Epigenetic and transcriptomic mechanisms of polyphenism in an aphid model

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2022

Zhe Yang Yim

School of Biological Sciences

Division of Evolution, Infection and Genomics.

Table of Contents

List of Figures 12

List of Tables 16

Abbreviations 18

Abstract 24

Declaration 26

Copyright statement 27

Contributions from collaborators 29

Acknowledgments 30

The author 31

Chapter 1: Introduction 33

1.1 Polyphenism, trade-offs and evolution 34

1.2 Life history theory 38

1.2.1 The trade-off among life-history traits 38

1.2.2 Phenotypic plasticity, ecology and evolution 41

b) The limitations and cost of plasticity 43

c) Type of environmental conditions and organism characteristics that aid the evolution of plasticity 44

d) Plasticity, colonisation and response towards environmental changes 45

e) Plasticity and genetic evolution 47

f) Plasticity and ecological speciation 48

g) The rate of plasticity evolution 49

h) The effect of plasticity on communities and ecosystems 50

1.3 The history of epigenetics 51

1.3.1 Mechanisms involved in epigenetics 54

1.3.2 Histone modifications 56

1.3.3 Methylation differences between vertebrates and invertebrates and within invertebrates 59

1.3.4 The role of epigenetics in insect development 64

1.3.5 Ecological epigenetics 65

1.3.6 miRNA regulation 69

a) Role of miRNA in insect growth and development 72

b) Role of miRNA in insect behaviour 72

c) Role of miRNA in insect oogenesis and embryogenesis 73

d) Role of miRNA in insects wing development 73

e) Role of miRNA in insect phenotypic plasticity 74

1.4 General introduction and history of *Buchnera* 75

1.4.1 Genome structure of *Buchnera aphidicola* 76

1.4.2 Genome evolution in *Buchnera* 77

1.4.3 Relationship between *Buchnera* and pea aphids 79

a) Lateral gene transfer 80

b) Host immunity 81

c) Symbiotic metabolism (nutritional relationship) 82

d) Regulation of gene expression in *Buchnera* and aphids 86

1.4.4 Secondary endosymbionts in aphids 87

a) *Rickettsia* 87

b) *Hamiltonella defensa* 88

c) *Wolbachia pipientis* 89

d) *Regiella insecticola* 90

1.5 Summary 91

1.6 Hypothesis 91

1.7 Thesis aims 92

Chapter 2: Methods 95

2.1 Materials 96

2.1.1 Aphid stock maintenance 96

2.1.2 Faba bean sowing (stock plant) 97

2.1.3 Stock plant and experiment plant maintenance 99

2.2 Mesocosm experiment 100

2.2.1 Alternative morph production through crowding condition 100

2.3 Microcosm experiment 101

2.3.1 Reproductive success of alternative morphs (F0 generation) and the subsequent offspring (F1) 101

2.4 Molecular analyses 102

2.4.1 Tissue collection 102

2.4.2 Quantitative Reverse Transcriptase qPCR 102

a) RNA extraction 102

b) RNA quantification and purity 102

c) RNA integrity assessment 103

d) Complementary DNA (cDNA) synthesis 103

e) Reference gene selection by geNorm analysis 104

f) Gene expression analysis using SYBR Green 107

2.4.3 Pyrosequencing analysis 111

a) DNA extraction 111

b) DNA quantification and purity 111

c) DNA integrity assessment 112

d) Bisulfite conversion 112

e) Pyrosequencing analysis 116

2.5 Western blot analysis 116

a) Protein lysate extraction 116

b) Protein loading quantification 116

c) Western blot 117

2.6 Statistics 118

2.7 Summary 119

Chapter 3: Response of pea aphid genotypes to environmental stress and the trade-offs between reproduction and dispersal morphs 120

3.1 Introduction 121

3.2 Methods 125

3.2.1 Aphid colony 125

3.2.2 Experiment plant maintenance 125

3.2.3 Crowded conditions for triggering of alternative morphs 125

3.2.4 Reproductive success of alternative morphs (F0 generation) and their subsequent offspring (F1 generation) 126

3.2.5 Statistics 126

3.3 Results 127

3.3.1 Morph production in different pea aphid genotypes 127

3.3.2 Total population size of different pea aphid genotypes 128

3.3.3 Plant fitness 129

3.3.4 Reproductive success of different pea aphid morphs (T1) 130

3.3.5 Reproductive success of offspring (F1) produced from different aphid morphs 132

3.4 Discussion 133

3.4.1 Aphid genotype response to environmental stimulus and the degree of alternative morphs production 134

3.4.2 Reproductive success in different pea aphid morphs 137

3.4.3 Reproductive success of offspring from different aphid morphs 139

3.5 Summary 141

Chapter 4: Expression of developmental genes and transcriptome profiles of dispersal and non-dispersal aphid in two aphid genotypes 143

4.1 Introduction 144

4.2 Methods 148

4.2.1. Materials 148

4.2.2 Candidate gene expression by qPCR 148

4.2.3 Transcriptome sequencing 149

4.2.4 Sequence assembly and annotation 149

4.2.5 Protein expression by Western blot 150

4.3 Results 151

4.3.1 RNA-seq analysis 151

4.3.2 Differentially expressed genes, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis between genotype and morphs 151

a) Differentially expressed genes between different group comparison 153

b) Differentially expressed genes, GO and KEGG between N116 winged and N116 wingless aphids 154

c) Differentially expressed genes, GO and KEGG between N127 pale and N127 red aphids 160

d) Differentially expressed gene, GO and KEGG between N127 red and N116 green wingless aphids 166

e) Differentially expressed genes, GO and KEGG between N127 pale and N116 winged aphids 172

f) Differentially expressed genes, GO and KEGG between N127 (red+pale) and N116 (winged+wingless) aphids 178

g) Differentially expressed genes, GO and KEGG between alternative morphs (N127 pale+N116 winged) and wild type (N127red+N116 wingless) 184

h) Differentially expressed genes, GO and KEGG between interaction of the aphid genotype and aphid morphs 189

4.3.3 Gene expression 194

a) geNorm analysis 194

b) RNA-seq gene validation 196

4.3.4 Protein expression 200

4.4 Discussion 201

4.4.1 Gene expression 203

4.4.2 Transcriptome profile expression 218

4.5 Summary 230

Chapter 5: DNA methylation of pea aphid developmental genes and methylome profile in different pea aphid morphs 232

5.1 Introduction 233

5.2 Methods 237

5.2.1 Materials 237

5.2.2. DNA methylation level by pyrosequencing 237

5.2.3 MBD-seq library preparation 237

5.2.4 Sequence assembly and annotation 238

5.2.5 Integration of RNA-seq and MBD-seq data 239

5.3 Results 240

5.3.1 DNA methylation levels 240

5.3.2 MBD-seq analysis 245

5.3.3 Methylome profiles of pea aphid genotypes and their morphs 245

a) Differentially methylated genes between different all group comparison 246

b) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N116 winged and N116 wingless aphids 248

c) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N127 pale and N127 red aphids 251

c) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N127 red and N116 green wingless aphids 254

d) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N127 pale and N116 winged aphids 258

e) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N127 (red+pale) and N116 (winged+wingless) aphids 262

