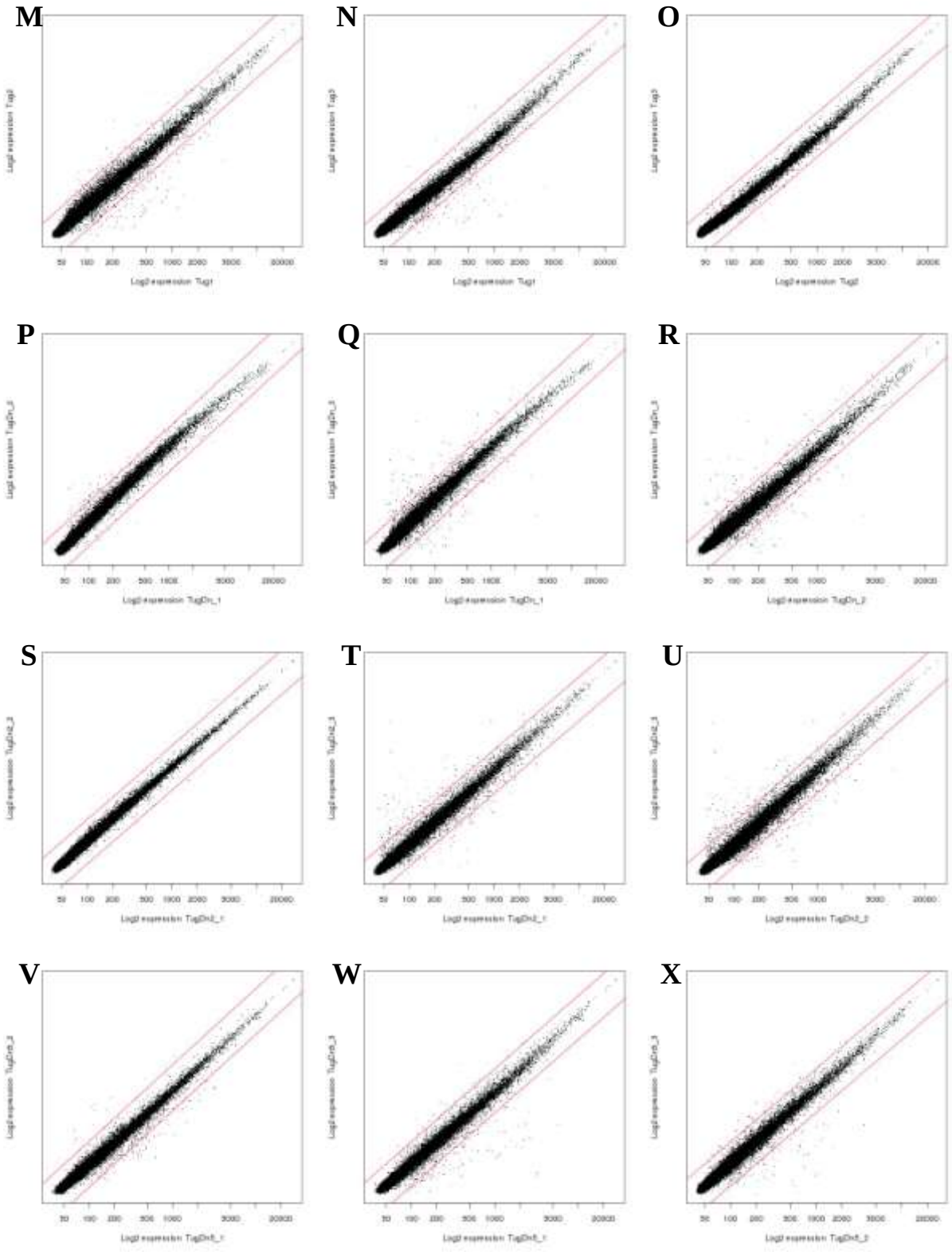
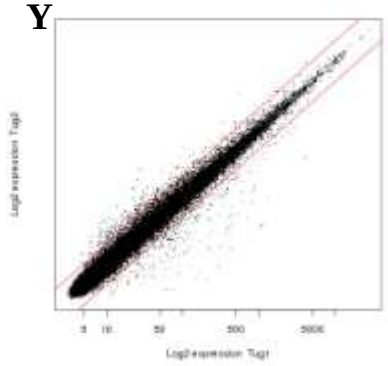
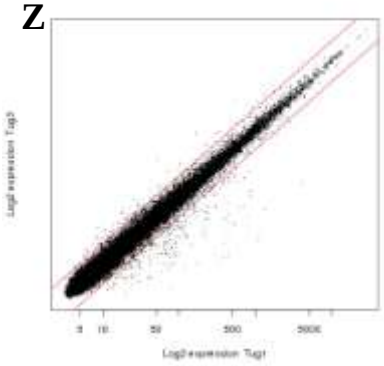


***Figure Appx 5.7* - 12 Slides cont.**

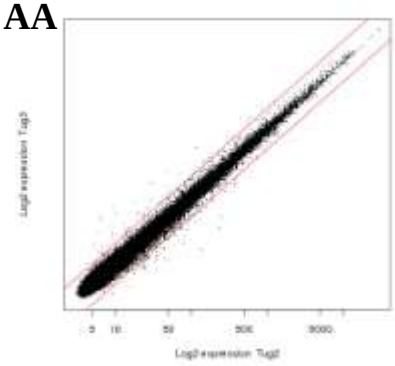
Y



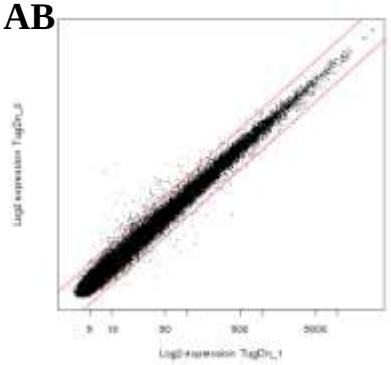
Z



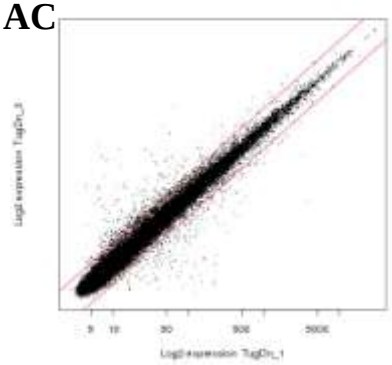
AA



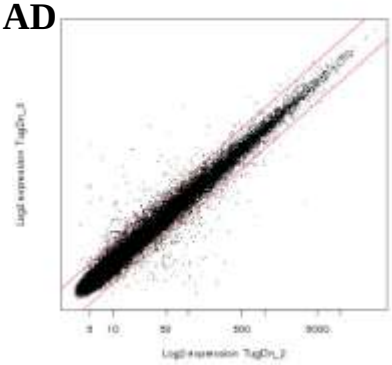
AB



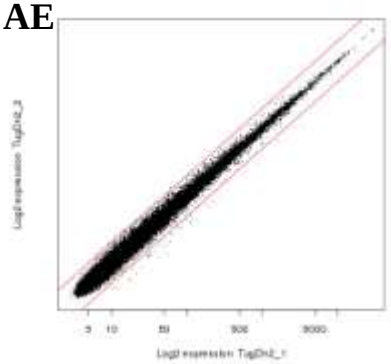
AC



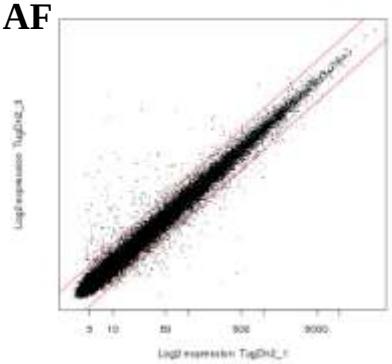
AD



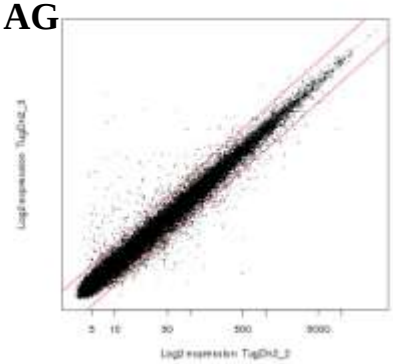
AE



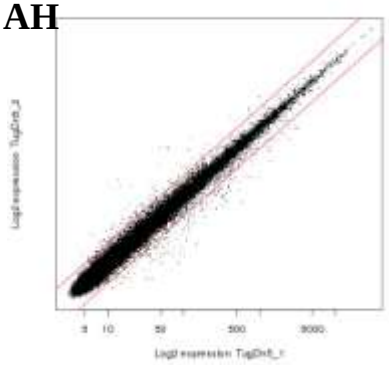
AF



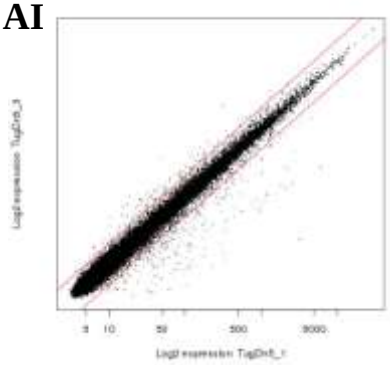
AG



AH



AI



AJ

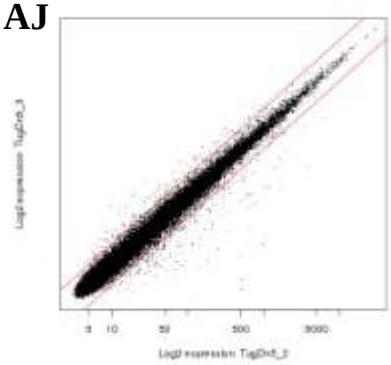
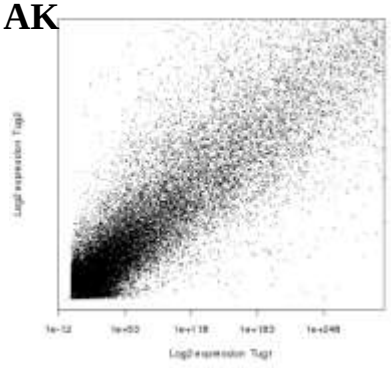
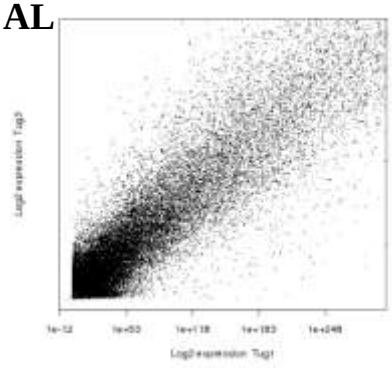


***Figure Appx 5.7* - 12 Slides cont.**

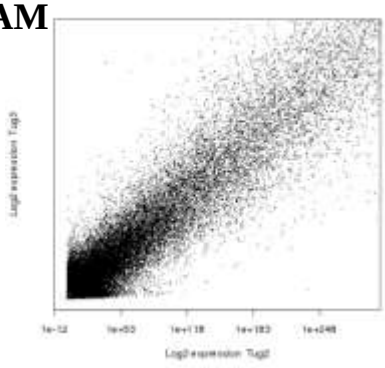
AK



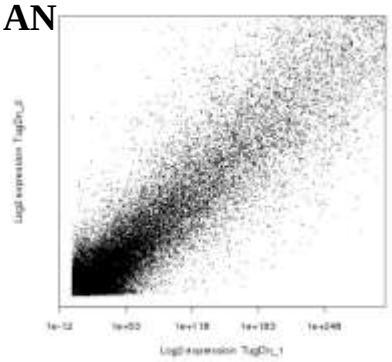
AL



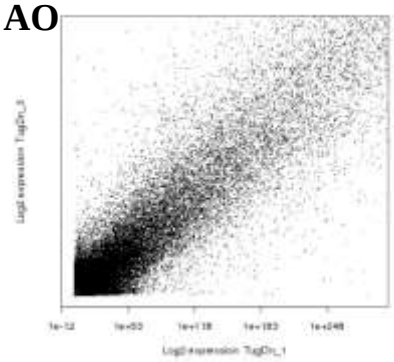
AM



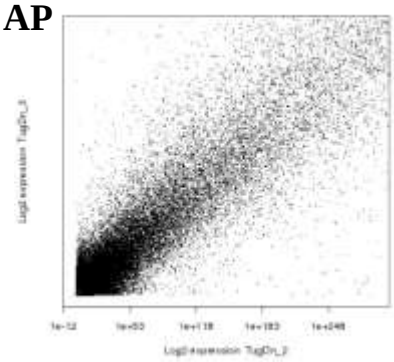
AN



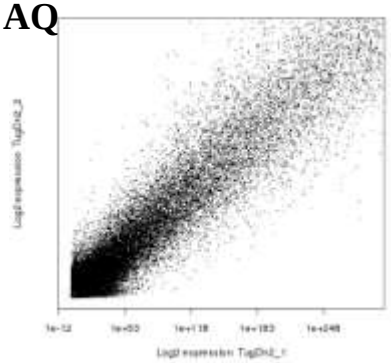
AO



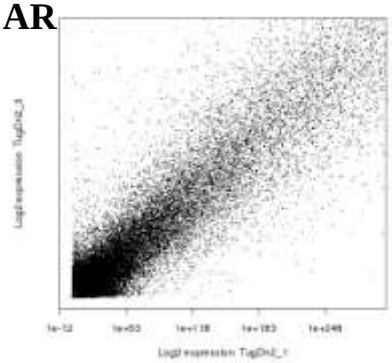
AP



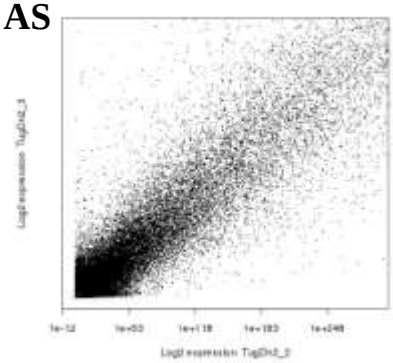
AQ



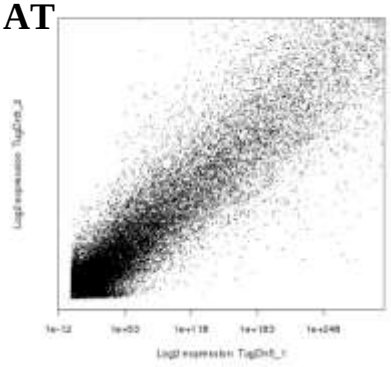
AR



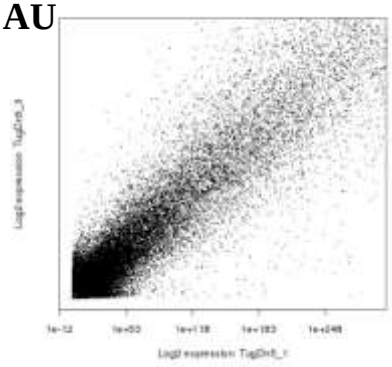
AS



AT



AU



AV

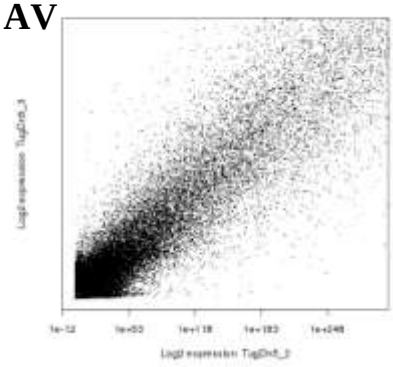
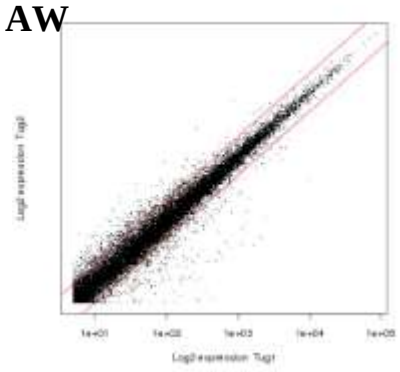
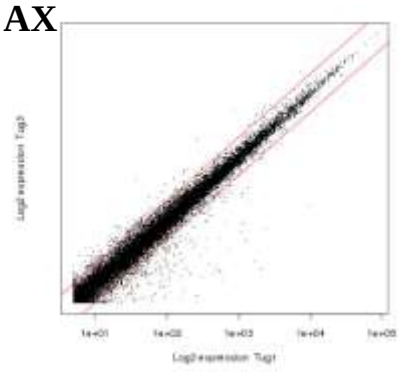


