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

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

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

"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?