5.3.4 Correlation between DNA methylation and gene expression between pea aphid morphs and genotypes 266

a) Differentially methylated genes and differentially expressed genes overlap between N116 winged and N116 wingless aphids 268

b) Differentially methylated genes and differentially expressed genes overlap between N127 pale and N127 red aphids 270

c) Differentially methylated genes and differentially expressed genes overlap between N127 red and N116 green wingless aphid 272

d) Differentially methylated genes and differentially expressed genes overlap between N127 pale and N116 winged aphids 274

e) Differentially methylated genes and differentially expressed genes overlap between N127 (red+pale) and N116 (winged+wingless) aphids 276

5.4 Discussion 278

5.4.1 DNA methylation level of candidate genes 278

5.4.2 Methylome profile expression 280

5.3.3 DNA methylation and gene expression 284

5.5 Summary 288

Chapter 6: miRNA-profile of dispersal and non-dispersal morphs in two aphid genotypes 290

6.1 Introduction 291

6.2 Methods 293

6.2.1. Materials 293

6.2.2 miRNA library preparation 293

6.2.3 Sequence assembly and miRNA annotation 294

6.3 Results 296

6.3.1 miRNA-seq analysis 296

6.3.2 Differentially expressed genes, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis between genotype and morphs 296

a) Differentially expressed miRNAs, GO and KEGG between N116 winged and N116 wingless aphids 298

b) Differentially expressed miRNAs, GO and KEGG between N127 red and N116 green wingless aphid 303

6.3.3 Integration of miRNA-seq and RNA-seq 310

a) miRNA targets genes that overlap with RNA-seq (differentially expressed genes) in N116 winged and N116 wingless aphids. 310

b) miRNA targets genes that overlap with RNA-seq (differentially expressed genes) in N127 red and N116 green wingless aphids. 312

6.4 Discussion 314

6.5 Summary 323

Chapter 7: General discussion 324

7.1 General discussion 325

7.2 Future directions 331

7.2.1 The trade-offs between different pea aphid morphs 331

7.2.2 The transcriptome profile and mechanisms of aphid development and polyphenism 332

7.2.3 The role of epigenetic mechanisms in regulating pea aphid wing polyphenism and development 333

7.3 Limitations 334

7.4 Summary 335

References 337

Word Count: 82700

List of Figures

Chapter 1

Figure 1.1 Reaction norm graph 34
Figure 1.2 Caste-polyphenism in carpenter ant 35
Figure 1.3 Life cycle and polyphenism in the pea aphid 37
Figure 1.4 The Y model in trade-off 40
Figure 1.5 Distribution and pattern of selection on plasticity in animals and plants 44
Figure 1.6 Plasticity in marine organism with different dispersal rate 45
Figure 1.7 Baldwin effect graph 46
Figure 1.8 Schematic flow for ecological speciation 48
Figure 1.9 Difference between Mendelian genetics and epigenetics 53
Figure 1.10 Summary of the mechanism of DNA methylation 54
Figure 1.11 Function of the DNMT family 56
Figure 1.12 Nucleosome structure 57
Figure 1.13 Summary of cellular process affected by histone modification 59
Figure 1.14 Distribution of the methylation pattern of genes across insects 60
Figure 1.15 DNMT family distribution pattern in different species of insects 62
Figure 1.16 DNA methylation level across insect order 63
Figure 1.17 Pathway of genetics and epigenetics in ecological context 65
Figure 1.18 miRNA biogenesis pathway 71
Figure 1.19 Phylogenetic tree from eubacteria and mitochondria using 16S rRNA gene 75
Figure 1.20 Genome of *Buchnera aphidicola* and reconstructed ancestor 78
Figure 1.21 Expression profile of genes thought to be laterally transferred 80
Figure 1.22 Simplified schematic diagram of amino acid production in Buchnera-aphid symbiosis 83
Figure 1.23 Amino acid pathway in Buchnera with aphid gene complementation 84

Chapter 2

Figure 2.1 Stock Aphid Maintenance 97
Figure 2.2 Features of stock plant maintenance 99
Figure 2.3 Mesocosm container set up for triggering alternative morphs in pea aphids 100
Figure 2.4 Transfer experiment set up 101
Figure 2.5 Experimental design for crowded conditions and transfer experiments 119

Chapter 3

Figure 3.1 Aphid life cycle 122
Figure 3.2 Percentage of morph production in two different pea aphid morphs 127
Figure 3.3 Total population size in two different pea aphid morphs 128
Figure 3.4 Effect of percentage of morphs on aphid total population size 129
Figure 3.5 Effect of plant survival rate on plant S:R ratio 130
Figure 3.6 Differences in reproduction between offspring from dispersal and non-dispersal mothers in two genotypes and the interactions between alternative morphs vs wild type and genotype 131
Figure 3.7 Differences in reproduction between offspring (F1) from dispersal and non-dispersal mothers in two genotypes 133

Chapter 4

Figure 4.1 The wing-formation gene network of pea aphids deduced from *Drosophila melanogaster* 145
Figure 4.2. Principal component analysis of gene expression of five pea aphid morphs 152
Figure 4.3. Venn diagram comparing DEG identified by RNA-seq between different aphid morphs 154
Figure 4.4 GO enrichments between N116 winged vs N116 wingless aphids 158
Figure 4.5 KEGG classification between N116 winged and N116 wingless aphids 159
Figure 4.6 GO enrichments between N127 pale and N127 red aphids 164
Figure 4.7 KEGG classification between N127 pale and N127 red aphids 165
Figure 4.8 GO enrichment between N127 red and N116 green wingless aphid 170
Figure 4.9 KEGG classification between N127 red and N116 green wingless aphid 171
Figure 4.10 GO enrichments between N127 pale and N116 winged aphids 176
Figure 4.11 KEGG classification between N127 pale and N116 winged aphids 177
Figure 4.12 GO enrichments between N127 (pale+red) and N116 (wingless+winged) aphids 182
Figure 4.13 KEGG classification between N127 (red+pale) and N116 (winged+wingless) aphids 183
Figure 4.14 GO enrichment between alternative morphs (N127 pale+N116 winged) and wildtype (N127 red+N116 wingless) 187
Figure 4.15 KEGG classification between alternative morphs (N127 pale+N116 winged) and wildtype (N127 red+N116 wingless) 188
Figure 4.16 GO enrichment in the interaction of aphid genotypes and aphid morphs 192

Figure 4.17 KEGG classification in the interaction of aphid genotypes and aphid morphs 193
Figure 4.18 Representative geNorm analysis using qBase+ software to identify the optimal number of reference genes 195
Figure 4.19 Validation through qPCR of RNA-seq differentially expressed genes between aphid morphs 198
Figure 4.20 Expression level of candidate genes between aphid morphs 199
Figure 4.21 Protein level of Hsp70 (against N116 wingless) 200