***Figure Appx 5.7* - 12 Slides cont.**

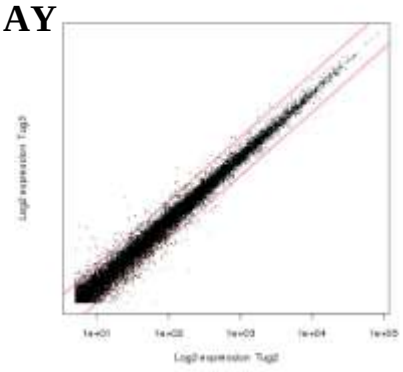
AW



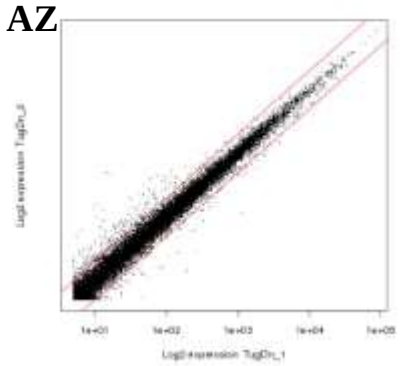
AX



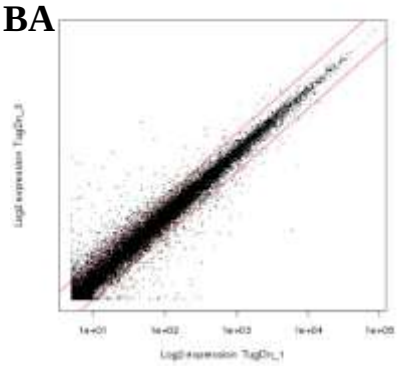
AY



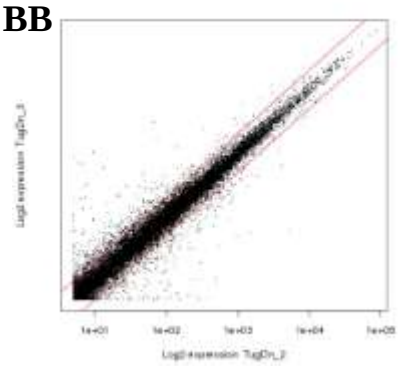
AZ



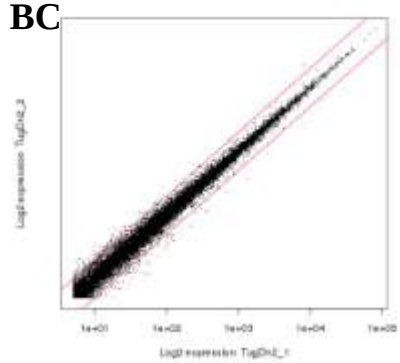
BA



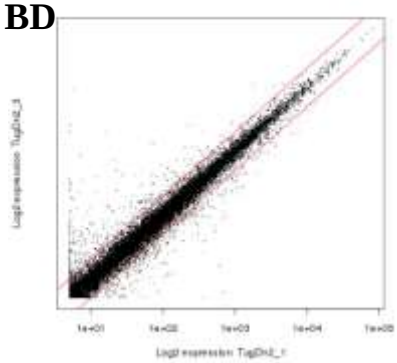
BB



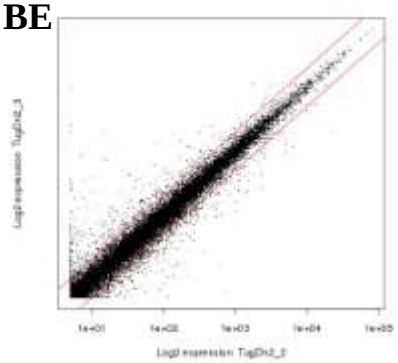
BC



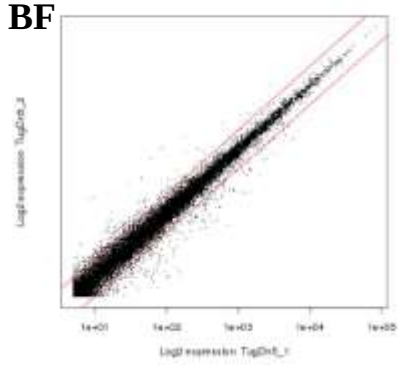
BD



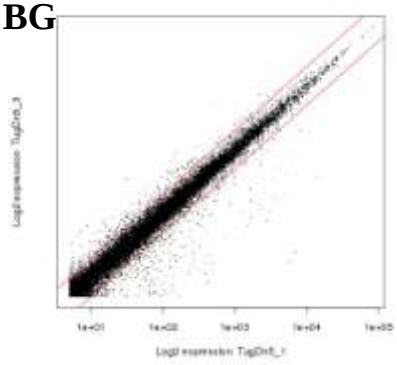
BE



BF



BG



BH

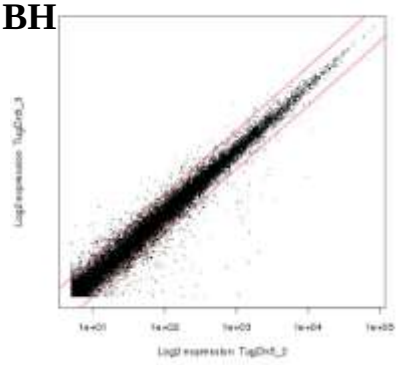


***Figure Appx 5.7* - 18 Slides**

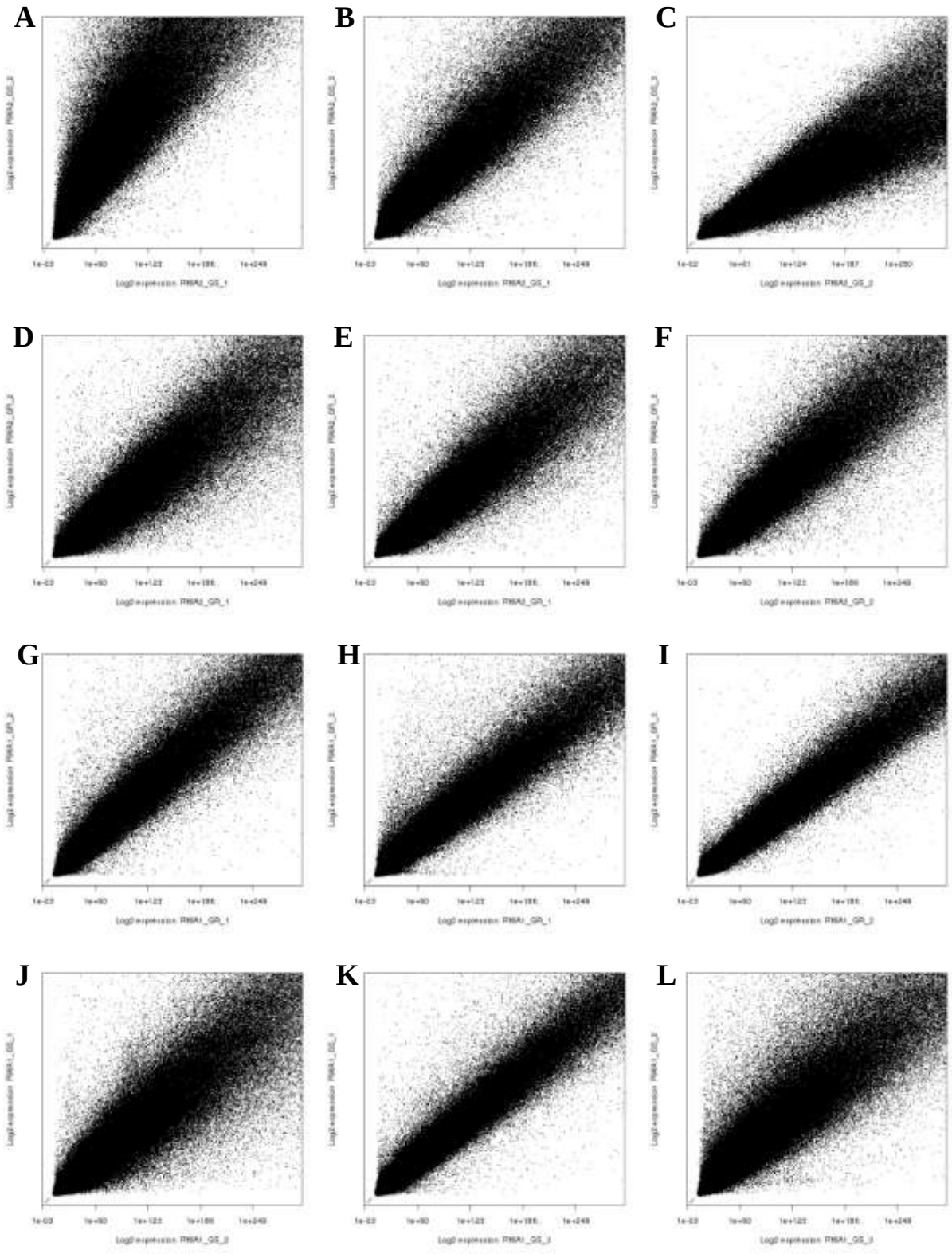


***Figure Appx 5.7* - 18 Slides cont.**

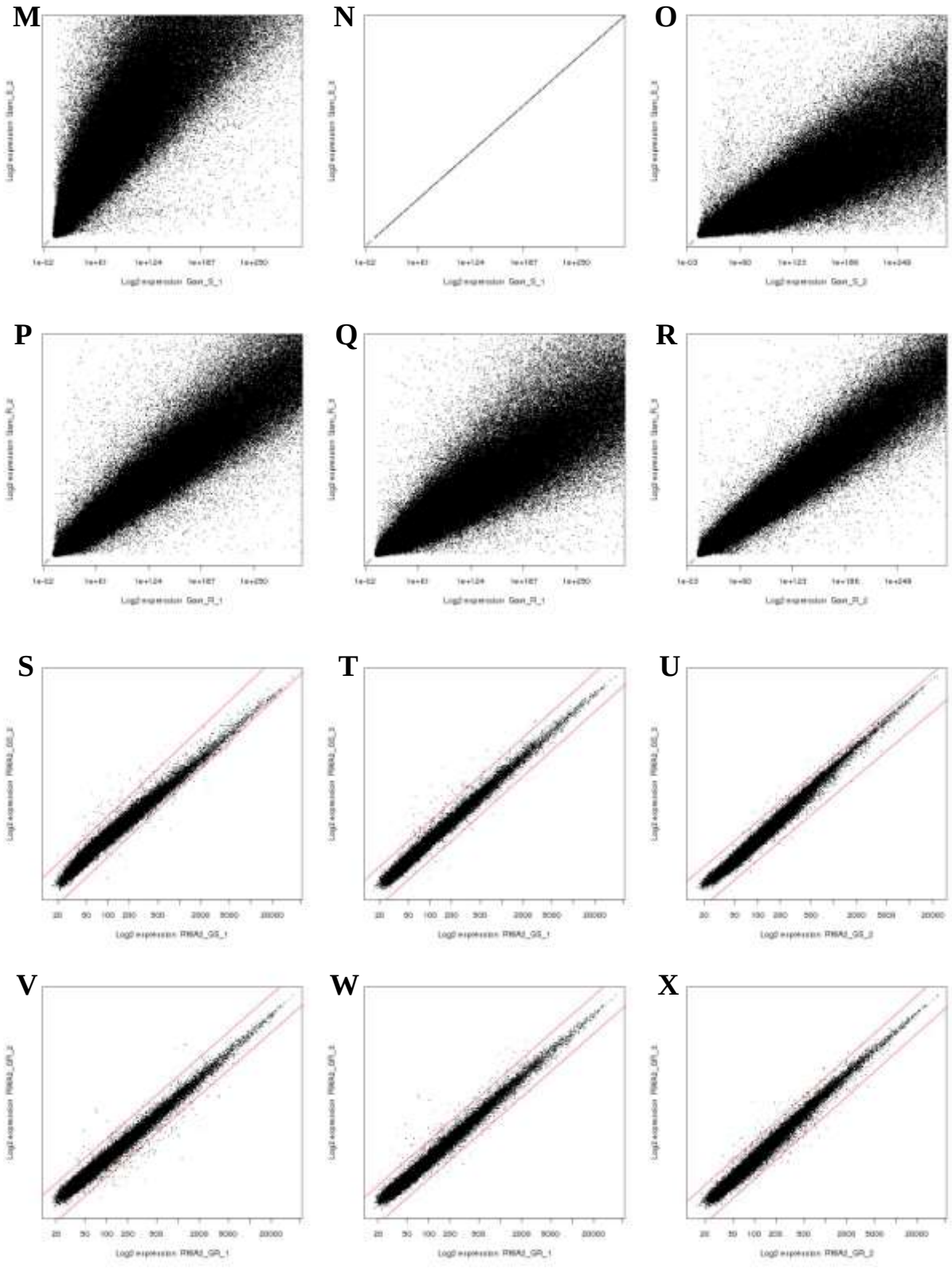
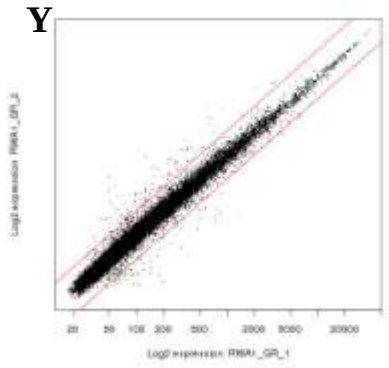
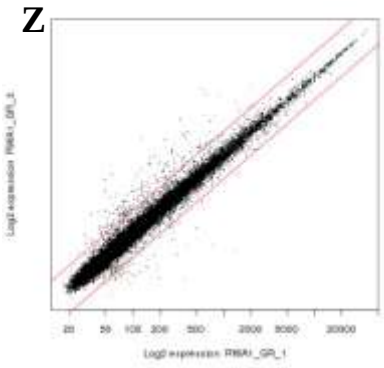


***Figure Appx 5.7* - 18 Slides cont.**

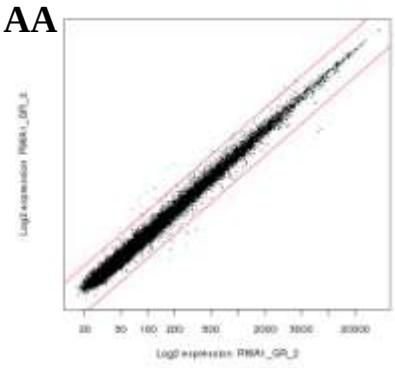
Y



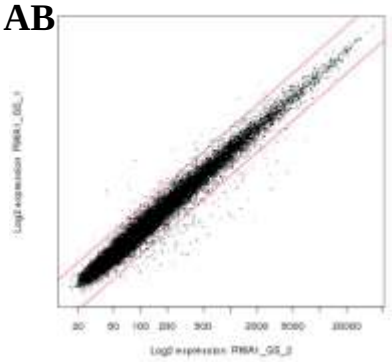
Z



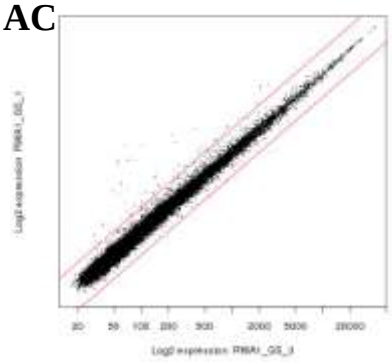
AA



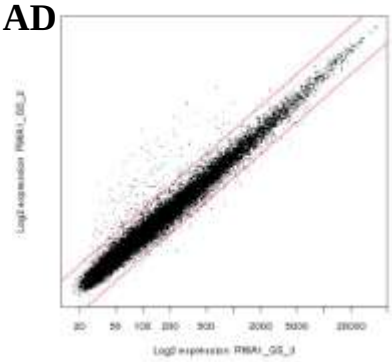
AB



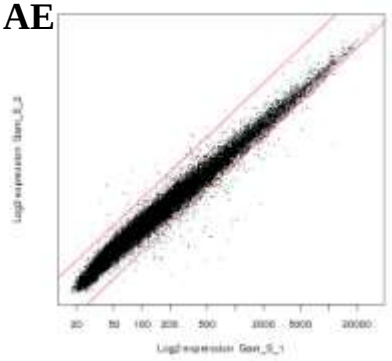
AC



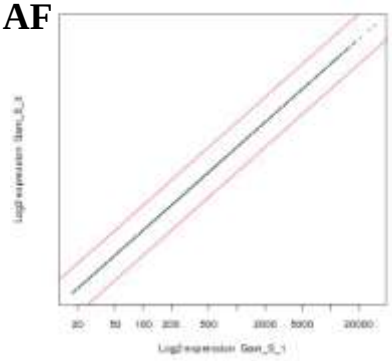
AD



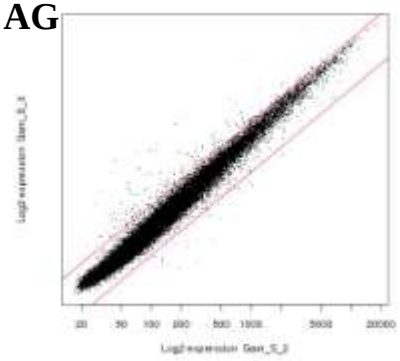
AE



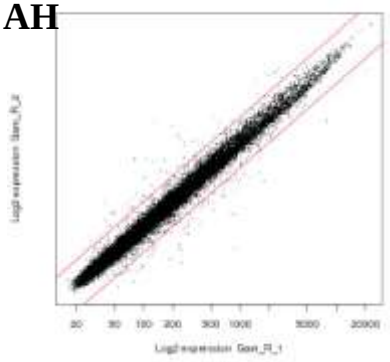
AF



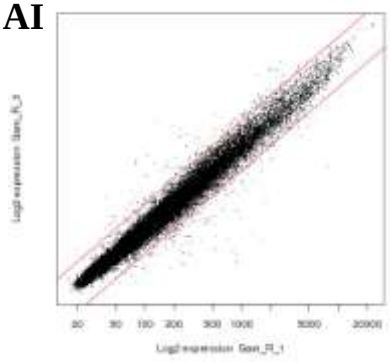
AG



AH



AI



AJ

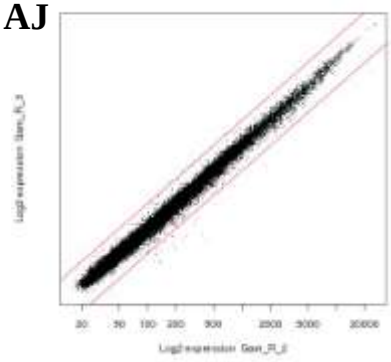
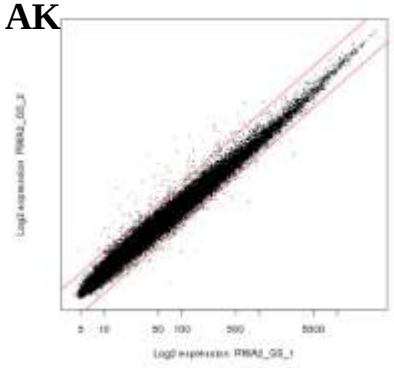
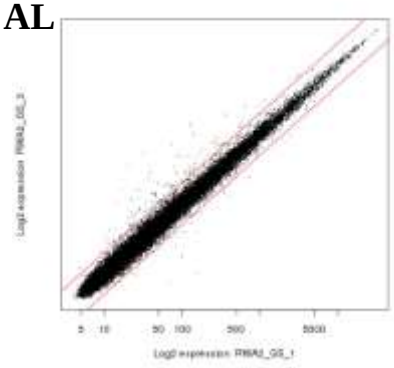