Chapter 5

Figure 5.1 The regulation framework for insect polyphenism 234
Figure 5.2 DNA methylation level of candidate genes between aphid morphs (from RNA-seq) 243
Figure 5.3 DNA methylation levels of candidate genes between aphid morphs 244
Figure 5.4. Principal component analysis of gene expression of four pea aphid Morphs 246
Figure 5.5. Venn diagram comparing DMG identified by MBD-seq between different aphid morphs 247
Figure 5.6 MBD-seq results for N116 winged vs N116 wingless aphids 250
Figure 5.7 MBD-seq results for N127 pale vs N127 red aphids 253
Figure 5.8 MBD-seq results for N127 red vs N116 green wingless aphids 256
Figure 5.9 GO enrichments analysis between N127 red vs N116 green wingless aphids 257
Figure 5.10 MBD-seq results for N127 pale vs N116 winged aphids 260
Figure 5.11 GO enrichments analysis between N127 pale vs N116 winged morphs 261
Figure 5.12 MBD-seq results for N127 (pale+red) and N116 (winged+wingless) aphids 264
Figure 5.13 GO enrichments analysis between N127 (pale+red) and N116 (wingless+winged) aphids 265
Figure 5.14. Correlation between DNA methylation and gene expression (gene body region only) for different group comparisons 267
Figure 5.15 MBD-seq and RNA-seq for N116 winged and N116 wingless aphids 269
Figure 5.16 MBD-seq and RNA-seq for N127 pale and N127 red aphids 271
Figure 5.17 MBD-seq and RNA-seq for N127 red and N116 wingless aphids 273
Figure 5.18 MBD-seq and RNA-seq for N127 pale and N116 winged aphids 275
Figure 5.19 MBD-seq and RNA-seq for N127 (pale+red) and N116 (winged+wingless) aphids 277

Chapter 6

Figure 6.1. Distribution of different type of RNA in the microRNA-seq 296
Figure 6.2. Venn diagram comparing number of differentially expressed miRNA identified by microRNA-seq between different aphid morphs 297
Figure 6.3 Number of genes regulated by each miRNA 299
Figure 6.4 GO enrichment analysis between N116 winged vs N116 wingless aphids 301
Figure 6.5 Number of genes regulated by each miRNA 304
Figure 6.6 GO enrichment analysis between N127 red vs N116 green wingless aphids 307
Figure 6.7 KEGG classification between N127 red and N116 green wingless aphids 308
Figure 6.8 Integration of miRNA with RNA-seq data between N116 winged and N116 wingless aphids 310
Figure 6.9 Integration of miRNA with RNA-seq data between N127 red and N116 green wingless aphids 312

List of Tables

Chapter 2

Table 2.1 Differences between life-history traits of two aphid genotypes (N116 vs N127) 96
Table 2.2 List of pea aphid reference genes used in the geNorm analysis 106
Table 2.3 Candidate genes for qPCR analysis. 108
Table 2.4 Candidate gene selected from Differentially expressed gene (DEG) list that are used for qPCR validation 109
Table 2.5. Candidate genes used for pyrosequencing analysis 113

Chapter 3

Table 2.1 Summary of statistical model used for the transfer experiment 1 and 2 analysis and the significant predictors for each gene and random effects in each model. 132

Chapter 4

Table 4.1 Summary of morph numbers used for molecular analysis in Chapter 4 148
Table 4.2 Number of raw reads and unique mapped reads obtained from RNA seq 151
Table 4.3 Differentially expressed genes between N116 winged and N116 wingless aphids 156
Table 4.4 Differentially expressed genes between N127 pale and N127 red aphids 162
Table 4.5 Differentially expressed genes between N127 red and N116 green wingless aphids 168
Table 4.6 Differentially expressed genes between N127 pale and N116 winged aphids 174
Table 4.7 Differentially expressed genes between N127 (red+pale) and N116 (winged+wingless) aphids 180
Table 4.8 Differentially expressed genes between alternative morphs (N127 pale+N116 winged) vs wildtype (N127 red+N116 wingless) 186
Table 4.9 Differentially expressed gene between interaction of aphid genotypes and aphid morphs 191
Table 4.10 Summary of statistical model used for gene expression analysis and the significant predictors for each gene 196

Chapter 5

Table 5.1 Summary of aphid numbers used for molecular analysis in Chapter 5 237

Table 5.2 Summary of statistical model used for gene expression analysis and the significant predictors for each genes 240

Table 5.3 Number of raw reads and unique mapped reads obtained from MBD- seq 245

Table 5.4 Top 10 most hypermethylated and hypomethylated in N116 winged vs N116 wingless aphids 249

Table 5.5 Top 10 most hypermethylated and hypomethylated in N127 pale vs N127 red aphids 252

Table 5.6 Top 10 most hypermethylated and hypomethylated in N127 red vs N116 green wingless aphids 255

Table 5.7 Top 10 most hypermethylated and hypomethylated in N127 pale vs N116 winged aphids 259

Table 5.8 Top 10 most hypermethylated and hypomethylated in N127 (red+pale) and N116 (winged+wingless) aphids 263

Chapter 6

Table 6.1 Summary of morph numbers used for molecular analysis in Chapter 6 292

Table 6.2 Number of raw reads and uniquely mapped reads obtained from miRNA seq 296

Table 6.3 Differentially expressed conserved and novel miRNA in N116 winged vs wingless aphids 300

Table 6.4 Differentially expressed conserved and novel miRNA in N127 red and N116 green wingless aphids 305

Abbreviations

5-Azac	5-azacytidine
DNMT	DNA methyltransferase
Dnmt1	DNA methyltransferase 1
Dnmt2	DNA methyltransferase 2
Dnmt3	DNA methyltransferase 3
Dnmt3L	DNA methyltransferase 3L
Dnmt3A	DNA methyltransferase 3B
Dnmt3B	DNA methyltransferase 3B
S phase	Synthesis phase
SAMe	S-adenosyl-L-Methionine
K	Lysine
H3K4	Histone 3 Lysine 4
H3K27	Histone 3 Lysine 27
H3K9	Histone 3 Lysine 9
miRNA	microRNA
UTR	Untranslated region
3'UTR	3' Untranslated region
5'UTR	5' Untranslated region
pri-miRNA	Primary miRNA
Dcr-1	Dicer 1
Ago	Argronaute
RISC	RNA-induced silencing complex
GXE	Gene x Environment
Eu	eursytomatous
St	Stenostomatous
Hh	Hedgehog
20E	20-hydroxyecdysone
dLMO	*Drosophila* LIM-only
16S rRNA	16 ribosomal RNA

MYA	Million years ago
23S rRNA	23 ribosomal RNA
5S rRNA	5 ribosomal RNA
LdcA	LD-carboxypeptidase
RlpA	Rare lipoprotein A
AmiD	N-acetylmuramoyl-L- alanine amidase
PGRPs	Peptidoglycan receptor proteins
IMD	Immunodeficiency
Ilve	Branched-chain amino acid transaminase
IlvA	Isoleucine pathway synthesis
MetB/C	Methionine biosynthesis pathway
APSE	Lambda-like bacteriophage
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
ME	2 Mercaptoethanol
gDNA	Genomic DNA
PCR	Polymerase chain reaction
NTC	No template control
-RT	No reverse transcriptase
Ct	Cycle threshold
SDHB	Succinate dehydrogenase B
NADH	NADH dehydrogenase
RPL12	Ribosomal protein L12
18S	18 ribosomal RNA
cDNA	Complementary DNA
fl	Flightin
Hsp83	Heat shock protein 83
SCAD	Short chain specific acyl-coA dehydrogenase
vg	Vestigial
en	Engrailed
ap1	Apterous 1

tor	Carotene dehydrogenase
TnC	Troponin C
Mad	Mothers against decapentaplegic homolog 4
Pburs	Partner of bursicon
DMAP1	DNA methyltransferase 1- associated protein 1
Kr-h1	Krueppel homolog 1
Rpd3	Histone deacetylase Rpd3
Treh	Trehalase
Shd	Ecdysone 20-monooxygenase
Apns1	Acyrthosiphon pisum densovirus 1
TBE	Tris-Borate-EDTA buffer
GLMs	General linear models
GLMMs	General linear mixed models
AM_WT	Alternative morph vs wild type
EFA	Exploratory factor analysis
SEM	Standard error mean
Dpp	Decapentaplegic
Sal	Spalt major
Ubx	Ultrabithorax
Ser	Serrate
Wg	Wingless
Ac	Achaete
Sc	Scute
Dll	Distalless
Hth	Homothorax
qPCR	Quantitative polymerase chain reaction
GTCF	Genome Technologies Core Facility
BIPAA	Bioinformatics Platform for Agroecosystem Arthropods
DEG	Differentially expressed genes
BP	Biological Process
MF	Molecular Function

CC	Cellular components
DFM	Direct flight muscles
IFM	Indirect flight muscles
DVM	Dorsoventral muscles
DLM	Dorsal longitudinal muscles
ODM	Oblique dorsal muscle
Sd	Scalloped
TnI	Troponin I
TnT	Troponin T
JH	Juvenile hormone
EcR	Ecdysone receptor
Tkv	Thickveins
HATs	Histone acetyltransferases
Atro	Atrophin
ApoD	Apolipoprotein D
P38	P38 mitogen-activated protein kinase
EH	Ecolsion hormone-like
Awd2	Abnormal wing disc 2
PO	Phenoloxidase 1
DMAD	DNA N6-methyl adenine demethylase
E(z)	Histone-lysine-N-methyltranferase E(z)
EP300	Histone acetyltransferase p300
LSD-1	Lipid storage surface-binding protein 1
IRS	Insulin receptor substrate 1
IGF	Insulin-like peptide receptor
FoxO	Forkhead protein
FH	Succinate dehydrogenase
OXCT1	Isocitrate dehydrogenase
PDHB	Pyruvate dehydrogenase E1 component subunit beta
PGK	Probable phosphoglycerate kinase
RR1	RR1 cuticle protein

CP	Cuticle protein
Cox10	Protoheme IX farnesyltransferase, mitochondrial like
HDAC8	Histone deacetylase 8
Kat-7	Histone acetyltansferase KAT7-like
E4	Esterase E4
P450	P450 4C1-like
Eip78C	Ecdysone induced protein 78C
Gus	Gustavus
Mth2	G-couple protein receptor Mth2
P38 MAPK	Mitogen-activated protein kinase p38b
Gr64f	Gustatory receptor for sugar taste 64-like
Tret1	Facilitated trehalose transporter Tret1
Hsp70	Heat shock protein 70
Idgf	Imaginal disc growth factor precursor
Hsp68	Heat shock protein 68
Hsp75	Heat shock protein 75
PEP	Phosphoenol pyruvate
Csp	Chemosensory protein-like precursor
Cp7	Cuticle protein 7
MBD-seq	Methyl-binding-domain sequencing
NA	Not available
DMR	Differentially methylated region
DMG	Differentially methylated genes
RIN	RNA integrity number
DRA	Dimer removal agent
Wnt2	Protein Wnt2
RR1	RR1 cuticle
Sap18	Histone deacetylase complex subunit SAP18
crtYB	Phytoene synthase
burs	Bursicon
JHBP	Juvenile hormone binding protein

Ds	Daschous
B-H1	Homeobox protein BH-1
Atx2	Ataxin2
Sax-3	Protein sax-3
Sca	Protein scabrous
N	Notch
PDS	Phytoene desaturase
P300	Histone acetyltransferase p300
Gld	Glucose dehydrogenase
Api	Aphicarus

Abstract

Polyphenism is the ability of an organism's genotype to produce multiple discrete phenotypes when exposed to environmental stress. Polyphenism is very important to organisms as it allows them to react quickly to changes in environmental stress and therefore increase their chance of survival in the changing environment. There are many different types of polyphenism, including phase polyphenism, wing polyphenism and others. Aphids are one of the key model organisms used in studying polyphenism. Under normal conditions, aphids reproduce asexually and all the offsprings are essentially clones of the mother. However, upon exposure to environmental stress such as crowding, predation, and depletion of food resources, aphids exhibit wing polyphenism whereby the offspring from the next generation will develop wings, which allows them to escape from poor environmental conditions. Even though the aphid model has been used to investigate polyphenism, the underlying mechanisms in regulating such wing polyphenism remain unclear.

The main experiment in this thesis involved exposing two different aphid genotypes (N116 and N127) to crowding conditions to trigger alternative morph production and the degree and number of adult alternative morphs were recorded (mesocosm experiment). Next, we carried out two transfer experiments to investigate the possibility of trade-offs between dispersal and reproduction. At the end of this mesocosm experiment, five 4th instar nymphs of each morph were transferred to individual plants to allow them to reproduce for another 12 days (Transfer experiment 1). The total number of offspring produced by each morph was recorded. At the end of transfer experiment 1, five 4th instar nymphs produced by each morph were transferred to another individual plant and allowed to reproduce for another 12 days and the total number of the population was recorded. The adult morphs collected from the mesocosm experiment were used to investigate the underlying mechanisms of wing polyphenism in aphids and the role of epigenetics mechanisms in regulating pea aphid wing polyphenisms through different techniques: qPCR, RNA-seq, MBD-seq, pyrosequencing and mIRNA-seq.