***Figure Appx 5.7* - 18 Slides cont.**

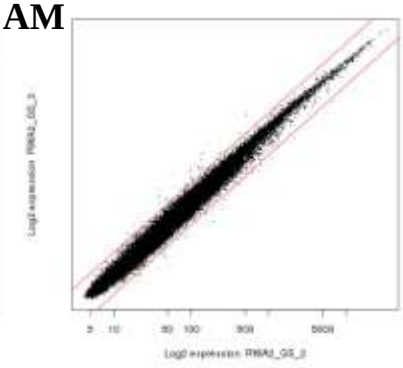
AK



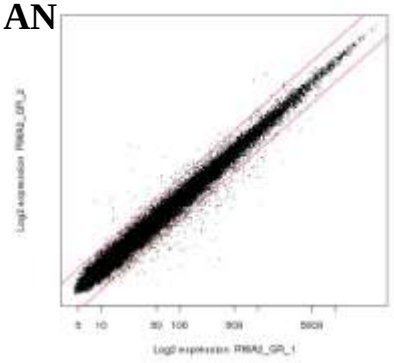
AL



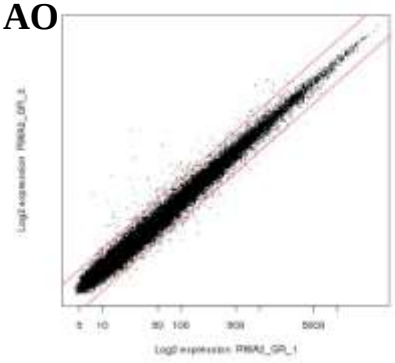
AM



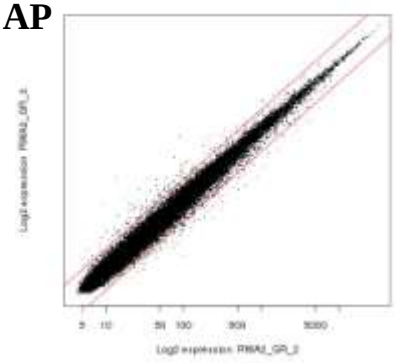
AN



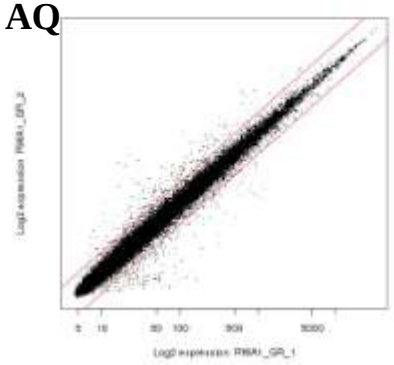
AO



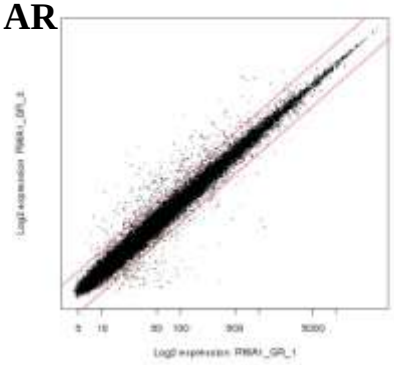
AP



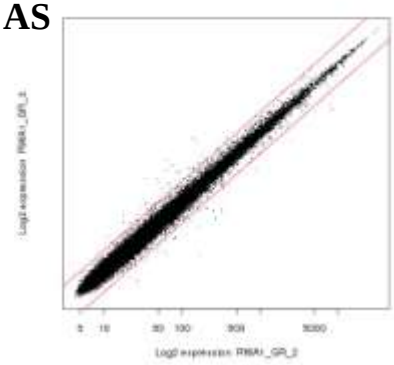
AQ



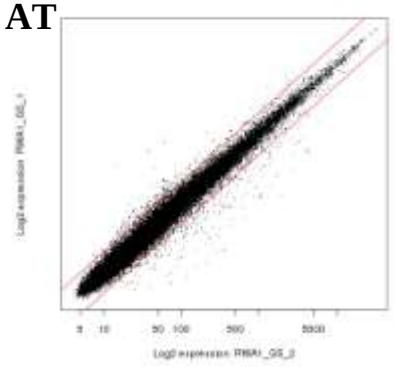
AR



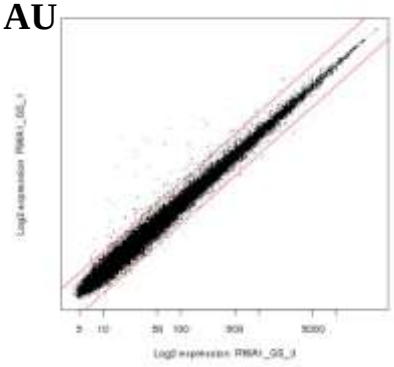
AS



AT



AU



AV

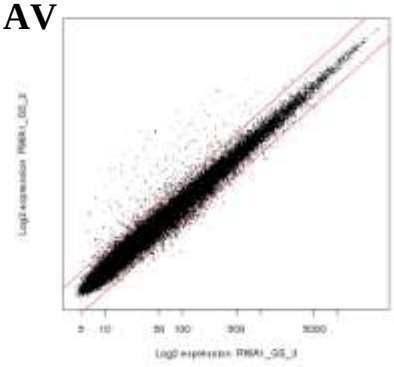
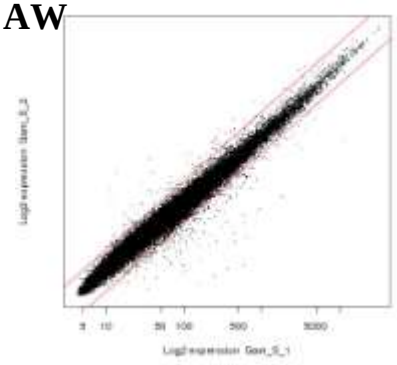
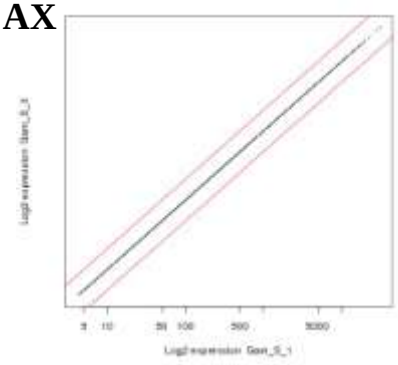


***Figure Appx 5.7* - 18 Slides cont.**

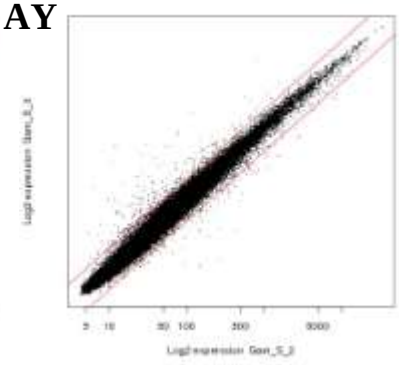
AW



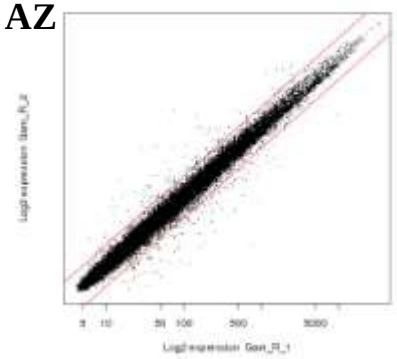
AX



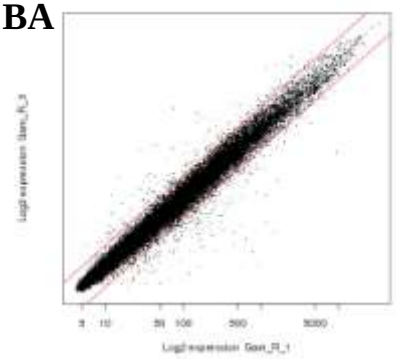
AY



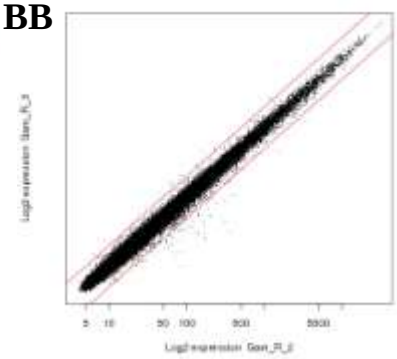
AZ



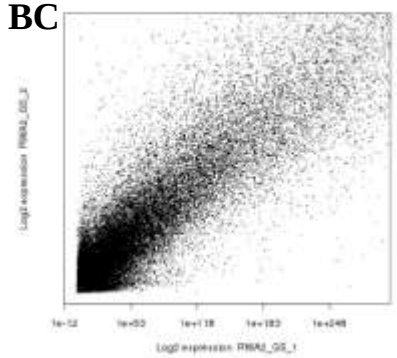
BA



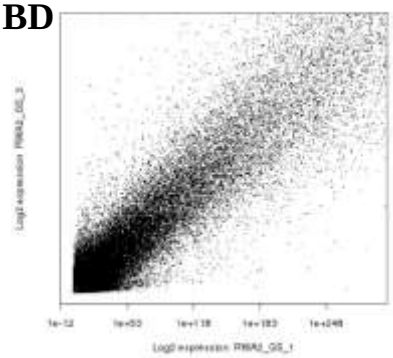
BB



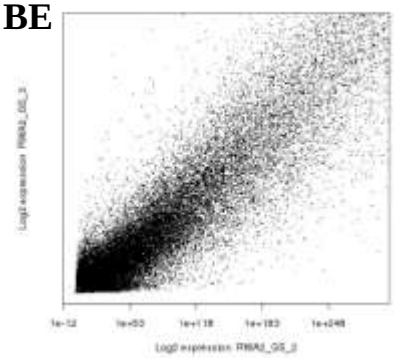
BC



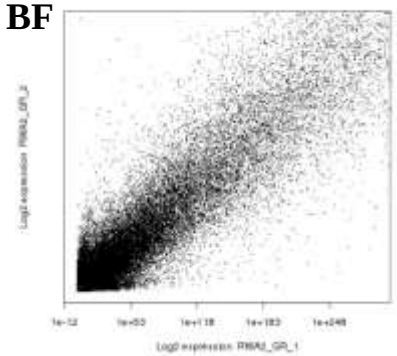
BD



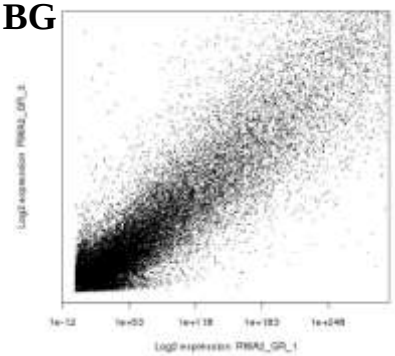
BE



BF



BG



BH

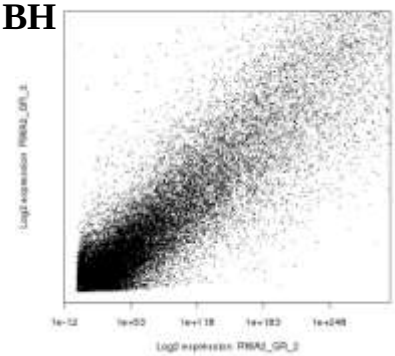
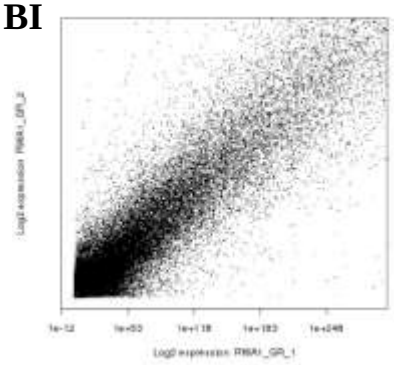
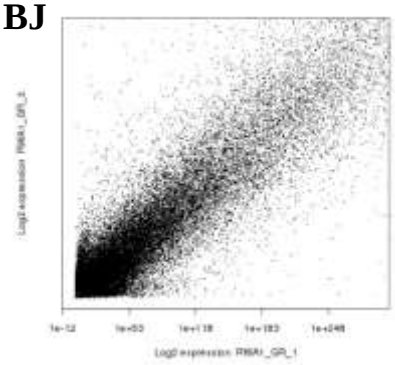


***Figure Appx 5.7* - 18 Slides cont.**

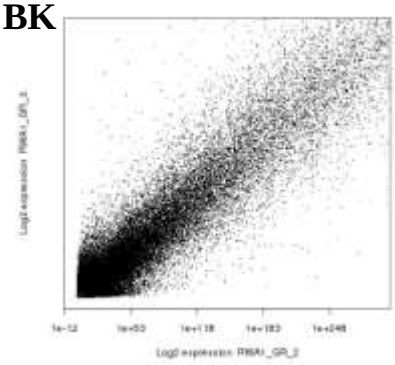
BI



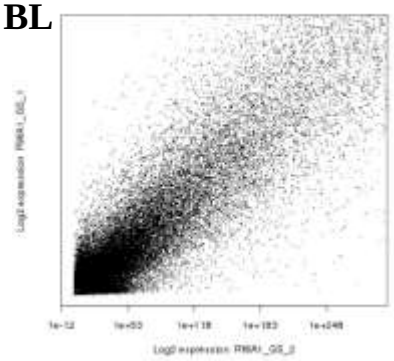
BJ



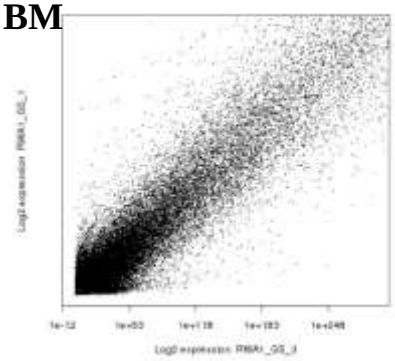
BK



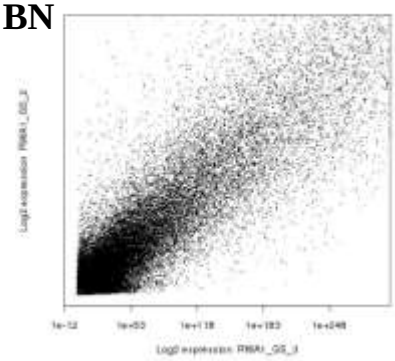
BL



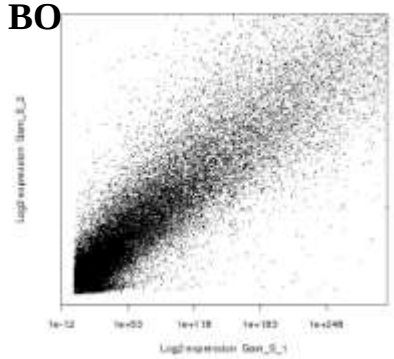
BM



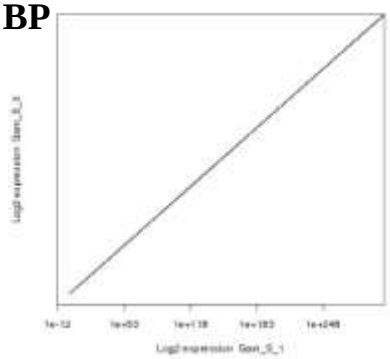
BN



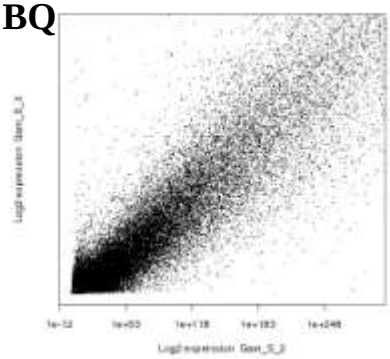
BO



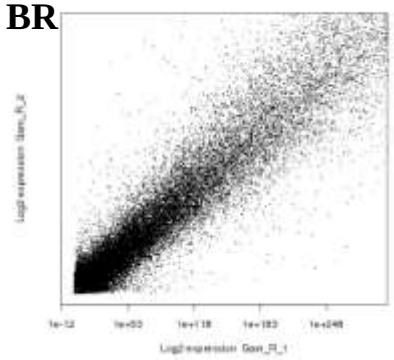
BP



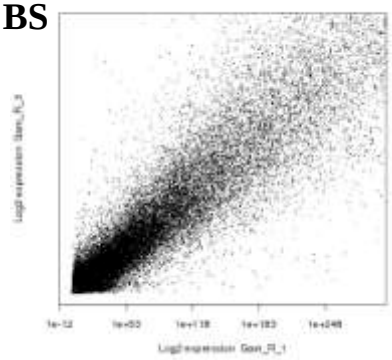
BQ



BR



BS



BT

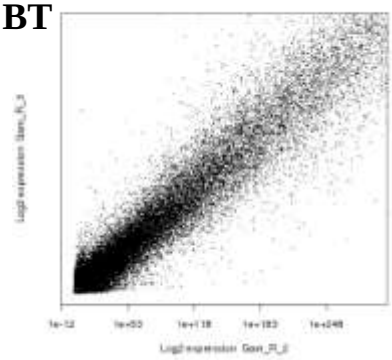
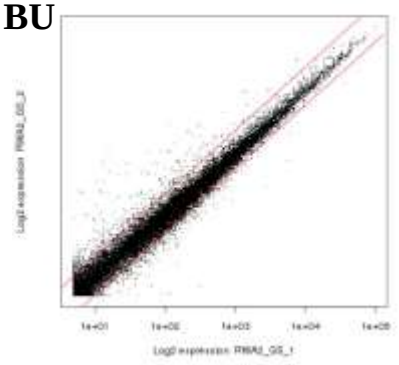
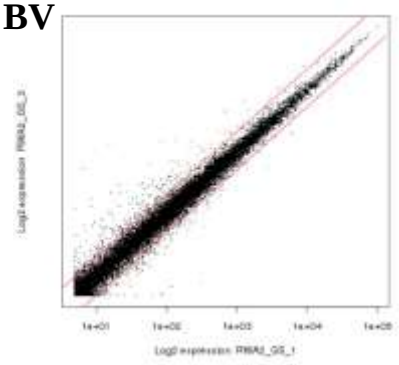