The two aphid genotypes reacted very differently to starvation conditions with N116 producing winged offspring in the next generation while N127 changed their body colour from red to pale. In addition, N127 also produced a higher percentage of alternative morphs compared to N116. We found that in both transfer experiments the dispersal morphs have lower reproduction compared to the wild type. This suggests a possible trade-off between dispersal and reproduction and such a trade-off could last for more than one generation. Our RNA-seq and qPCR identify several genes that could be involved in regulating wing development, reproduction and stress response in the pea aphid. Pyrosequencing revealed that among the genes that were differentially expressed between morphs only a few were differentially methylated. Further, our MBD-seq shows that the pea aphid has the highest methylation in the gene body region. However, integrating the MBD-seq and RNA-seq suggests that there was no clear correlation between DNA methylation and gene expression but some genotypes and morphs show stronger positive trends. In addition, we found that many conserved and some novel miRNAs were differentially expressed between the aphid morphs. Target gene prediction results from these miRNAs suggest that they target many genes that were involved in wing development, reproduction and wing development. Together the results from this thesis provide insight into genes that could be regulating wing polyphenism, stress response and reproduction in pea aphids and the role of different epigenetic mechanisms in regulating this polyphenism.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s

regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses

Contributions from collaborators

All experiments were performed by YZY other than:

- **Chapter 4:** RNA-seq were extracted by YZY and library preparation, sequencing and bioinformatics were performed by the Genome Technologies Core Facilities (GTCF)
- **Chapter 5:** MBD-seq was performed in two batches with the first batch of sequencing being done by Liam Hanson from Manchester Metropolitan University and bioinformatics analysis help by Dr Ian Donaldson.
- **Chapter 6:** miRNA seq (library preparation, sequencing and some part of bioinformatics analysis) were performed by the company RealSeqBiosciencs, USA.

Acknowledgments

Firstly, I would like to express my special thanks to my supervisor Dr. Reinmar Hager for welcoming me to Manchester and supporting and guiding me throughout my PhD. I have learned so much from my supervisor throughout this whole PhD journey which will help me to become a better scientist in the future. Next, I would like to express my gratitude to my external supervisor Dr. Christopher Murgatroyd for all the guidance in the lab and advice during my PhD which has helped to enhance my knowledge in NGS and in the field of epigenetics. I would also like to express many thanks to Dr. Mouhammad Shadi Khudr for all the advice and help in the lab during this whole PhD journey.

The whole PhD journey has been challenging especially with the pandemic going on. However, I have largely benefitted from the help of some amazing PhD students that I met during my PhD journey: Becky Woods, Hawa Jahan and Conrado Quilles who have contributed to my PhD project in one way or another.

At the end of a long day, I would like to thank a few of my best friends that I met in Manchester: Jozef Bača and Gracie Ducloux who have been so supportive and helped me get through each week. Also, I would like to express me special thanks to the International Society for making me feel at home during my stay in Manchester. I can't thank you all enough and hope we can celebrate soon. Lastly, I would love to say a big thank you to my family especially my parents and grandparents for the unconditional love, and unwavering support from afar and I hope I would make all of you proud.

The author

i) Publication

Yim Z. Y, Khudr M.S, Quiles C.G, Jahan H, Alfutimie A, Hager R. 2022. Clone-specific aphid resilience under competition and in the absence of light reveals surprising phenotypic plasticity (Preprint and submit to Insect Science)

ii) Conference presentation

DNA methylation and gene expression patterns show differential responses to stress in aphids developing different alternative morphs - XI International Anniversary Symposium on Aphids**. (Katowice, Poland) (Oral Presentation)**

Epigenetic and gene expression differences are associated with reproductive success in aphid morphs (talk) - Doctoral Academy Graduate Society PhD Conference **(University of Manchester, United Kingdom) (Oral Presentation)**

Epigenetic and gene expression differences are associated with reproductive success in aphid morphs – Congress of the European Society for Evolutionary Biology (ESEB) **(Prague, Czech Republic) (Poster Presentation)**

Molecular mechanisms associated with a trade-off in a parasite host-plant system – Dragons and Elevators event- School of Biological Sciences, Division of Evolution and Genomic Sciences **(University of Manchester, United Kingdom) (Poster Presentation)**

iii) Prizes

First prize in Scientific Image Competition – School of Biological Sciences, Division of Evolution and Genomic Sciences

First prize in group presentation for mock research grant proposal – Dragons and Elevators event- School of Biological Sciences, Division of Evolution and Genomic Sciences

iv) Travel grants

University of Manchester Doctoral Academy conference support fund, 2022 (£500)

Scholarship to attend the XI International Anniversary Symposium on Aphids in Poland (Full scholarship including registration, accommodation and food)

Chapter 1: Introduction

1.1 Polyphenism, trade-offs and evolution

Phenotypic plasticity is commonly defined as the ability of an organism genotype to produce a different phenotype in response to different environmental stimuli (West-Eberhard 2003). Phenotype changes encompass a broad range such as behavioural, physiological (metabolic rates) and morphological changes. Reaction norm is commonly used to describe the phenotypic variation of the same genotype when exposed to a range of environmental conditions (Arnold et al. 2019) (Figure 1.1).

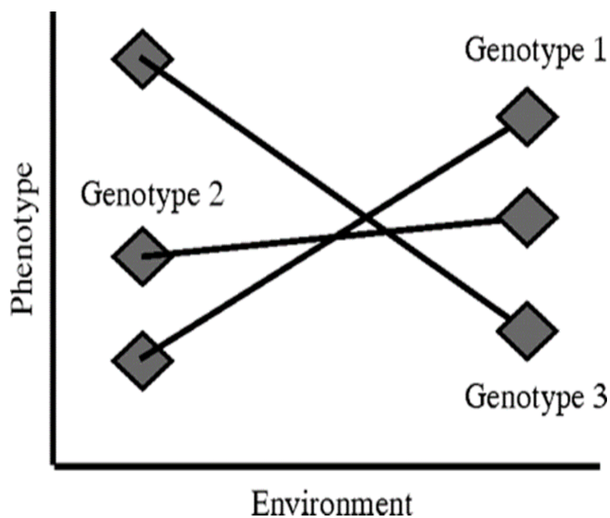


Figure 1.1 Reaction norm graph. The slope represents the distribution pattern of the plasticity (with either a positive response or negative response towards the environment with the lines representing the norm of reaction (change in the mean value of trait/phenotype in different conditions). For example, in genotype 1 and 3 both shows plasticity in the environment but with very different patterns, while genotype 2 shows no plasticity changes. (Adapted from Pigluicci et al., 2006).

A reaction norm can be used to derive the magnitude and pattern phenotypic variation of an organism produced from a single genotype when exposed to multiple different environments (Manuck 2010). However, different genotypes in a single species can show very different reaction norms in response to a specific phenotypic trait and environmental

stimulus. Due to the complexity of the interrelationships between genetic and environmental factors (GxE) in determining a trait, a wide range of reaction norms exist.

One of the most extreme cases of phenotypic variation is known as polyphenism, which is commonly defined as the ability of a single genotype to give rise to multiple phenotypes when exposed to different environmental conditions. Many different types of polyphenism are found among insects. One of the most commonly found polyphenism is observed in the Florida carpenter ant (*Camponotus floridanus*) which is involved in caste differentiation (Figure 1.2).