***Figure Appx 5.7* - 18 Slides cont.**

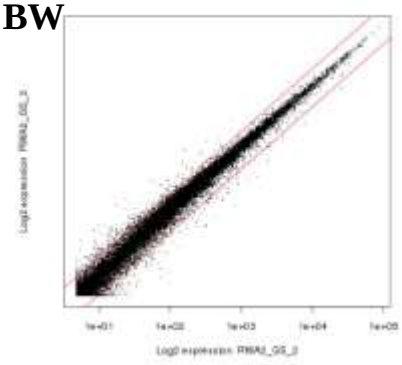
BU



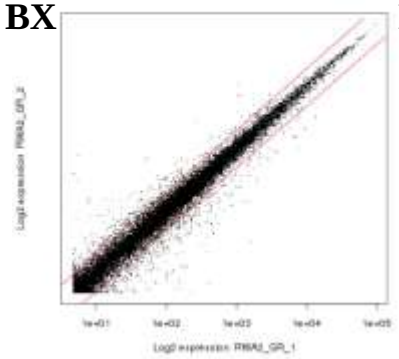
BV



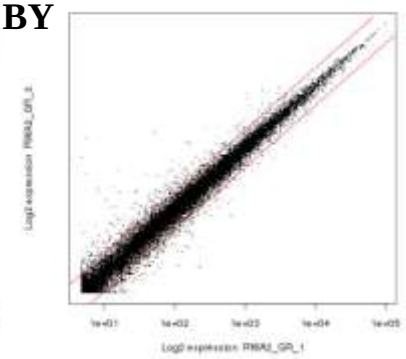
BW



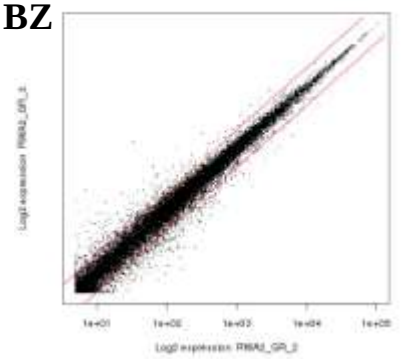
BX



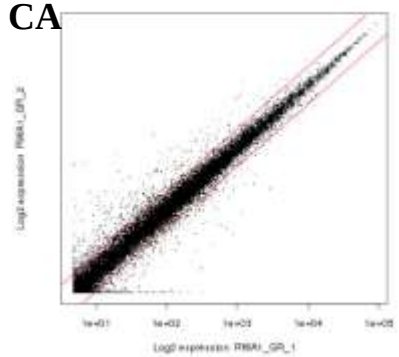
BY



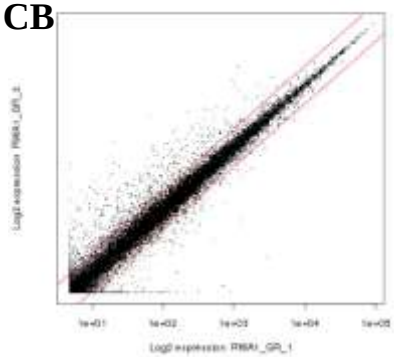
BZ



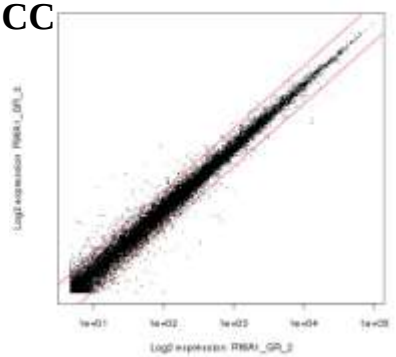
CA



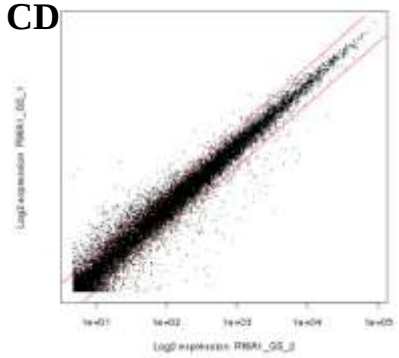
CB



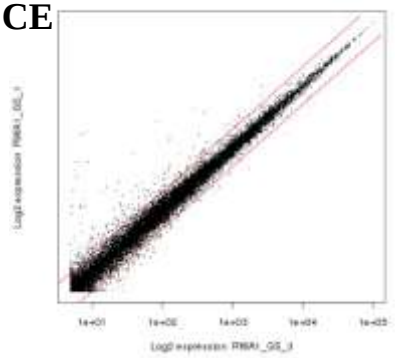
CC



CD



CE



CF

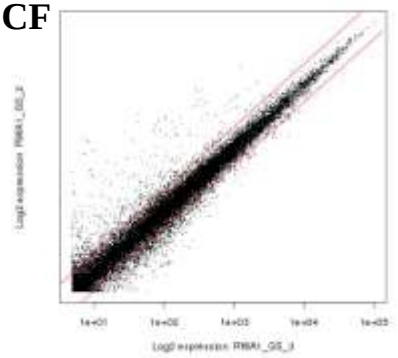


Figure Appx 5.7 **- 18 Slides cont.**

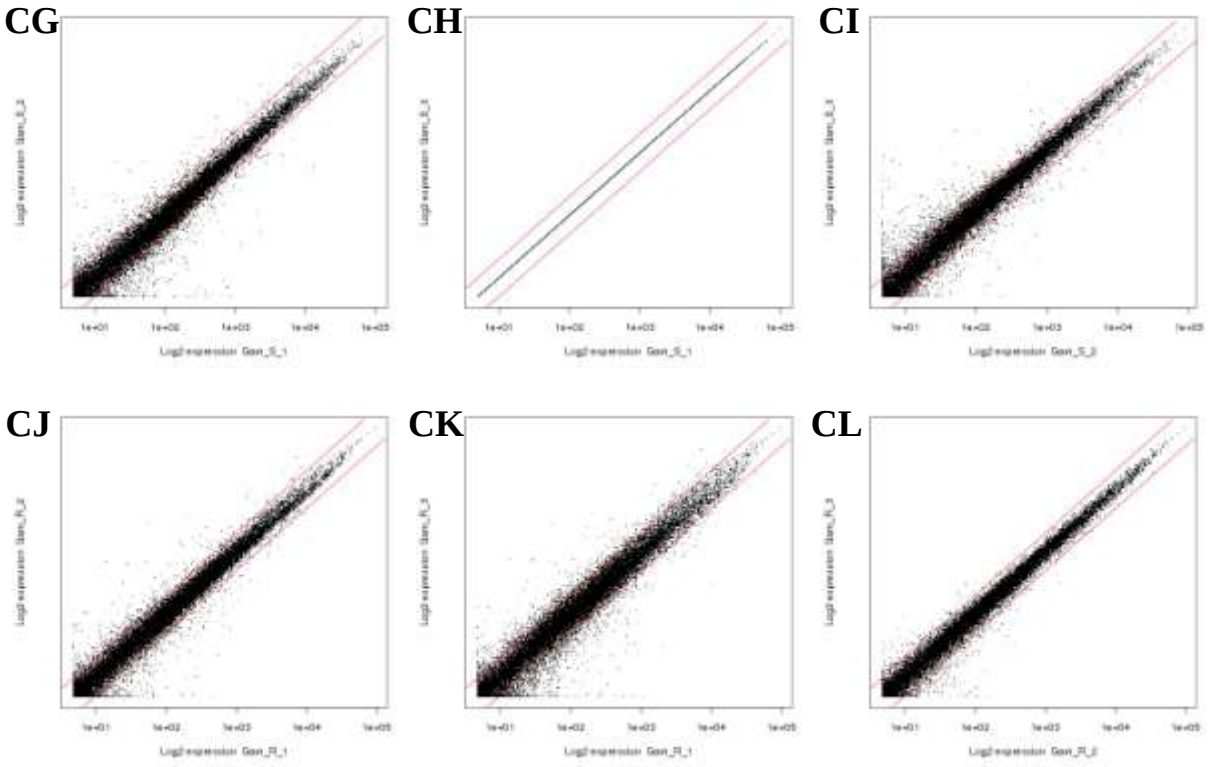


Figure Appx 5.7 The $\log_2$ expression slide comparison plots for all the normalized datasets. The raw/pre-normalized dataset is also produced and included for comparison. The files are written to the folder */R results/Graphs of raw and normalized data/Log2 plots/*. The 12 Slides files are: (A) rawdata_Tug1_vs_Tug2_log2.jpg, (B) rawdata_Tug1_vs_Tug3_log2.jpg, (C) rawdata_Tug2_vs_-Tug3_log2.jpg, (D) rawdata_TugDn_1_vs_TugDn_2_log2.jpg, (E) rawdata_TugDn_1_vs_TugDn_-3_log2.jpg, (F) rawdata_TugDn_2_vs_TugDn_3_log2.jpg, (G) rawdata_TugDn2_1_vs_TugDn2_-2_log2.jpg, (H) rawdata_TugDn2_1_vs_TugDn2_3_log2.jpg, (I) rawdata_TugDn2_2_vs_TugDn2_-3_log2.jpg, (J) rawdata_TugDn5_1_vs_TugDn5_2_log2.jpg, (K) rawdata_TugDn5_1_vs_TugDn5_-3_log2.jpg, (L) rawdata_TugDn5_2_vs_TugDn5_3_log2.jpg, (M) esetVSN_Tug1_vs_Tug2_log2.-jpg, (N) esetVSN_Tug1_vs_Tug3_log2.jpg, (O) esetVSN_Tug2_vs_Tug3_log2.jpg, (P) esetVSN_-TugDn_1_vs_TugDn_2_log2.jpg, (Q) esetVSN_TugDn_1_vs_TugDn_3_log2.jpg, (R) esetVSN_-TugDn_2_vs_TugDn_3_log2.jpg, (S) esetVSN_TugDn2_1_vs_TugDn2_2_log2.jpg, (T) esetVSN_-TugDn2_1_vs_TugDn2_3_log2.jpg, (U) esetVSN_TugDn2_2_vs_TugDn2_3_log2.jpg, (V) eset-VSN_TugDn5_1_vs_TugDn5_2_log2.jpg, (W) esetVSN_TugDn5_1_vs_TugDn5_3_log2.jpg, (X) esetVSN_TugDn5_2_vs_TugDn5_3_log2.jpg, (Y) esetRMA_Tug1_vs_Tug2_log2.jpg, (Z) eset-RMA_Tug1_vs_Tug3_log2.jpg, (AA) esetRMA_Tug2_vs_Tug3_log2.jpg, (AB) esetRMA_TugDn-_1_vs_TugDn_2_log2.jpg, (AC) esetRMA_TugDn_1_vs_TugDn_3_log2.jpg, (AD) esetRMA_-

TugDn_2_vs_TugDn_3_log2.jpg, (AE) esetRMA_TugDn2_1_vs_TugDn2_2_log2.jpg, (AF) esetRMA_TugDn2_1_vs_TugDn2_3_log2.jpg, (AG) esetRMA_TugDn2_2_vs_TugDn2_3_log2.jpg, (AH) esetRMA_TugDn5_1_vs_TugDn5_2_log2.jpg, (AI) esetRMA_TugDn5_1_vs_TugDn5_3_log2.jpg, (AJ) esetRMA_TugDn5_2_vs_TugDn5_3_log2.jpg, (AK) esetMAS_Tug1_vs_Tug2_log2-.jpg, (AL) esetMAS_Tug1_vs_Tug3_log2.jpg, (AM) esetMAS_Tug2_vs_Tug3_log2.jpg, (AN) esetMAS_TugDn_1_ vs_TugDn_2_log2.jpg, (AO) esetMAS_TugDn_1_vs_TugDn_3_log2.jpg, (AP) esetMAS_TugDn_2_vs_TugDn_3_log2.jpg, (AQ) esetMAS_TugDn2_1_vs_TugDn2_2_log2.jpg, (AR) esetMAS_TugDn2_1_vs_TugDn2_3_log2.jpg, (AS) esetMAS_TugDn2_2_vs_TugDn2_3_log2.jpg, (AT) esetMAS_TugDn5_1_vs_TugDn5_2_log2.jpg, (AU) esetMAS_TugDn5_1_vs_TugDn5_3_log2.jpg, (AV) esetMAS_TugDn5_2_vs_TugDn5_3_log2.jpg, (AW) esetGCRMA_Tug1_vs_Tug2_log2.jpg, (AX) esetGCRMA_Tug1_vs_Tug3_log2.jpg, (AY) esetGCRMA_Tug2_vs_Tug3_log2.jpg, (AZ) esetGCRMA_TugDn_1_vs_TugDn_2_log2.jpg, (BA) esetGCRMA_TugDn_1_vs_TugDn_3_log2.jpg, (BB) esetGCRMA_TugDn_2_vs_TugDn_3_log2.jpg, (BC) esetGCRMA_TugDn2_1_vs_TugDn2_2_log2.jpg, (BD) esetGCRMA_TugDn2_1_vs_TugDn2_3_log2.jpg, (BE) esetGCRMA_TugDn2_2_vs_TugDn2_3_log2.jpg, (BF) esetGCRMA_TugDn5_1_vs_TugDn5_2_log2.jpg, (BG) esetGCRMA_TugDn5_1_vs_TugDn5_3_log2.jpg, (BH) esetGCRMA_TugDn5_2_vs_TugDn5_3_log2.jpg.

The 18 Slides files are: (A) rawdata_RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (B) rawdata_RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (C) rawdata_RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (D) rawdata_RWA2_GR_1_vs_RWA2_GR_2_log2.jpg, (E) rawdata_RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (F) rawdata_RWA2_GR_2_vs_RWA2_GR_3_log2.jpg, (G) rawdata_RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (H) rawdata_RWA1_GR_1_vs_RWA1_GR_3_log2.jpg, (I) rawdata_RWA1_GR_2_vs_RWA1_GR_3_log2.jpg, (J) rawdata_RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (K) rawdata_RWA1_GS_3_vs_RWA1_GS_1_log2.jpg, (L) rawdata_RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (M) rawdata_Gam_S_1_vs_Gam_S_2_log2.jpg, (N) rawdata_Gam_S_1_vs_Gam_S_3_log2.jpg, (O) rawdata_Gam_S_2_vs_Gam_S_3_log2.jpg, (P) rawdata_Gam_R_1_vs_Gam_R_2_log2.jpg, (Q) rawdata_Gam_R_1_vs_Gam_R_3_log2.jpg, (R) rawdata_Gam_R_2_vs_Gam_R_3_log2.jpg, (S) esetVSN_RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (T) esetVSN_RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (U) esetVSN_RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (V) esetVSN_RWA2_GR_1_vs_RWA2_GR_2_log2.jpg, (W) esetVSN_RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (X) esetVSN_RWA2_GR_2_vs_RWA2_GR_3_log2.jpg, (Y) esetVSN_RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (Z) esetVSN_RWA1_GR_1_vs_RWA1_GR_3_log2.-

jpg, (AA) esetVSN_RWA1_GR_2_vs_RWA1_GR_3_log2.jpg, (AB) esetVSN_RWA1_GS_2_-vs_RWA1_GS_1_log2.jpg, (AC) esetVSN_RWA1_GS_3_vs_RWA1_GS_1_log2.jpg, (AD) esetVSN_RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (AE) esetVSN_Gam_S_1_vs_Gam_S_2_log2.jpg, (AF) esetVSN_Gam_S_1_vs_Gam_S_3_log2.jpg, (AG) esetVSN_Gam_S_2_vs_Gam_S_3_log2.-jpg, (AH) esetVSN_Gam_R_1_vs_Gam_R_2_log2.jpg, (AI) esetVSN_Gam_R_1_vs_Gam_R_3_-log2.jpg, (AJ) esetVSN_Gam_R_2_vs_Gam_R_3_log2.jpg, (AK) esetRMA_RWA2_GS_1_vs_-RWA2_GS_2_log2.jpg, (AL) esetRMA_RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (AM) esetRMA_RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (AN) esetRMA_RWA2_GR_1_vs_RWA2_GR_-2_log2.jpg, (AO) esetRMA_RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (AP) esetRMA_RWA2_-GR_2_vs_RWA2_GR_3_log2.jpg, (AQ) esetRMA_RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (AR) esetRMA_RWA1_GR_1_vs_RWA1_GR_3_log2.jpg, (AS) esetRMA_RWA1_GR_2_vs_-RWA1_GR_3_log2.jpg, (AT) esetRMA_RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (AU) esetRMA_RWA1_GS_3_vs_RWA1_GS_1_log2.jpg, (AV) esetRMA_RWA1_GS_3_vs_RWA1_GS_-2_log2.jpg, (AW) esetRMA_Gam_S_1_vs_Gam_S_2_log2.jpg, (AX) esetRMA_Gam_S_1_vs_-Gam_S_3_log2.jpg, (AY) esetRMA_Gam_S_2_vs_Gam_S_3_log2.jpg, (AZ) esetRMA_Gam_R_-1_vs_Gam_R_2_log2.jpg, (BA) esetRMA_Gam_R_1_vs_Gam_R_3_log2.jpg, (BB) esetRMA_-Gam_R_2_vs_Gam_R_3_log2.jpg, (BC) esetMAS_RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (BD) esetMAS_RWA2_GS_1_vs_RWA2_GS_3_log2.jpg, (BE) esetMAS_RWA2_GS_2_vs_RWA2_-GS_3_log2.jpg, (BF) esetMAS_RWA2_GR_1_vs_RWA2_GR_2_log2.jpg, (BG) esetMAS_-RWA2_GR_1_vs_RWA2_GR_3_log2.jpg, (BH) esetMAS_RWA2_GR_2_vs_RWA2_GR_3_log2.-jpg, (BI) esetMAS_RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (BJ) esetMAS_RWA1_GR_1_vs_-RWA1_GR_3_log2.jpg, (BK) esetMAS_RWA1_GR_2_vs_RWA1_GR_3_log2.jpg, (BL) esetMAS_RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (BM) esetMAS_RWA1_GS_3_vs_RWA1_GS_-1_log2.jpg, (BN) esetMAS_RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (BO) esetMAS_Gam_S_1_-vs_Gam_S_2_log2.jpg, (BP) esetMAS_Gam_S_1_vs_Gam_S_3_log2.jpg, (BQ) esetMAS_Gam_-S_2_vs_Gam_S_3_log2.jpg, (BR) esetMAS_Gam_R_1_vs_Gam_R_2_log2.jpg, (BS) esetMAS_-Gam_R_1_vs_Gam_R_3_log2.jpg, (BT) esetMAS_Gam_R_2_vs_Gam_R_3_log2.jpg, (BU) esetGCRMA_RWA2_GS_1_vs_RWA2_GS_2_log2.jpg, (BV) esetGCRMA_RWA2_GS_1_vs_-RWA2_GS_3_log2.jpg, (BW) esetGCRMA_RWA2_GS_2_vs_RWA2_GS_3_log2.jpg, (BX) esetGCRMA_RWA2_GR_1_vs_RWA2_GR_2_log2.jpg, (BY) esetGCRMA_RWA2_GR_1_vs_-RWA2_GR_3_log2.jpg, (BZ) esetGCRMA_RWA2_GR_2_vs_RWA2_GR_3_log2.jpg, (CA) esetGCRMA_RWA1_GR_1_vs_RWA1_GR_2_log2.jpg, (CB) esetGCRMA_RWA1_GR_1_vs_-

RWA1_GR_3_log2.jpg, (CC) esetGCRMA_RWA1_GR_2_vs_RWA1_GR_3_log2.jpg, (CD) esetGCRMA_RWA1_GS_2_vs_RWA1_GS_1_log2.jpg, (CE) esetGCRMA_RWA1_GS_3_vs_-RWA1_GS_1_log2.jpg, (CF) esetGCRMA_RWA1_GS_3_vs_RWA1_GS_2_log2.jpg, (CG) eset-GCRMA_Gam_S_1_vs_Gam_S_2_log2.jpg, (CH) esetGCRMA_Gam_S_1_vs_Gam_S_3_log2.jpg, (CI) esetGCRMA_Gam_S_2_vs_Gam_S_3_log2.jpg, (CJ) esetGCRMA_Gam_R_1_vs_Gam_R_-2_log2.jpg, (CK) esetGCRMA_Gam_R_1_vs_Gam_R_3_log2.jpg, (CL) esetGCRMA_Gam_R_-2_vs_Gam_R_3_log2.jpg.

Figure Appx 5.8 **- 12 Slides**

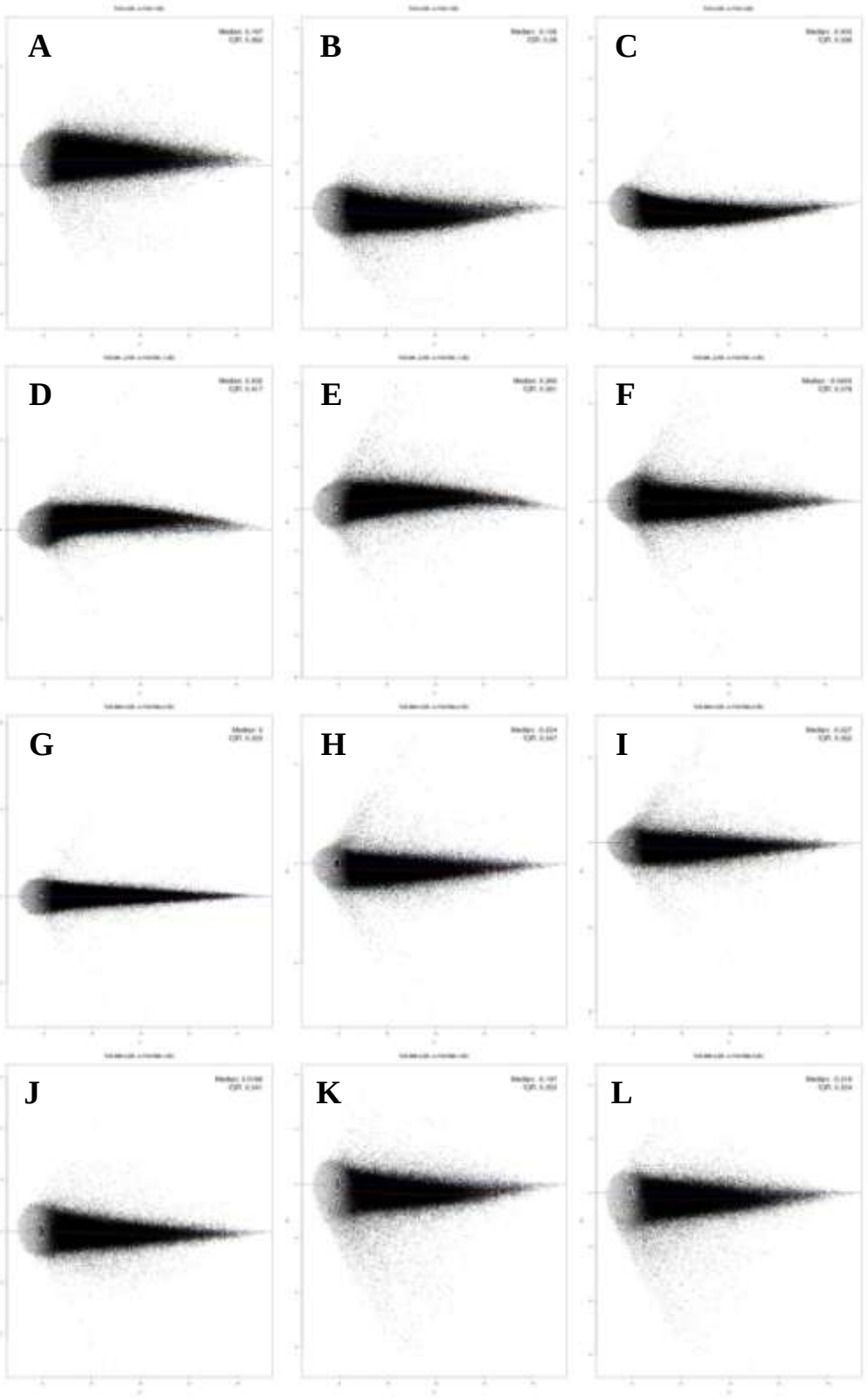


***Figure Appx 5.8* - 12 Slides cont.**

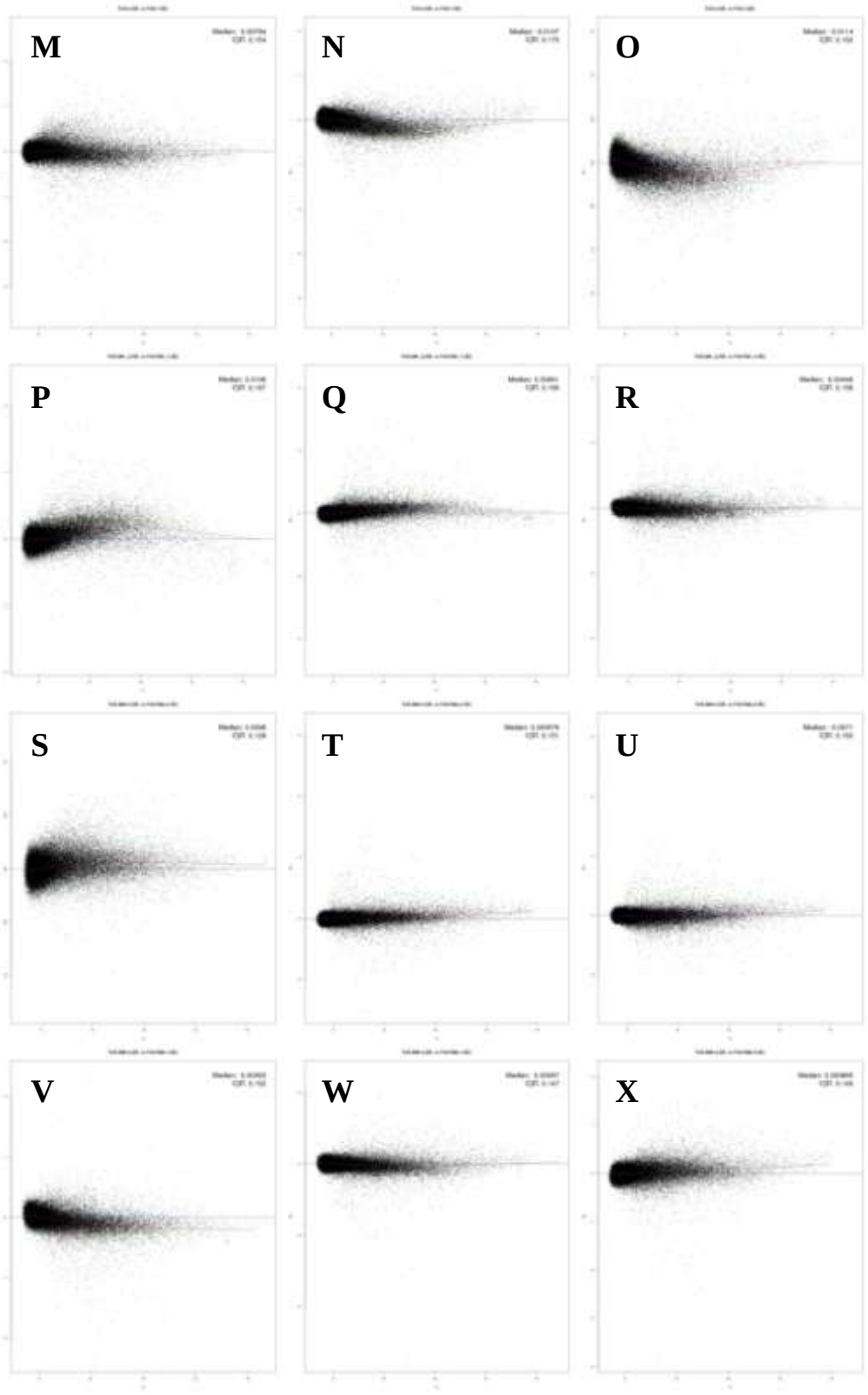


***Figure Appx 5.8* - 12 Slides cont.**

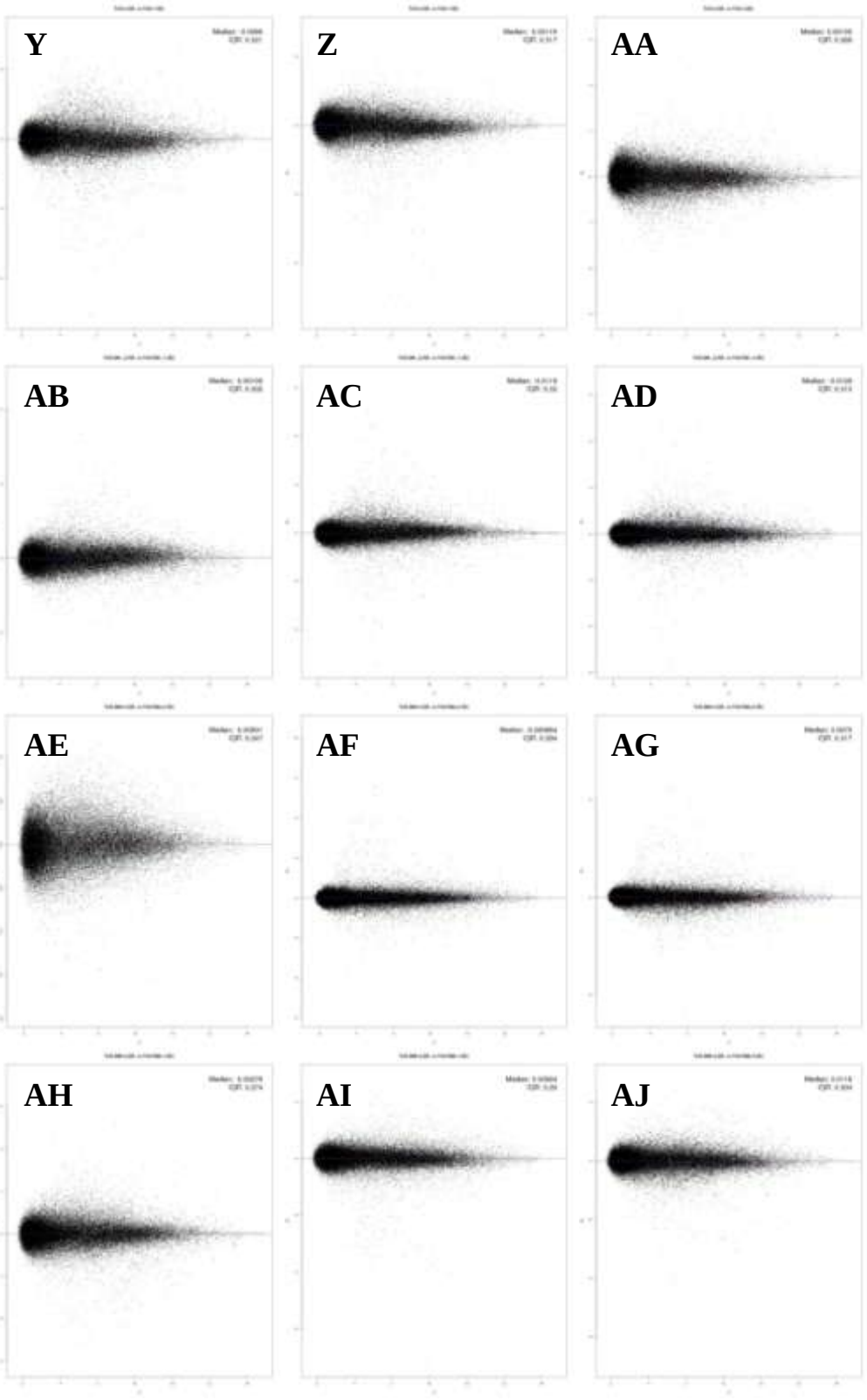


***Figure Appx 5.8* - 12 Slides cont.**

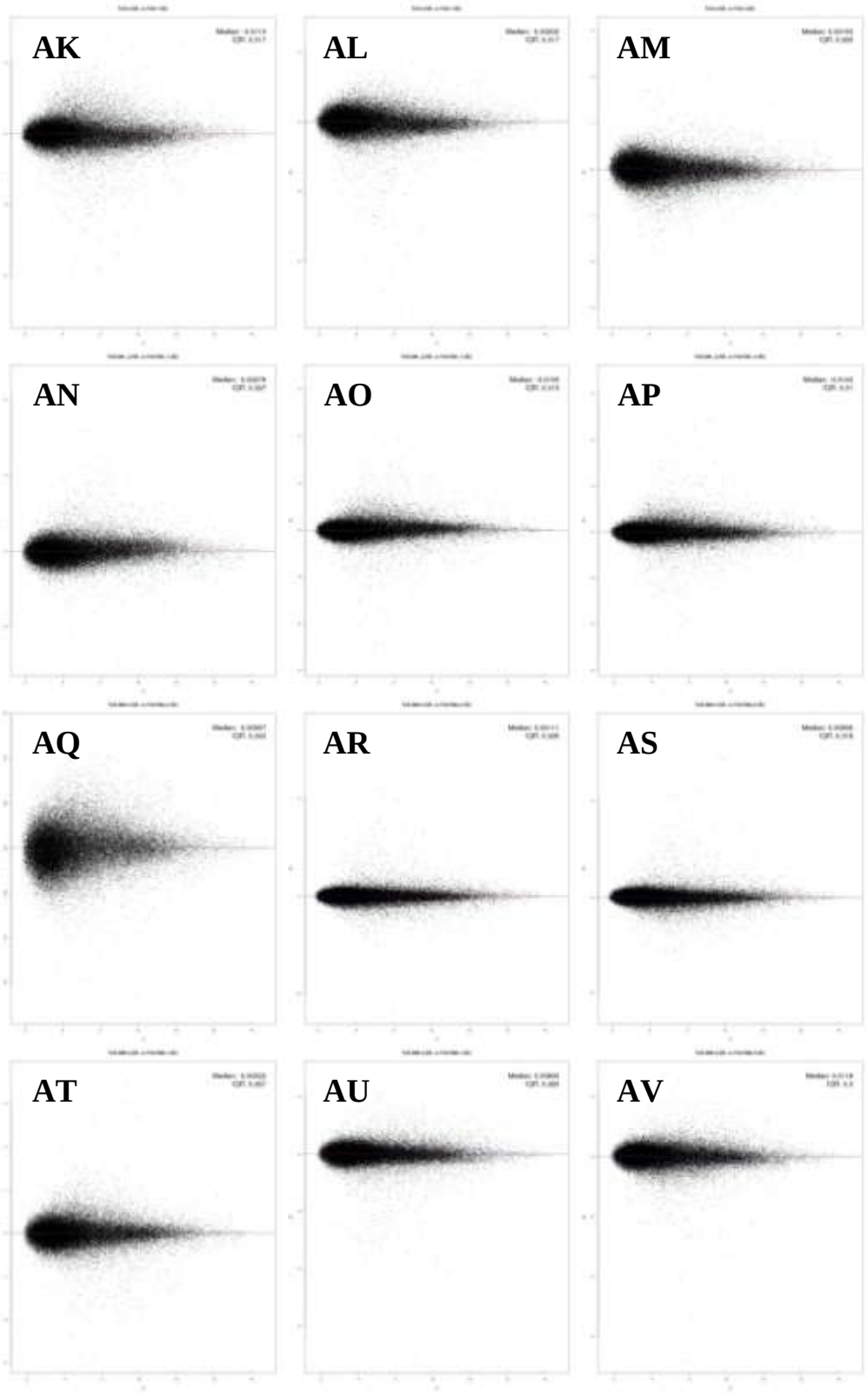


***Figure Appx 5.8* - 12 Slides cont.**

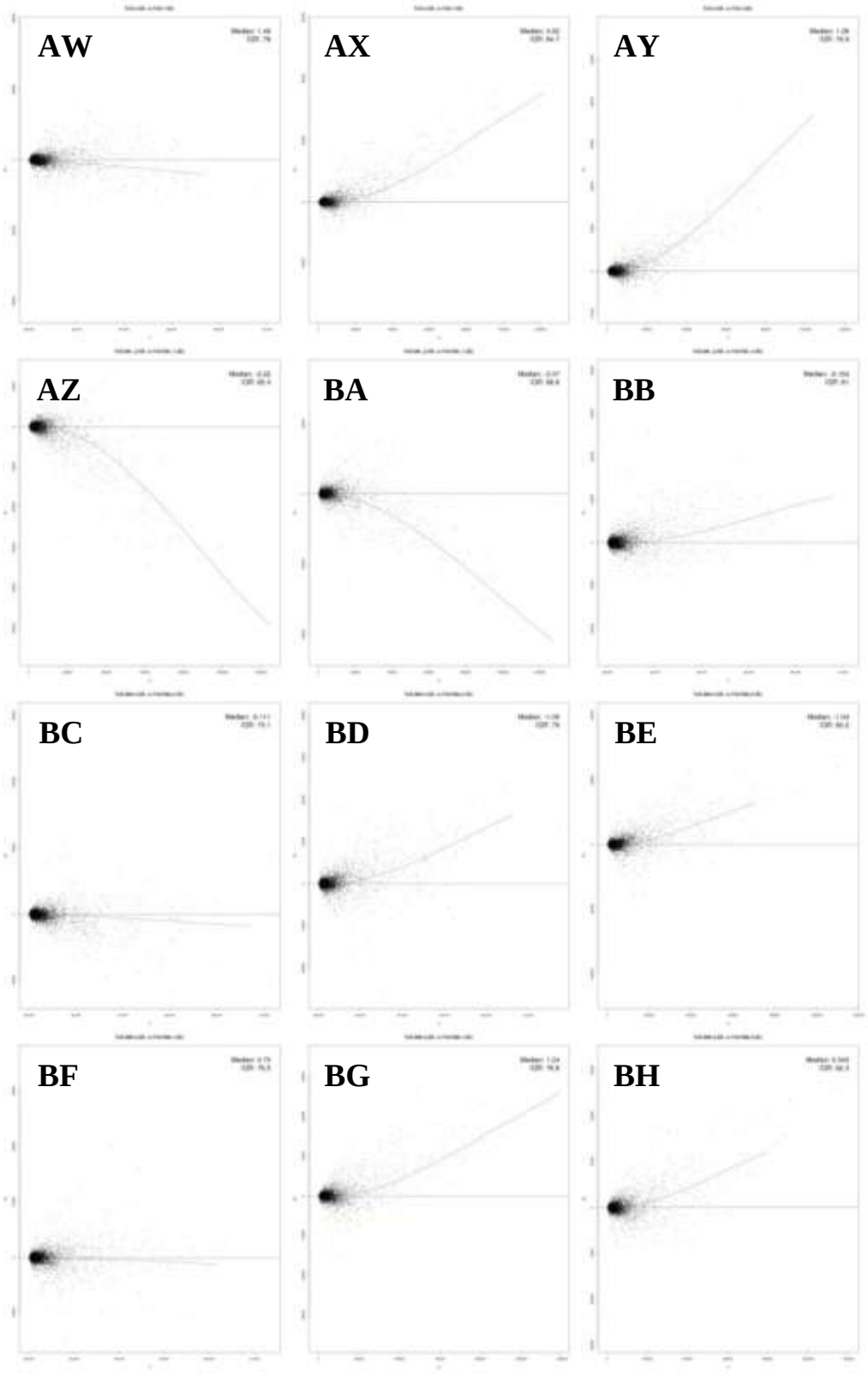


***Figure Appx 5.8* - 12 Slides cont.**

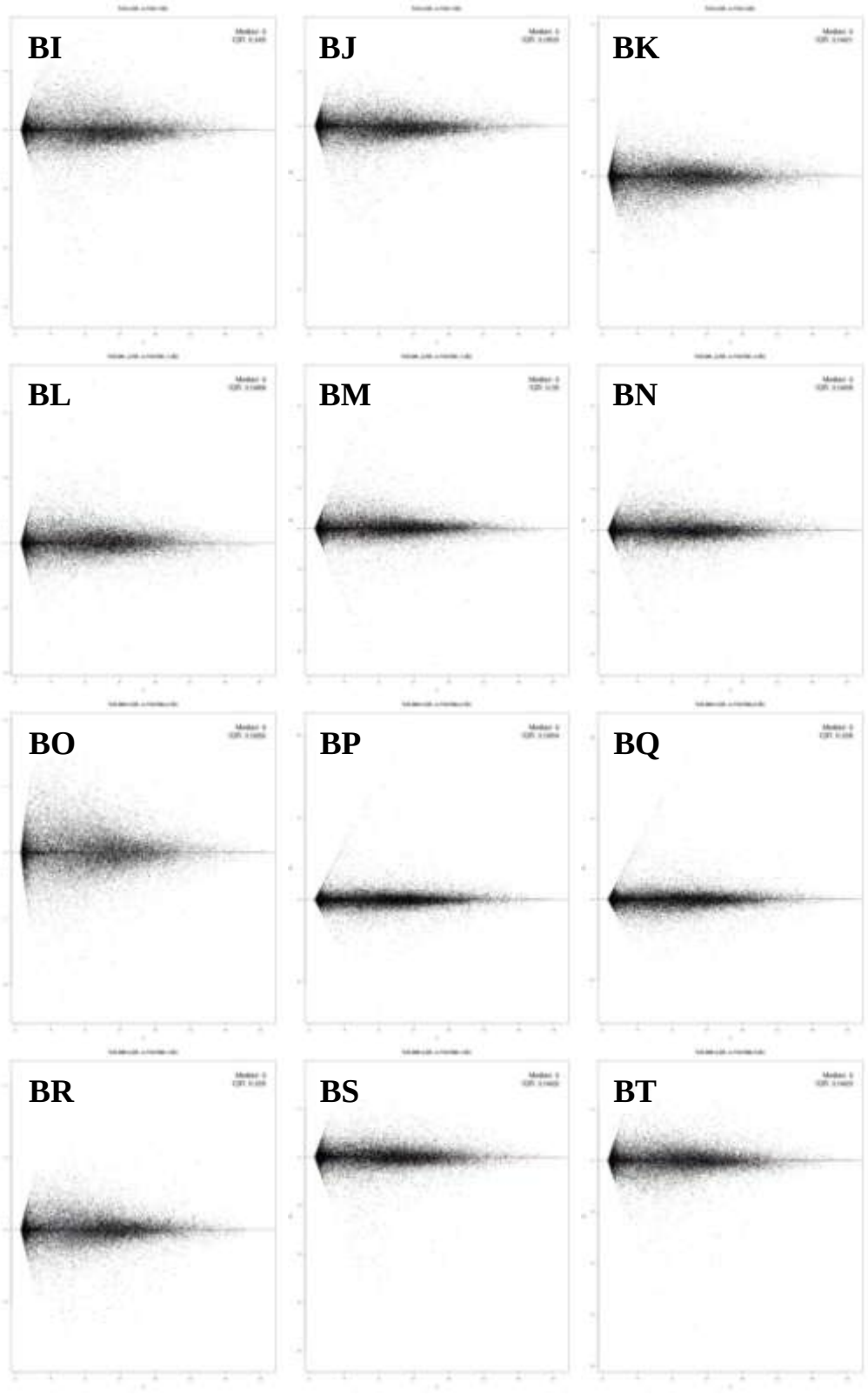


Figure Appx 5.8 **- 18 Slides**

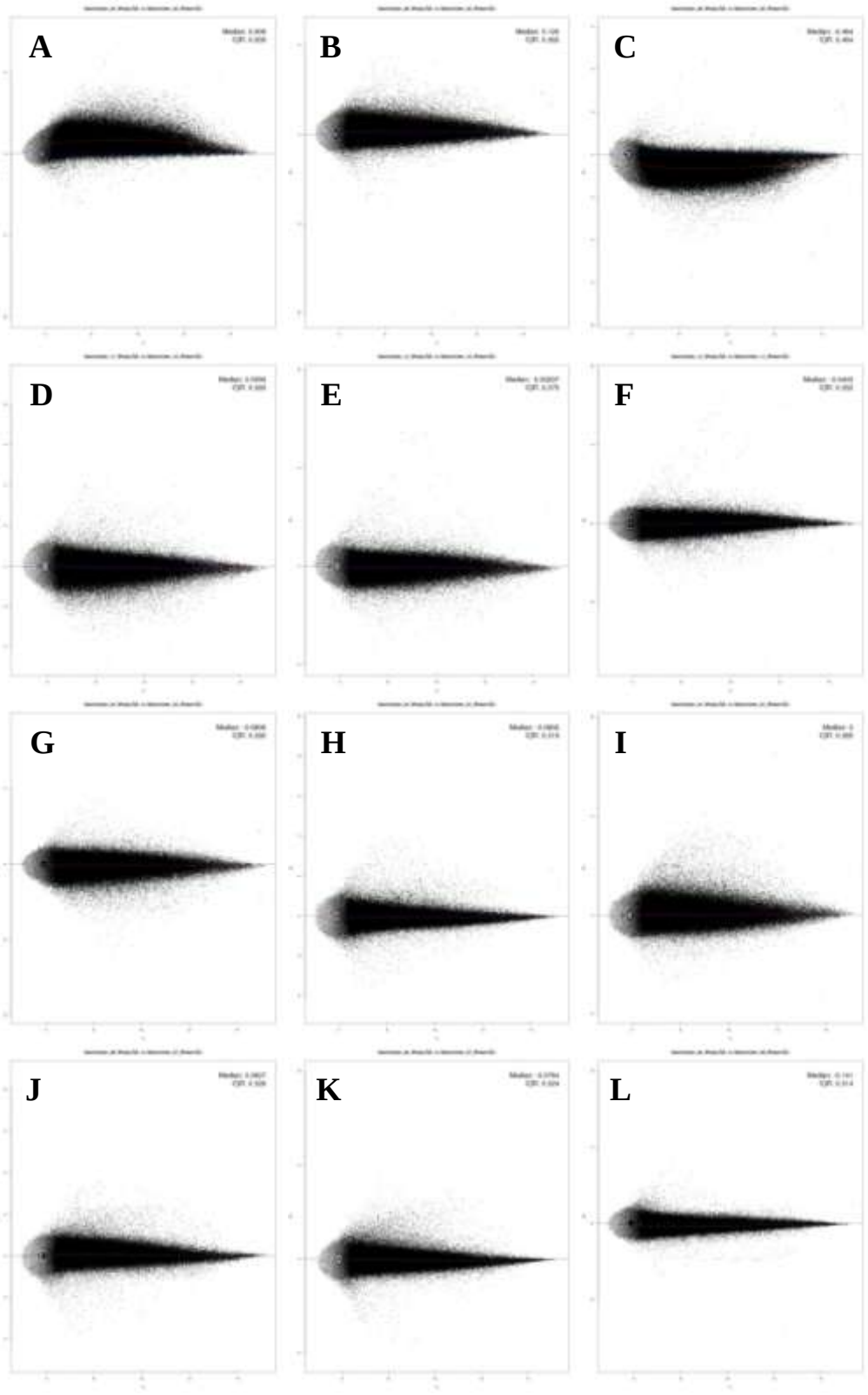


***Figure Appx 5.8* - 18 Slides cont.**

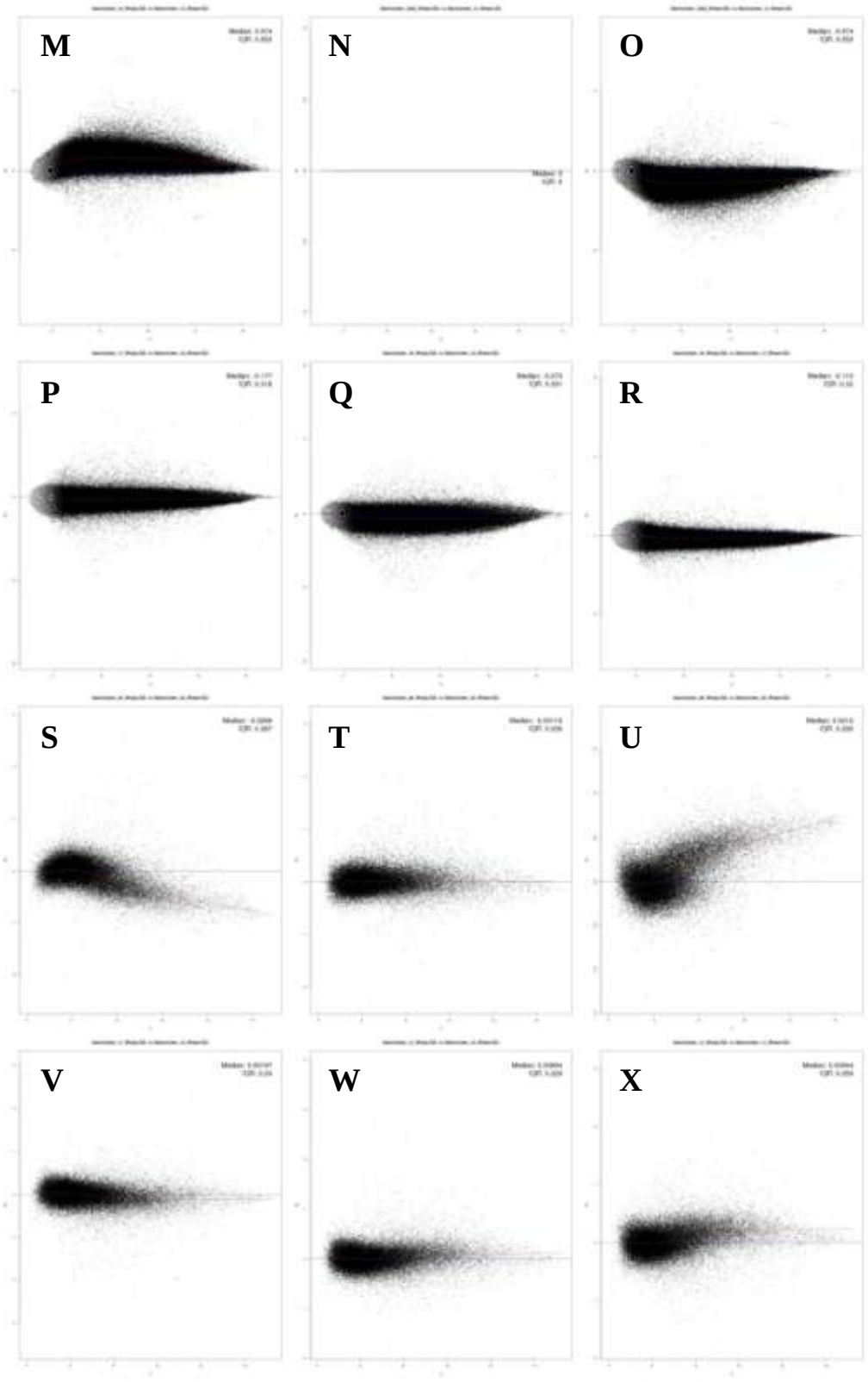


***Figure Appx 5.8* - 18 Slides cont.**

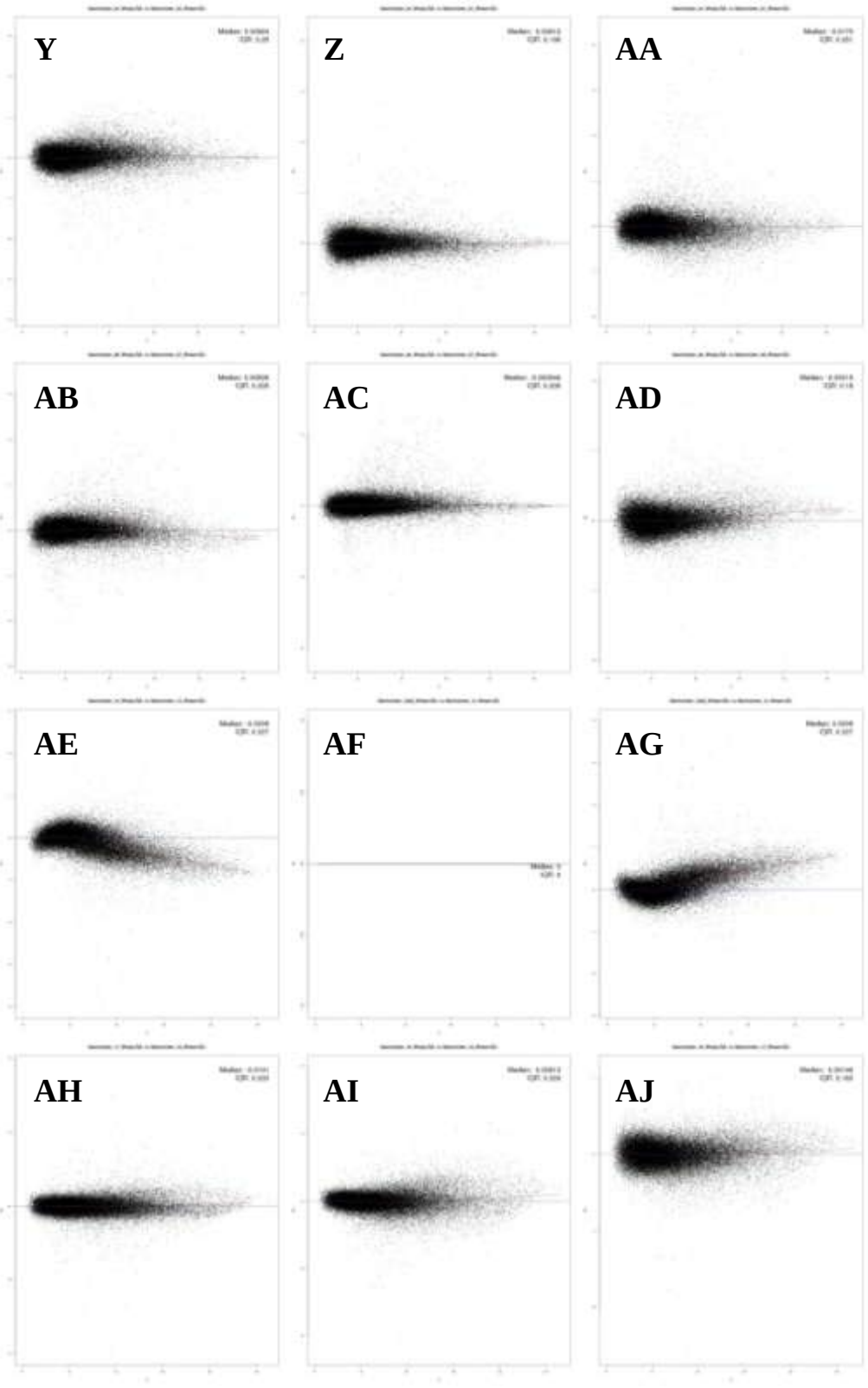


***Figure Appx 5.8* - 18 Slides cont.**

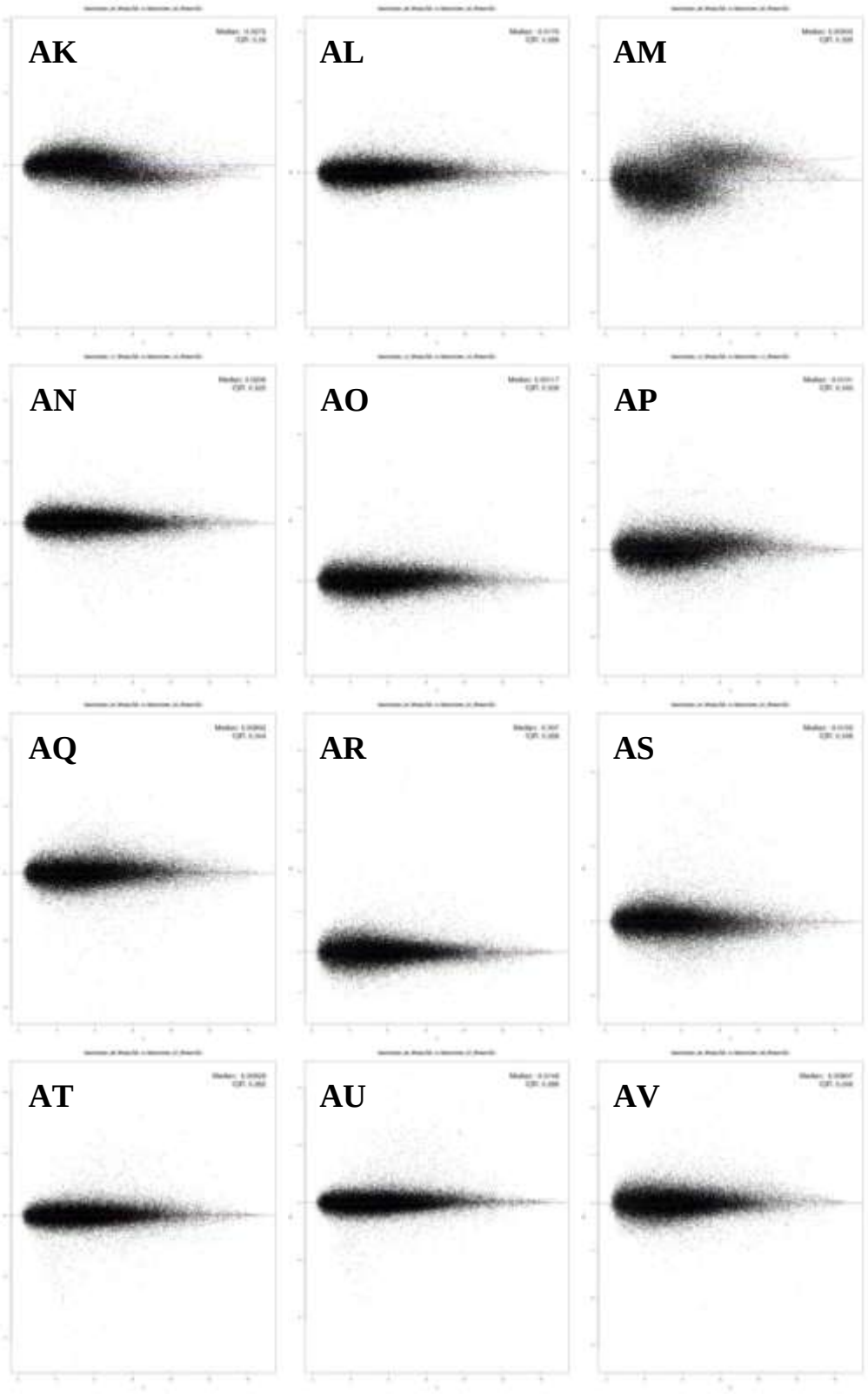


***Figure Appx 5.8* - 18 Slides cont.**

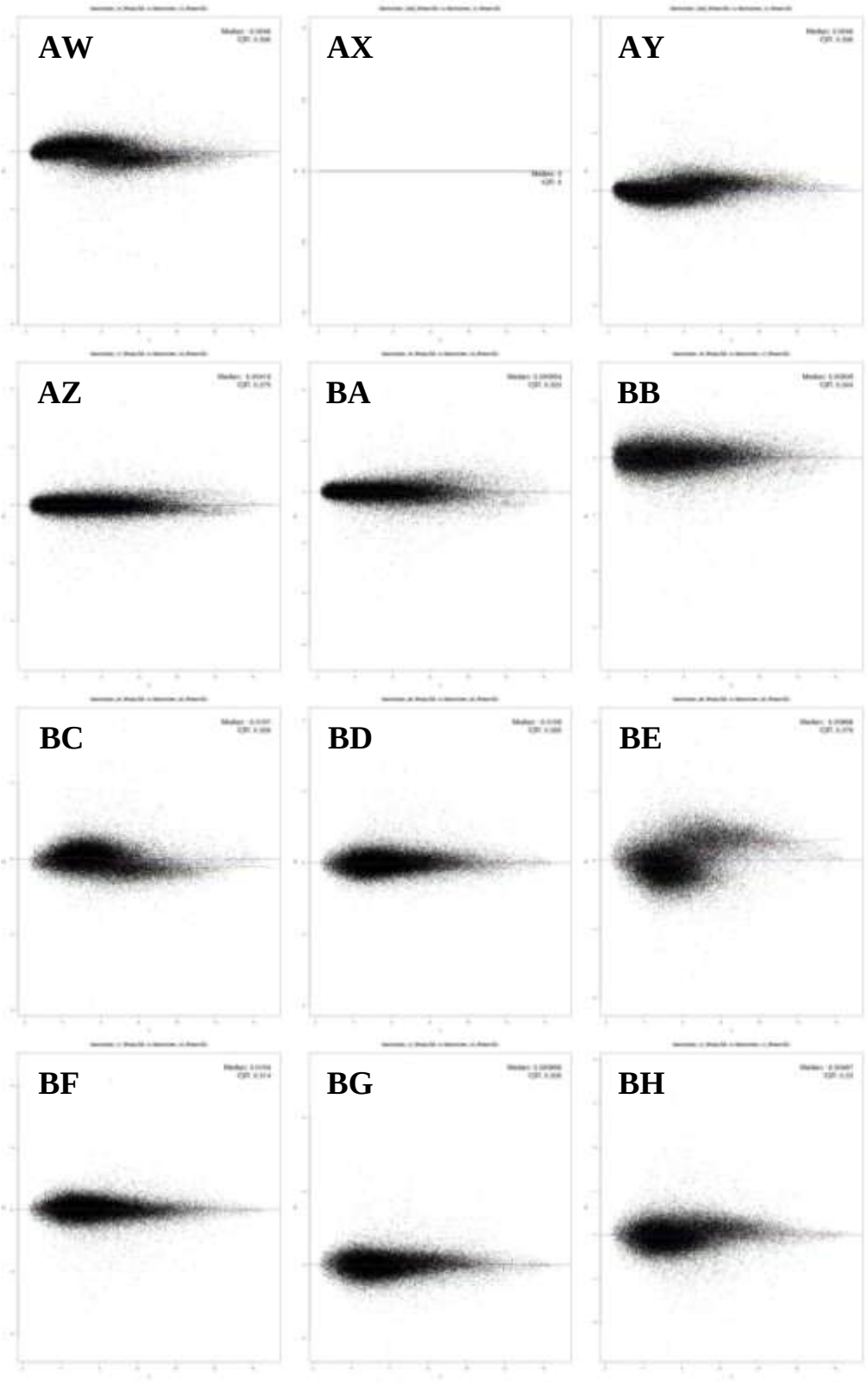


***Figure Appx 5.8* - 18 Slides cont.**

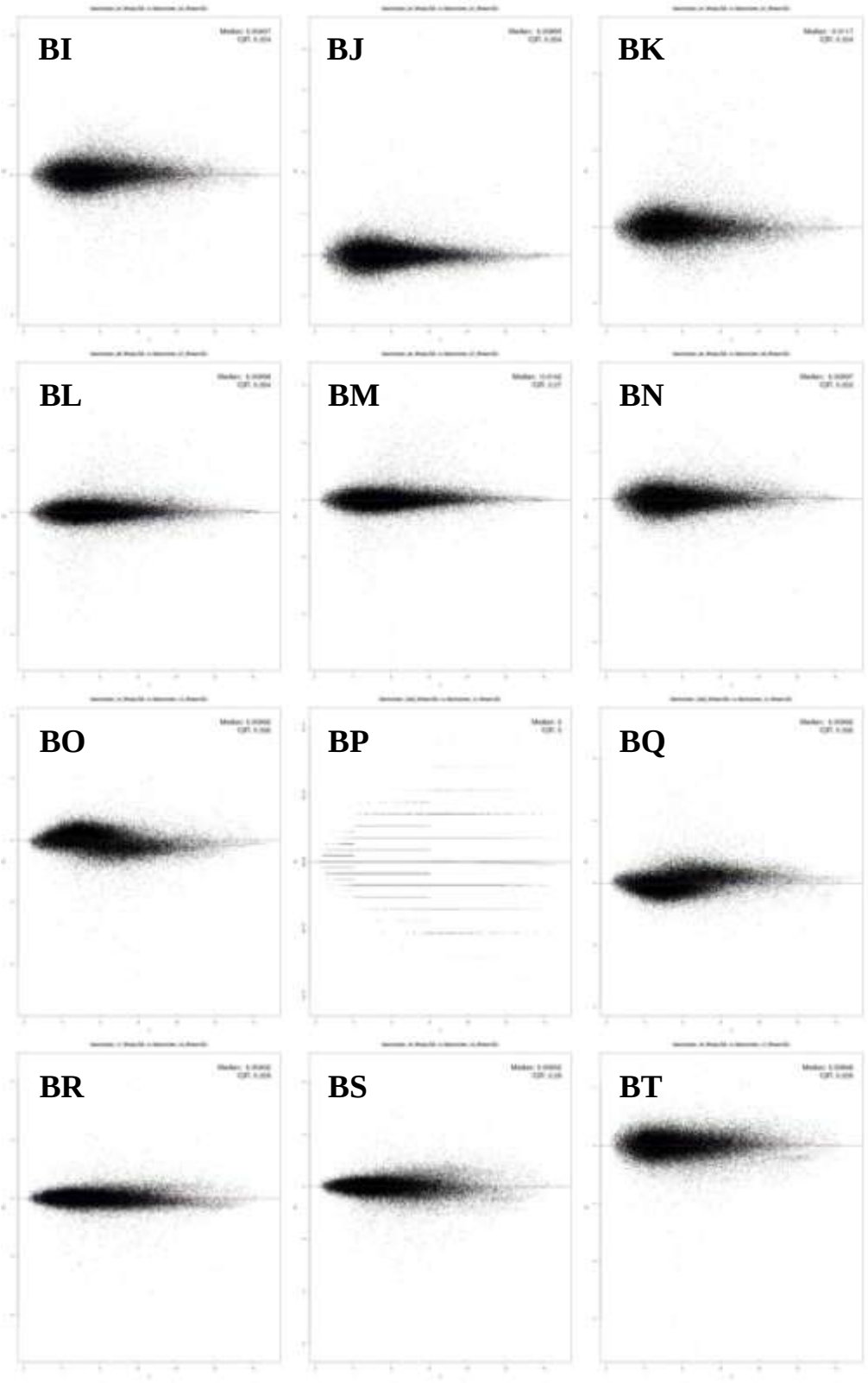


***Figure Appx 5.8* - 18 Slides cont.**

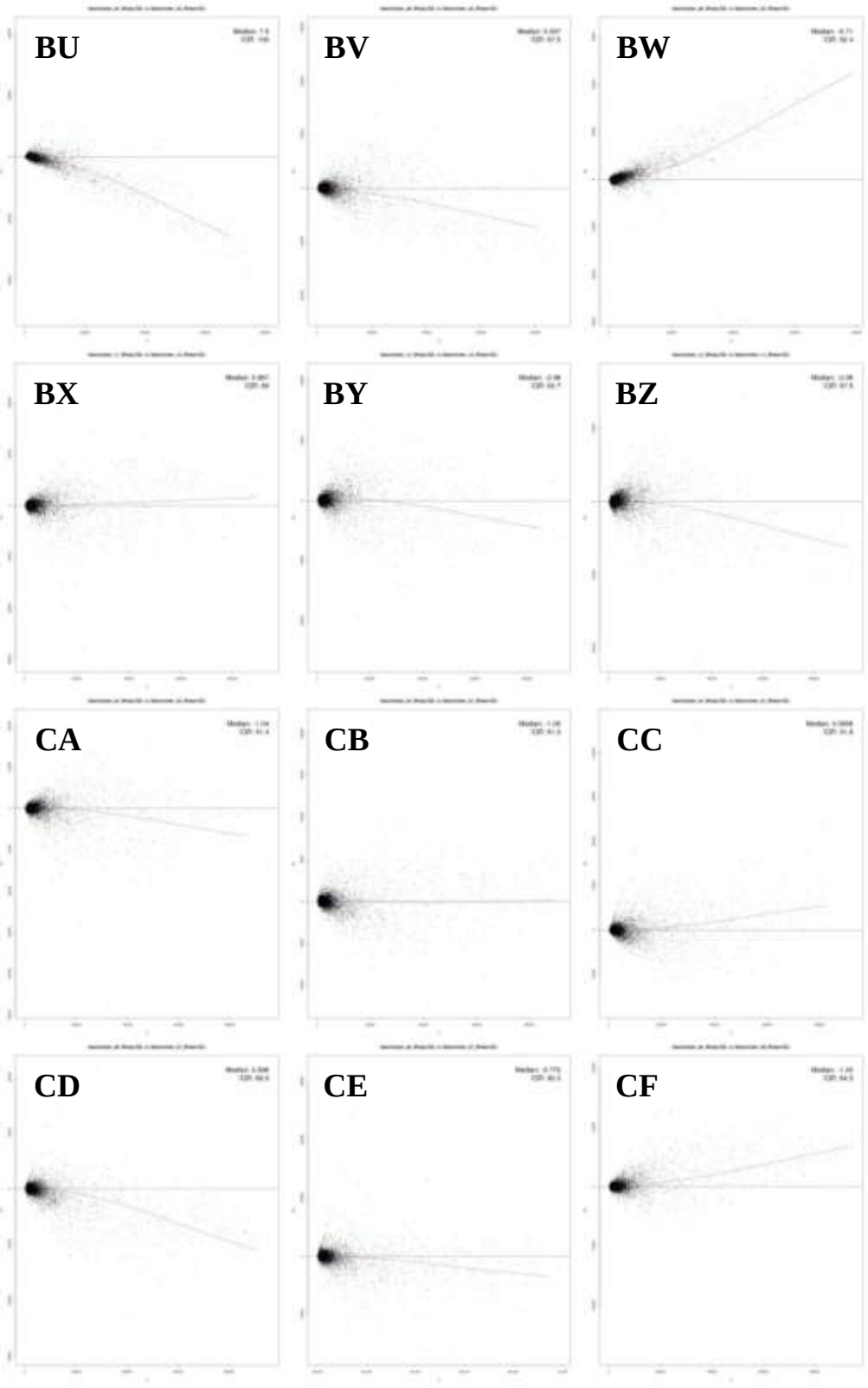


***Figure Appx 5.8* - 18 Slides cont.**

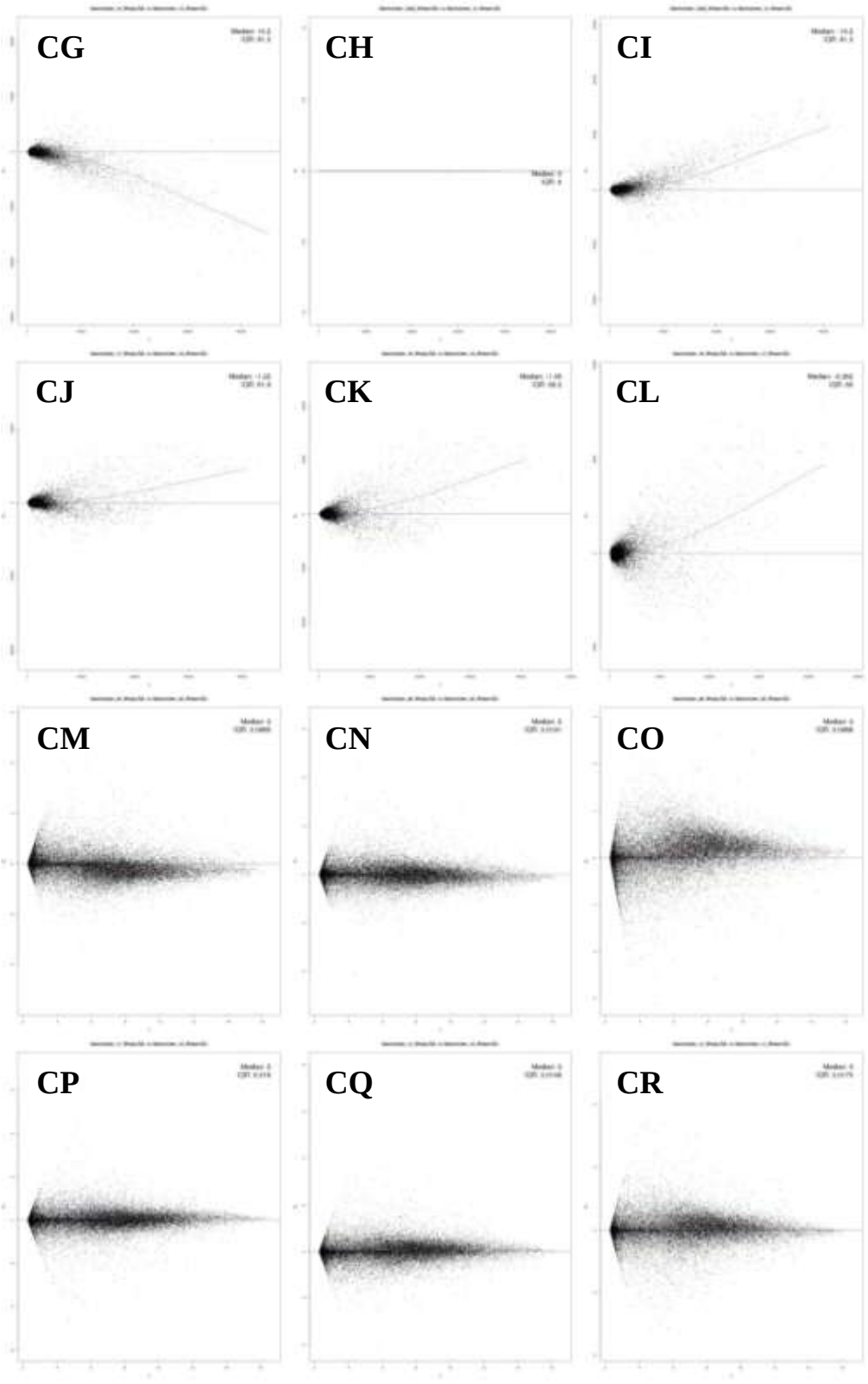


***Figure Appx 5.8* - 18 Slides cont.**

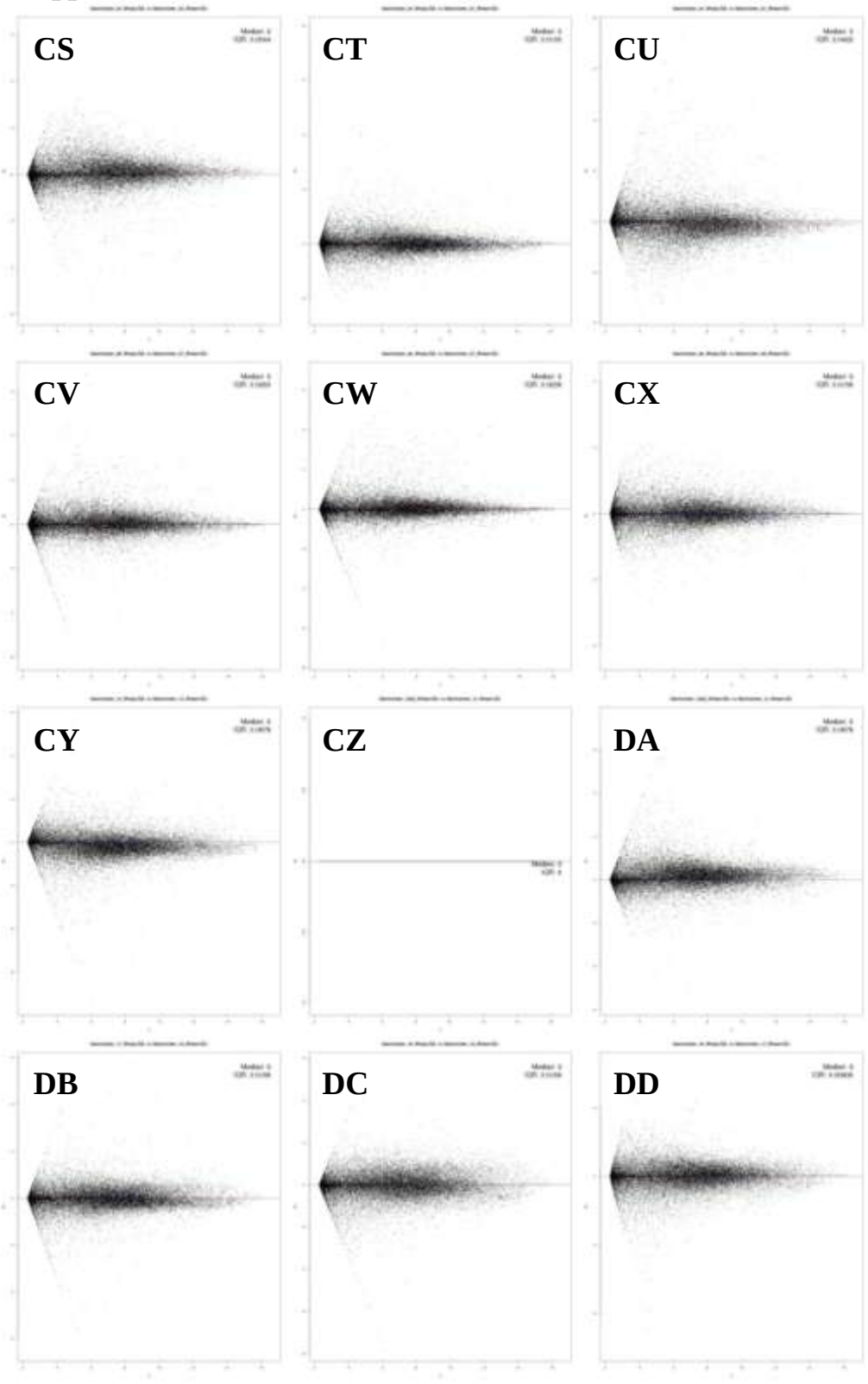


Figure Appx 5.8 The MA-plot comparisons of the different slides after normalization by the different methods. The raw/pre-normalized plots are also included for comparison. The files are