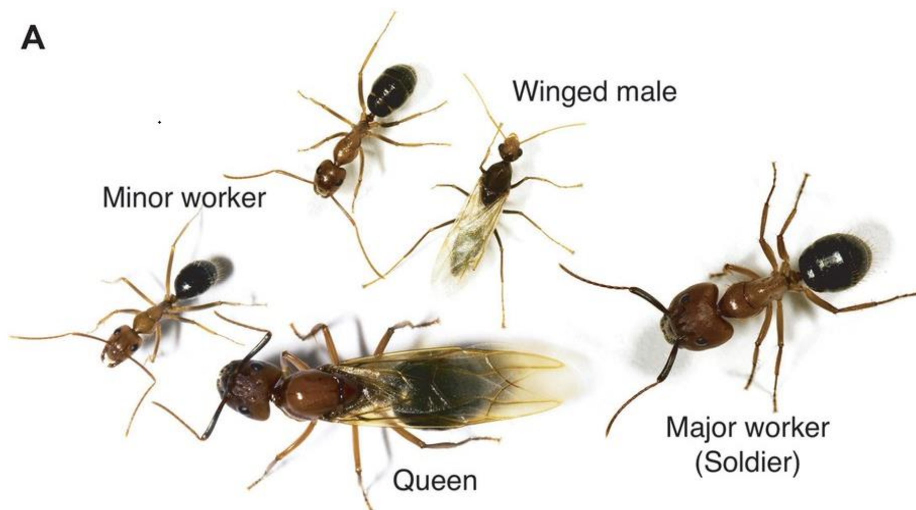


Figure 1.2. Caste-polyphenism in carpenter ant. The different castes in the carpenter ant with minor/major worker, winged male and the queen. (Adapted from Chittka et al. 2012).

In these ants, female embryos can typically result in either producing minor or major worker ants, or a queen (Chittka et al. 2012). This caste differentiation in the embryo depends on environmental stimuli such as diet and temperature, and that minor and major ant workers share a more similar methylation profile in comparison to ant queens (Bonasio et al., 2012). Major and minor ant workers are important for foraging and scouting activities; with both worker types exhibit unique histone acetylation patterns comparison with queen ants (Simola et al. 2016). Simola shown that inhibiting histone acetyltransferase inhibitor C646

caused defects in the foraging and scouting activities of major carpenter ant workers (Simola et al., 2016).

The second type of polyphenism is known as mouth-form polyphenism, which is most studied in roundworm (*Pristionchus pacificus*). Roundworm typically exhibit two different types of versatile teeth, i.e. denticles, which are commonly known as eursytomatous (Eu) or stenostomatous (St) (Ragsdale et al. 2013). Eu individual typically exhibit bigger and claw-like shaped dorsal teeth, which increase the beetle's opportunities for finding prey and feeding on other nematodes. In comparison, St individuals have smaller and less profound dorsal teeth, which typically results in lower predation success. Mouth determination generally happens during the larval stage and is irreversible. The St and Eu types commonly feed on bacteria; however, the Eu individuals can also rely on other food resources such as nematodes, thereby increasing their chance of survival when food availability is low (Bento et al., 2010).

The third type of polyphenism is commonly known as dispersal polyphenism. The primary difference seen in dispersal polyphenism is the presence or absence of winged musculature (wing polyphenism). One of the most common examples of wing polyphenism can be observed in pea aphids (*Acyrthosiphon pisum*), a type of sap-sucking insect that commonly feeds on legumes (Figure 1.3).

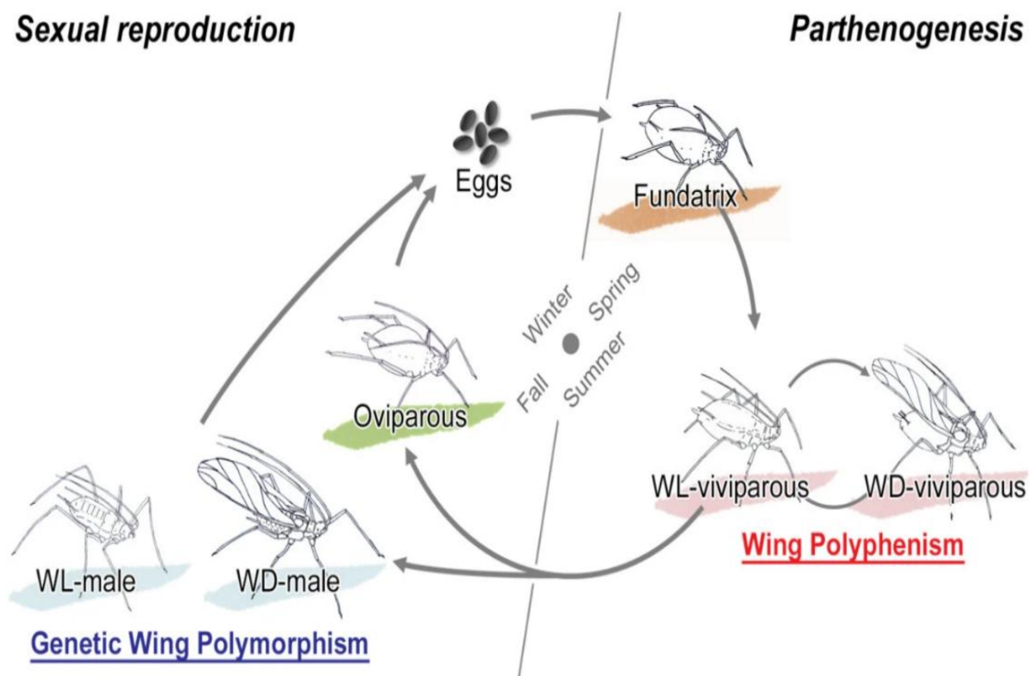


Figure 1.3. Life cycle and polyphenism in pea aphid. Pea aphids show wing polyphenism mainly during spring and summer time and the trigger of wing polyphenism depends on various environmental stimulus (food quality, population density, predation). In fall and winter, the mothers switch to sexual reproduction (sexual polyphenism) to produce males which will mate with the female mother to produce overwintering eggs. (Adapated from Ogawa and Miura 2014).

The pea aphid mother generally produces through asexual parthenogenesis and in normal conditions the offspring produced are typically wingless (Simpson et al. 2011). However, when the asexual mother is exposed to a variety of different environments, e.g. increased population density and low food quality, the offspring produced in the subsequent generation will typically be winged. Epigenetic processes such as DNA methylation occur in pea aphids; however, the extent to which the morph is regulated by epigenetics remains unclear (Zhang et al. 2019). Apart from wing polyphenism, pea aphids also exhibit another form of polyphenism known as sexual polyphenism. This happens during summer and autumn when pea aphids switch from asexual to sexual polyphenism, whereby the mother will produce both winged male and wingless female offspring. In contrast to the winged

female, the development of winged males occurs due to genetic differences in the aphicarus (api) on the X-chromosomes (Braendle et al. 2006).