written to the subfolder */R results/Graphs of raw and normalized data/MA plots/*. The 12 slides files are: (A) rawdata_Tug1_vs_Tug2.jpg, (B) rawdata_Tug1_vs_Tug3.jpg, (C) rawdata_Tug2_vs_Tug3-.jpg, (D) rawdata_TugDn_1_vs_TugDn_2.jpg, (E) rawdata_TugDn_1_vs_TugDn_3.jpg, (F) raw-data_TugDn_2_vs_TugDn_3.jpg, (G) rawdata_TugDn2_1_vs_TugDn2_2.jpg, (H) rawdata_-TugDn2_1_vs_TugDn2_3.jpg, (I) rawdata_TugDn2_2_vs_TugDn2_3.jpg, (J) rawdata_TugDn5_1_-vs_TugDn5_2.jpg, (K) rawdata_TugDn5_1_vs_TugDn5_3.jpg, (L) rawdata_TugDn5_2_vs_-TugDn5_3.jpg, (M) esetVSN_Tug1_vs_Tug2.jpg, (N) esetVSN_Tug1_vs_Tug3.jpg, (O) eset-VSN_Tug2_vs_Tug3.jpg, (P) esetVSN_TugDn_1_vs_TugDn_2.jpg, (Q) esetVSN_TugDn_1_vs_-TugDn_3.jpg, (R) esetVSN_TugDn_2_vs_TugDn_3.jpg, (S) esetVSN_TugDn2_1_vs_Tug-Dn2_2.jpg, (T) esetVSN_TugDn2_1_vs_TugDn2_3.jpg, (U) esetVSN_TugDn2_2_vs_TugDn2-_3.jpg, (V) esetVSN_TugDn5_1_vs_TugDn5_2.jpg, (W) esetVSN_TugDn5_1_vs_TugDn5_3.jpg, (X) esetVSN_TugDn5_2_vs_TugDn5_3.jpg, (Y) esetRMA_Tug1_vs_Tug2.jpg, (Z) esetRMA_-Tug1_vs_Tug3.jpg, (AA) esetRMA_Tug2_vs_Tug3.jpg, (AB) esetRMA_TugDn_1_vs_Tug-Dn_2.jpg, (AC) esetRMA_TugDn_1_vs_TugDn_3.jpg, (AD) esetRMA_TugDn_2_vs_TugDn_-3.jpg, (AE) esetRMA_TugDn2_1_vs_TugDn2_2.jpg, (AF) esetRMA_TugDn2_1_vs_TugDn2_3-.jpg, (AG) esetRMA_TugDn2_2_vs_TugDn2_3.jpg, (AH) esetRMA_TugDn5_1_vs_TugDn5_2.jpg, (AI) esetRMA_TugDn5_1_vs_TugDn5_3.jpg, (AJ) esetRMA_TugDn5_2_vs_TugDn5_3.jpg, (AK) esetPLM_Tug1_vs_Tug2.jpg, (AL) esetPLM_Tug1_vs_Tug3.jpg, (AM) esetPLM_Tug2_vs_-Tug3.jpg, (AN) esetPLM_TugDn_1_ vs_TugDn_2.jpg, (AO) esetPLM_TugDn_1_vs_TugDn_3.jpg, (AP) esetPLM_TugDn_2_vs_TugDn_3.jpg, (AQ) esetPLM_TugDn2_1_vs_TugDn2_2.jpg, (AR) esetPLM_TugDn2_1_vs_TugDn2_3.jpg, (AS) esetPLM_TugDn2_2_vs_TugDn2_3.jpg, (AT) esetPLM_TugDn5_1_vs_TugDn5_2.jpg, (AU) esetPLM_TugDn5_1_vs_TugDn5_3.jpg, (AV) esetPLM_TugDn5_2_vs_TugDn5_3.jpg, (AW) esetMAS_Tug1_vs_Tug2.jpg, (AX) esetMAS_-Tug1_vs_Tug3.jpg, (AY) esetMAS_Tug2_vs_Tug3.jpg, (AZ) esetMAS_TugDn_1_vs_TugDn_-2.jpg, (BA) esetMAS_TugDn_1_vs_TugDn_3.jpg, (BB) esetMAS_TugDn_2_vs_TugDn_3.jpg, (BC) esetMAS_TugDn2_1_vs_TugDn2_2.jpg, (BD) esetMAS_TugDn2_1_vs_TugDn2_3.jpg, (BE) esetMAS_TugDn2_2_vs_TugDn2_3.jpg, (BF) esetMAS_TugDn5_1_vs_TugDn5_2.jpg, (BG) eset-MAS_TugDn5_1_vs_TugDn5_3.jpg, (BH) esetMAS_TugDn5_2_vs_TugDn5_3.jpg, (BI) eset-GCRMA_Tug1_vs_Tug2.jpg, (BJ) esetGCRMA_Tug1_vs_Tug3.jpg, (BK) esetGCRMA_-Tug2_vs_Tug3.jpg, (BL) esetGCRMA_TugDn_1_vs_TugDn_2.jpg, (BM) esetGCRMA_TugDn_-1_vs_TugDn_3.jpg, (BN) esetGCRMA_TugDn_2_vs_TugDn_3.jpg, (BO) esetGCRMA_TugDn2_-1_vs_TugDn2_2.jpg, (BP) esetGCRMA_TugDn2_1_vs_TugDn2_3.jpg, (BQ) esetGCRMA_-

TugDn2_2_vs_TugDn2_3.jpg, (BR) esetGCRMA_TugDn5_1_vs_TugDn5_2.jpg, (BS) esetGCRMA_TugDn5_1_vs_TugDn5_3.jpg, (BT) esetGCRMA_TugDn5_2_vs_TugDn5_3.jpg.