1.2 Life history theory

One of the most well-known theories in evolutionary ecology is known as life-history theory. The main aim of this theory is to explain diversity in life histories across different species. The study of life history is essential in evolutionary context because the adaptation of a trait by natural selection depends on the Darwinian fitness among individuals. Thus, it is important to study the life history traits that commonly affect the fitness (survival and reproduction) of organisms (Fabian and Flatt 2012). The fundamental aspect of life-history traits is to understand the influence of natural selection and other evolutionary forces on how organisms allocate their resources to reproduction and survival (Redneck 2010). Life history theory mainly revolves around the analysis of the evolution and the interaction of life history traits. A wide range of life-history traits can be analysed such as size at birth, rate of survival, lifespan, growth pattern, and reproduction rate, among others.

1.2.1 The trade-off among life-history traits

A crucial component that is essential to life-history theory is determining the trade-offs that occur between these traits (Hill et al. 1999). In an ideal world, the fitness of an organism will be at a maximum if the organism can maintain maximal survival and reproduction across all ages and stages. However, this will lead to a hypothetical organism known as a 'Darwinian demon', representing the absence of biological constraint on evolution (Law 1979). Without biological contrast on evolution, an organism will have maximum fitness, whereby they will be able to reproduce once they are born, reproduce an infinite number of offspring, and live indefinitely. In reality, we know that such an organism cannot exist, as there are limits to resources; additionally,

life history traits are affected by trade-offs and different constraints such that life history traits are not maximised by natural selection and thus limiting the fitness of an organisms (Stearns 1992).

In simplest case, trade-offs are often defined as a negative relationship between two or more traits (Stearns 1989). However, trade-offs can occur at two either genetic or physiological level. In genetic level, trade-offs are usually caused by allele with antagonistic pleiotropic effect. The negative phenotypic or genetic correlation between fitness component across individuals in a population is usually used to described trade-offs. For example, If the correlation is due to genetics, the negative genetic correlation will be predicted to limit the evolution of another traits. Therefore, genetic trade-offs is usually observed in a population when the evolutionary change that of a trait that increase fitness is linked to the evolutionary changes in another trait that decrease fitness. Genetics correlation can usually be quantified using genetic breeding designs. A classic example is the direct artificial selection for increase adult lifespan in a genetically variable laboratory population of fruit flies result in the evolution of increased adult lifespan (after around 10 generation) (Zwaan et al. 1995). However, the increase in lifespan is coupled with a decrease in early reproduction in fruit flies which suggest that the lifespan and early reproduction are negatively couple on genetic level (e.g. through antagonistic pleiotropic alleles).

On the other hand, trade-offs can also occur at physiological level which usually revolved around the concept of competitive allocation of a finite pool of energy into one life-history trait result in a decrease in energy allocation for another traits. However, it is important to bear in mind that physiological trade-offs might also involve a genetic basis since that different genotypes might varies in their resource allocation. Further, at population level trade-offs in physiological level does not always translate into genetic (evolutionary). For example, when the physiological trade off (within individual) is genetically fixed across in

the individual across the population, the two life history traits will showed the same negative physiological relationship whereby the genetic correlation among individuals would be zero (Stearns 1989)

A classic model for describing physiological trade-offs is known as the Y-model. The main principle of this model is that under circumstances where resources are finite, it is impossible to increase two different traits at a given time (King et al. 2011) (Figure 1.4).

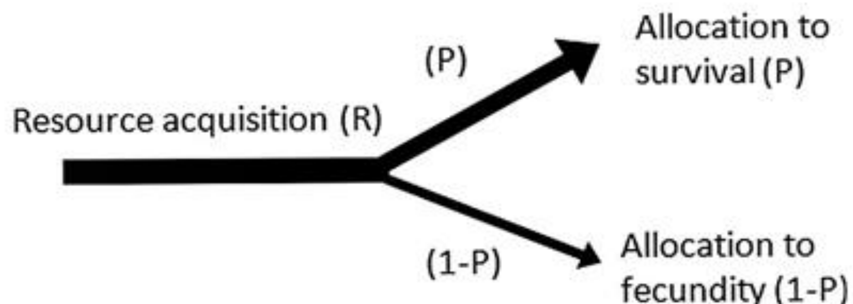


Figure 1.4. Y model in trade-off. R= Resource allocation, P= Allocation of resources for survival, (1-P) = resource allocate to fecundity. The width in arrows represents the proportion of resource allocated for each trait (thicker arrow= more resource allocation) (Adapted from Lailvaux and Husak 2014).

One of the most common examples for explaining the Y-model can be observed between somatic and reproductive efforts. An organism must choose to invest energy in either somatic (growth and maintenance) or reproductive effort (increase in the number of offspring). One example of this that can be observed in the wild are chimpanzees (*Pan*

troglodytes schweinfurthii), where the infant typically shows a long period of dependency on the mother. Research showed that when the mother chimpanzee invested more energy in reproduction efforts, the juvenile typically showed a decreased in growth (body size) (Thompson et al. 2016). Another classic trade-off often observed in reproduction can be seen between parenting and mating efforts (Dawson 1996; Ratikainen et al. 2018). In bi-parental care species that are socially monogamous, the male typically faces a trade-off between investing energy in parental care and seeking out extra-pair mating. According to the parental-mating effort, when the male is of higher quality (larger) he will have a greater chance at extra-pair mating; accordingly, they will reduce the time spent on parental care (Westneat et al. 1990). By analysing field data from the hair-crested drongos (*Dicrurus hottentottus*), Lei results supports the idea of trade-off in mating and parental care trade-off. During the incubation period, male drongos with longer tarsi exhibited increased efforts in seeking out extra-pair mating then caring for their young (Lei et al. 2020). However, these trade-offs only happened when nearby fertile female drongos were high in number. Trade-offs are often observed between predation risk and breeding investment. A classic example of such a trade-off can be observed involving the threespined stickleback (*Gasterostues aculeates*). When predators were introduced to the living environment of these animals, the males exhibited a decrease in breeding and nuptial colouration (Candolin 1998). However, these reproduction decisions were implemented when the future breeding probability was low, where the male will increase breeding despite the increase in predation risk.