The 18 Slides files are: (A) rawdata_RWA2_GS_1_vs_RWA2_GS_2.jpg, (B) rawdata_RWA2_GS_1_vs_RWA2_GS_3.jpg, (C) rawdata_RWA2_GS_2_vs_RWA2_GS_3.jpg, (D) rawdata_RWA2_GR_1_vs_RWA2_GR_2.jpg, (E) rawdata_RWA2_GR_1_vs_RWA2_GR_3.jpg, (F) rawdata_RWA2_GR_2_vs_RWA2_GR_3.jpg, (G) rawdata_RWA1_GR_1_vs_RWA1_GR_2.jpg, (H) rawdata_RWA1_GR_1_vs_RWA1_GR_3.jpg, (I) rawdata_RWA1_GR_2_vs_RWA1_GR_3.jpg, (J) rawdata_RWA1_GS_2_vs_RWA1_GS_1.jpg, (K) rawdata_RWA1_GS_3_vs_RWA1_GS_1.jpg, (L) rawdata_RWA1_GS_3_vs_RWA1_GS_2.jpg, (M) rawdata_Gam_S_1_vs_Gam_S_2.jpg, (N) rawdata_Gam_S_1_vs_Gam_S_3.jpg, (O) rawdata_Gam_S_2_vs_Gam_S_3.jpg, (P) rawdata_Gam_R_1_vs_Gam_R_2.jpg, (Q) rawdata_Gam_R_1_vs_Gam_R_3.jpg, (R) rawdata_Gam_R_2_vs_Gam_R_3.jpg, (S) esetVSN_RWA2_GS_1_vs_RWA2_GS_2.jpg, (T) esetVSN_RWA2_GS_1_vs_RWA2_GS_3.jpg, (U) esetVSN_RWA2_GS_2_vs_RWA2_GS_3.jpg, (V) esetVSN_RWA2_GR_1_vs_RWA2_GR_2.jpg, (W) esetVSN_RWA2_GR_1_vs_RWA2_GR_3.jpg, (X) esetVSN_RWA2_GR_2_vs_RWA2_GR_3.jpg, (Y) esetVSN_RWA1_GR_1_vs_RWA1_GR_2.jpg, (Z) esetVSN_RWA1_GR_1_vs_RWA1_GR_3.jpg, (AA) esetVSN_RWA1_GR_2_vs_RWA1_GR_3.jpg, (AB) esetVSN_RWA1_GS_2_vs_RWA1_GS_1.jpg, (AC) esetVSN_RWA1_GS_3_vs_RWA1_GS_1.jpg, (AD) esetVSN_RWA1_GS_3_vs_RWA1_GS_2.jpg, (AE) esetVSN_Gam_S_1_vs_Gam_S_2.jpg, (AF) esetVSN_Gam_S_1_vs_Gam_S_3.jpg, (AG) esetVSN_Gam_S_2_vs_Gam_S_3.jpg, (AH) esetVSN_Gam_R_1_vs_Gam_R_2.jpg, (AI) esetVSN_Gam_R_1_vs_Gam_R_3.jpg, (AJ) esetVSN_Gam_R_2_vs_Gam_R_3.jpg, (AK) esetRMA_RWA2_GS_1_vs_RWA2_GS_2.jpg, (AL) esetRMA_RWA2_GS_1_vs_RWA2_GS_3.jpg, (AM) esetRMA_RWA2_GS_2_vs_RWA2_GS_3.jpg, (AN) esetRMA_RWA2_GR_1_vs_RWA2_GR_2.jpg, (AO) esetRMA_RWA2_GR_1_vs_RWA2_GR_3.jpg, (AP) esetRMA_RWA2_GR_2_vs_RWA2_GR_3.jpg, (AQ) esetRMA_RWA1_GR_1_vs_RWA1_GR_2.jpg, (AR) esetRMA_RWA1_GR_1_vs_RWA1_GR_3.jpg, (AS) esetRMA_RWA1_GR_2_vs_RWA1_GR_3.jpg, (AT) esetRMA_RWA1_GS_2_vs_RWA1_GS_1.jpg, (AU) esetRMA_RWA1_GS_3_vs_RWA1_GS_1.jpg, (AV) esetRMA_RWA1_GS_3_vs_RWA1_GS_2.jpg, (AW) esetRMA_Gam_S_1_vs_Gam_S_2.jpg, (AX) esetRMA_Gam_S_1_vs_Gam_S_3.jpg, (AY) esetRMA_Gam_S_2_vs_Gam_S_3.jpg, (AZ) esetRMA_Gam_R_1_vs_Gam_R_2.jpg, (BA) esetRMA_Gam_R_1_vs_Gam_R_3.jpg, (BB) esetRMA_Gam_R_2_vs_Gam_R_3.jpg, (BC) esetPLM_RWA2_GS_1_vs_RWA2_GS_2.jpg, (BD) esetPLM_RWA2_GS_1_vs_RWA2_GS_3.jpg, (BE) esetPLM_RWA2_GS_2_vs_RWA2_GS_3.jpg, (BF)

esetPLM_RWA2_GR_1_vs_RWA2_GR_2.jpg, (BG) esetPLM_RWA2_GR_1_vs_RWA2_GR_-3.jpg, (BH) esetPLM_RWA2_GR_2_vs_RWA2_GR_3.jpg, (BI) esetPLM_RWA1_GR_1_vs_-RWA1_GR_2.jpg, (BJ) esetPLM_RWA1_GR_1_vs_RWA1_GR_3.jpg, (BK) esetPLM_RWA1_-GR_2_vs_RWA1_GR_3.jpg, (BL) esetPLM_RWA1_GS_2_vs_RWA1_GS_1.jpg, (BM) eset-PLM_RWA1_GS_3_vs_RWA1_GS_1.jpg, (BN) esetPLM_RWA1_GS_3_vs_RWA1_GS_2.jpg, (BO) esetPLM_Gam_S_1_vs_Gam_S_2.jpg, (BP) esetPLM_Gam_S_1_vs_Gam_S_3.jpg, (BQ) esetPLM_Gam_S_2_vs_Gam_S_3.jpg, (BR) esetPLM_Gam_R_1_vs_Gam_R_2.jpg, (BS) eset-PLM_Gam_R_1_vs_Gam_R_3.jpg, (BT) esetPLM_Gam_R_2_vs_Gam_R_3.jpg, (BU) esetMAS_-RWA2_GS_1_vs_RWA2_GS_2.jpg, (BV) esetMAS_RWA2_GS_1_vs_RWA2_GS_3.jpg, (BW) esetMAS_RWA2_GS_2_vs_RWA2_GS_3.jpg, (BX) esetMAS_RWA2_GR_1_vs_RWA2_GR_2.-jpg, (BY) esetMAS_RWA2_GR_1_vs_RWA2_GR_3.jpg, (BZ) esetMAS_RWA2_GR_2_vs_-RWA2_GR_3.jpg, (CA) esetMAS_RWA1_GR_1_vs_RWA1_GR_2.jpg, (CB) esetMAS_RWA1_-GR_1_vs_RWA1_GR_3.jpg, (CC) esetMAS_RWA1_GR_2_vs_RWA1_GR_3.jpg, (CD) eset-MAS_RWA1_GS_2_vs_RWA1_GS_1.jpg, (CE) esetMAS_RWA1_GS_3_vs_RWA1_GS_1.jpg, (CF) esetMAS_RWA1_GS_3_vs_RWA1_GS_2.jpg, (CG) esetMAS_Gam_S_1_vs_Gam_S_2.jpg, (CH) esetMAS_Gam_S_1_vs_Gam_S_3.jpg, (CI) esetMAS_Gam_S_2_vs_Gam_S_3.jpg, (CJ) eset-MAS_Gam_R_1_vs_Gam_R_2.jpg, (CK) esetMAS_Gam_R_1_vs_Gam_R_3.jpg, (CL) esetMAS_-Gam_R_2_vs_Gam_R_3.jpg, (CM) esetGCRMA_RWA2_GS_1_vs_RWA2_GS_2.jpg, (CN) eset-GCRMA_RWA2_GS_1_vs_RWA2_GS_3.jpg, (CO) esetGCRMA_RWA2_GS_2_vs_RWA2_GS_-3.jpg, (CP) esetGCRMA_RWA2_GR_1_vs_RWA2_GR_2.jpg, (CQ) esetGCRMA_RWA2_GR_1_-vs_RWA2_GR_3.jpg, (CR) esetGCRMA_RWA2_GR_2_vs_RWA2_GR_3.jpg, (CS) eset-GCRMA_RWA1_GR_1_vs_RWA1_GR_2.jpg, (CT) esetGCRMA_RWA1_GR_1_vs_RWA1_-GR_3.jpg, (CU) esetGCRMA_RWA1_GR_2_vs_RWA1_GR_3.jpg, (CV) esetGCRMA_RWA1_-GS_2_vs_RWA1_GS_1.jpg, (CW) esetGCRMA_RWA1_GS_3_vs_RWA1_GS_1.jpg, (CX) eset-GCRMA_RWA1_GS_3_vs_RWA1_GS_2.jpg, (CY) esetGCRMA_Gam_S_1_vs_Gam_S_2.jpg, (CZ) esetGCRMA_Gam_S_1_vs_Gam_S_3.jpg, (DA) esetGCRMA_Gam_S_2_vs_Gam_S_3.jpg, (DB) esetGCRMA_Gam_R_1_vs_Gam_R_2.jpg, (DC) esetGCRMA_Gam_R_1_vs_Gam_R_3.jpg, (DD) esetGCRMA_Gam_R_2_vs_Gam_R_3.jpg.

Table Appx 5.2 Selected GeneIDs/probe sets shown to be up regulated (12 slide experiment), either generally or for specific slide comparisons (TuD-Tug, Tu2-Tug, *etc*.). Values under the different normalization methods (VSN, RMA, *etc*.) indicate the number of times the specific probe set was obtained across the six slide comparisons within that specific normalization method, *i.e.* 6 indicates that the geneID was obtained in all six slide comparisons. The slide comparisons indicate the number of times the geneID/probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences. *Table 5.1* shows examples of up regulated GeneIDs/probe sets.

Probe set ID	***VSN***	***RMA***	***PLM***	***MAS5***	***GCRMA***	***Occ. 95%***	***TuD - Tug***	***Tu2 - Tug***	***Tu5 - Tug***	***Tu2 - TuD***	***Tu5 - TuD***	***Tu5 - Tu2***	***Occ. 99%***	***Occ. 99.9%***
TaAffx.110208.1.S1_at	6	6	6	6	6	**30**	5	5	5	5	5	5	29	22
TaAffx.26346.1.S1_at	5	6	5	5	6	**27**	5	5	5	5	5	2	25	18
Ta.4593.1.A1_at	6	5	5	5	5	**26**	5	5	5	5	5	1	22	15
Ta.10581.1.A1_at	1	1	1	1	1	**5**	5	0	0	0	0	0	4	1
Ta.11261.1.S1_at	1	1	1	1	1	**5**	5	0	0	0	0	0	4	2
Ta.11832.1.S1_at	1	1	1	1	1	**5**	5	0	0	0	0	0	4	1
Ta.7584.1.S1_s_at	1	2	1	1	1	**6**	0	5	0	1	0	0	3	0
Ta.21925.1.S1_at	1	2	2	1	1	**7**	0	5	0	2	0	0	2	0
Ta.7361.1.A1_at	2	1	1	1	2	**7**	0	5	0	2	0	0	0	0
Ta.27788.2.S1_a_at	2	2	2	1	1	**8**	0	0	5	0	0	3	0	0
TaAffx.29630.1.S1_at	2	2	2	1	1	**8**	0	0	5	0	0	3	5	0
Ta.2107.1.S1_s_at	2	2	2	2	2	**10**	0	0	5	0	0	5	3	0
Ta.10008.1.A1_at	1	1	1	1	1	**5**	0	0	0	5	0	0	3	0
Ta.10426.1.A1_at	1	1	1	1	1	**5**	0	0	0	5	0	0	4	0
Ta.10552.1.S1_at	1	1	1	1	1	**5**	0	0	0	5	0	0	0	0
Ta.1040.1.S1_at	1	1	1	1	1	**5**	0	0	0	0	5	0	0	0
Ta.22932.1.S1_x_at	2	1	2	1	1	**7**	0	0	0	0	5	2	2	0
Ta.959.1.S1_at	2	2	1	1	1	**7**	0	0	0	0	5	2	2	0
Ta.10520.1.S1_at	1	1	1	1	1	**5**	0	0	0	0	0	5	1	0
Ta.10883.2.S1_at	1	1	1	1	1	**5**	0	0	0	0	0	5	3	0
Ta.10891.1.S1_at	1	1	1	1	1	**5**	0	0	0	0	0	5	3	0

Table Appx 5.3 GeneIDs/probe sets, selected in *Table 5.1*, that were still shown to be differentially regulated after FDR (Benjamini-Hochberg method) or FWER (Bonferroni) correction. Values under the different normalization methods (VSN, RMA, *etc.*) indicate the number of times the specific probe set was obtained across the six slide comparisons (TuD-Tug, Tu2-Tug, *etc.*) within that specific normalization method, *i.e.* 6 indicates that the geneID was obtained in all six slide comparisons. The slide comparisons indicate the number of times the geneID/probe set was obtained after normalization with the five methods, for that specific comparison, *i.e.* a 5 indicates that all five normalization methods detected that probe set for the specific slide comparison. Values summed for the normalization methods equals the values summed for the slide comparisons and are given by the "occurrence" (Occ.) value at specific confidences. *Table 5.2* shows examples of differentially up regulated GeneIDs/probe sets after FDR/FWER correction.

Probe set ID	***VSN***	***RMA***	***PLM***	***MAS5***	***GCRMA***	***Occ. 95%***	***TuD - Tug***	***Tu2 - Tug***	***Tu5 - Tug***	***Tu2 - TuD***	***Tu5 - TuD***	***Tu5 - Tu2***	***Occ. 99%***	***Occ. 99.9%***
After FDR														
TaAffx.110208.1.S1_at	3	4	5	3	4	**19**	0	3	5	1	5	5	16	11
TaAffx.26346.1.S1_at	3	3	3	3	2	**14**	0	4	5	0	5	0	8	3
Ta.4593.1.A1_at	3	3	3	3	1	**13**	0	4	5	0	4	0	7	4
TaAffx.29630.1.S1_at	0	1	1	0	0	**2**	0	0	2	0	0	0	0	0
Ta.2107.1.S1_s_at	0	0	1	0	0	**1**	0	0	1	0	0	0	0	0
Ta.22932.1.S1_x_at	1	0	1	0	0	**2**	0	0	0	0	2	0	0	0
Ta.959.1.S1_at	1	1	0	0	0	**2**	0	0	0	0	2	0	0	0
Ta.10883.2.S1_at	1	1	1	0	0	**3**	0	0	0	0	0	3	0	0
Ta.10891.1.S1_at	1	0	1	0	0	**2**	0	0	0	0	0	2	0	0
After FWER														
TaAffx.110208.1.S1_at	1	1	3	1	1	**7**	0	0	5	0	1	1	5	2
TaAffx.26346.1.S1_at	0	1	0	0	0	**1**	0	0	1	0	0	0	0	0

CURRICULUM VITAE

Zacharias Hendrik Swanevelder was born June 1st, 1978 in Ventersdorp, North-West Province. He attended the High School Verwoerdburg where he matriculated with five distinctions in 1996. He obtained his *Baccalaureus Scientiae* degree *cum laude* at the University of South Africa in 1999, majoring in Botany and Biochemistry. The degree *Baccalaureus Scientiae Honores*, with specialisation in Plant Physiology, was awarded *cum laude* in the Department of Botany, University of Pretoria in 2000. For his achievements he was awarded Academic Colours by the University of Pretoria and received the *Margaretha Mess*-medal for the best BSc (Hons) Botany student (2000).

In his *Magister Scientiae* entitled, Diversity and population structure of *Clivia miniata* Lindl. (Amaryllidaceae): Evidence from molecular genetics and ecology, he focused on the development of molecular markers, the ecology and taxonomy of the indigenous genus *Clivia*, and conducted a molecular marker analysis of the genus. The degree was obtained *cum laude* in 2003, where after he again was bestowed Academic Colours from the University of Pretoria. His MSc produced four scientific papers in peer review journals, entitled: A new species of *Clivia* (Amaryllidaceae) endemic to the Pondoland Centre of Endemism, South Africa; Amaryllidaceae: a new variety in the genus *Clivia*; Amaryllidaceae: a new variety of *Clivia robusta* and Amaryllidaceae: a natural hybrid in the genus *Clivia*, as well as other articles in popular journals. He is also the recipient of a NRF PhD Prestigious Bursary.

In his PhD, entitled: Aphid-Plant interactions and the possible role of an endosymbiont in aphid biotype development, he investigates essential amino acid biosynthetic plasmid genes of the aphid endosymbiont, *Buchnera aphidicola*, from different Russian wheat aphid biotypes and the possible role of this bacterium in the development of the new aphid biotypes. The aphid-plant interaction, from a plant's perspective, is also investigated on a molecular level with special focus on the influences that statistical normalization methods have on identifying differentially regulated genes after Russian wheat aphid infestation. The study resulted in the peer reviewed paper: Limited endosymbiont variation in *Diuraphis noxia* (Hemiptera: Aphididae) biotypes from the USA and South Africa, and also contributed to two additional papers entitled: Deciphering defense strategies that are elucidated in wheat containing different *Dn* resistance genes, and Transcript profiling of wheat genes expressed during feeding by two different biotypes of *Diuraphis noxia*. In his PhD he presented at various international and national research meetings, where he also received a special

award for a presentation in 2008 at the Joint Meeting of the WERA66 (Integrated Management of Russian Wheat Aphid and other Cereal Arthropod Pests) and IPRI (International Plant Resistance to Insects Workshop). In 2009, he was the author of a book entitled: *Clivias*: Nature and Nurture.