1.2.2 Phenotypic plasticity, ecology and evolution

In recent years, an emerging field of study has been the interaction between phenotypic plasticity, ecology and evolution (eco-evolutionary dynamics). The interaction between ecology and evolutionary changes can occur in two different directions. First, evolutionary changes resulting from ecological changes can occur across variable time frames, ranging

from a small number of generations to hundreds of years. On the other hand, evolutionary changes can have a significant impact on ecosystems (Fussmann et al. 2007; Kinnison and Hairston 2007). A fundamental principle of eco-evolutionary dynamics is the interaction between an organism's phenotype and the environment. Fundamental questions exist in this context that must be addressed to understand the relationship between ecological evolution and phenotypic plasticity.

a) The adaptive nature of phenotypic plasticity

The adaptive nature of phenotype plasticity will first be discussed. Organisms can react to environmental stimuli by showing a variation in phenotype; however, not all plasticity characteristics are adaptive (Grether 2005). It is thus essential to quantify the adaptive significance of plasticity. One way in which adaptive significance can be quantified is by inducing plasticity followed by monitoring the fitness changes in organisms. One example of this is the thickness and size of a snail (*Physella acute*) shell. Auld and Relyea 2011 induced the snail through the presence/absence of non-lethal predatory crayfish. Next, the phenotypes (both alone and combined) were exposed to selection through introducing lethal crayfish followed by quantifying the linear and non-linear selection differentials. The introduction of non-lethal crayfish results in an increase in the mall and shell thickness of the snails. Further, the predation rate (number of snails killed) of predator-induced snail were much lower compared to the uninduced snails suggesting that induced snail has higher survival rate compared to those that were uninduced. This study suggest that predator induced changes of shell thickness could be adaptive. Although it is important to consider the adaptive value of plasticity in one environment, to determine if the plasticity is adaptive overall, the adaptive value across multiple environments should be measured.

b) The limitations and cost of plasticity

Second, the cost and limitations of plasticity must be addressed. Plasticity has both proximate and ultimate limits. The latter refers that some phenotypic changes can never be achieved through plasticity and selected phenotypic changes will thus never evolve. In the cases of proximate limits, the plasticity that can at any current point be observed within a population will not be sufficient for full adaptation. A good example of proximity limitations in plasticity can be observed during breeding time among the migratory Icelandic black-tailed godwit population (*Limosa limosa*) (Both and Visser 2001). Black-tailed godwits can change their breeding period in response to an abundance of caterpillars; however, the birds cannot breed until they reach the new place; accordingly, this migratory timing serves as a limitation on the benefit achieved through plasticity during the breeding time (Gill et al. 2014). Further, The cost and limitations of plasticity should be context-dependent, whereby the cost might be strong only in stressful environment and the plastic response are large.. (Figure 1.5). An example of relationship between cost of plasticity and degree of plasticity is observed in the island population of common frog (*Rana temporaria*) which varies in development time and phenotypic plasticity (Lind and Johansson 2009). Lind found that cost was only observed in frogs with high plasticity which suggest that the cost of plasticity might only present in the highly plastic population. Next, an example of relationship between cost of plasticity and stressful condition is investigated in the meta-analysis by van Buskirk and Steiner (2009). In animals the cost of plasticity is only observed when the environmental conditions are stressful (Figure 1.5). However, it is important to bear in mind that these observations is not universal and might varies depending on the organisms.

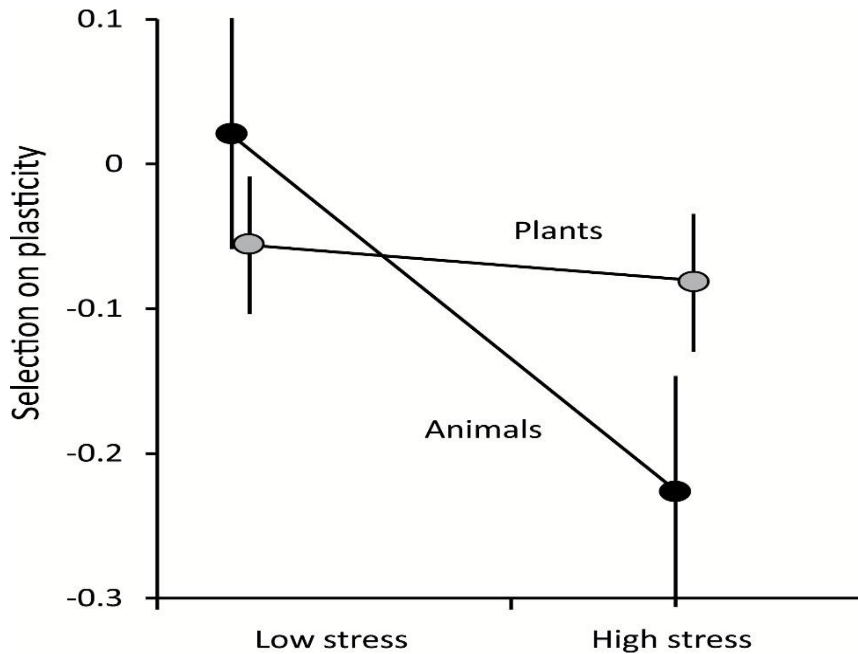


Figure 1.5. Distribution and pattern of selection on plasticity in animals and plants. Negative value= cost on plasticity, positive value= benefit of plasticity. In high stress, cost of plasticity is higher in animals, while in low stress plasticity is more beneficial in animals. In contrast, cost and benefits of plasticity in plants are not affected by stress levels which suggest that the cost of plasticity varies between organisms, traits and environment (Adapted from Hendry 2017).

c) Type of environmental conditions and organism characteristics that aid the evolution of plasticity

The third aspects involve determining the type of environmental conditions and organisms that are favourable for enhancing plasticity to evolve. As discussed above, the costs and limits differ across traits, environments, and organisms. Accordingly, we predict that the evolution of plasticity will also vary in the same manner. Using a theoretical model, researchers showed that the evolution of plasticity and selective conditions were closely related (Stomp et al. 2008; Ezard et al. 2014). The model also showed consistency in a few environmental aspects that favour evolution of plasticity. First aspect is the spatial variation whereby plasticity are more likely to evolve with higher spatial heterogeneity. Second aspect is the dispersal rate whereby population with higher dispersal favours evolution of plasticity as they might experience bigger spatial variation.

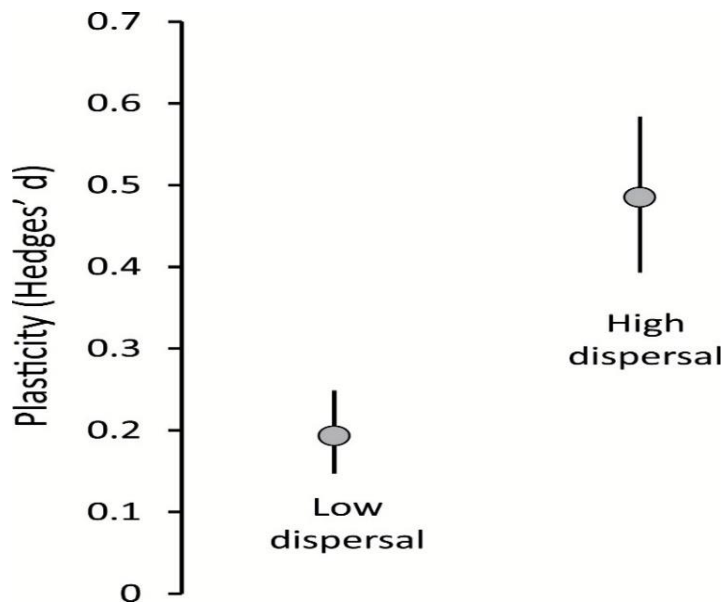


Figure 1.6. Plasticity in marine organisms with different dispersal rate. Marine organisms with high dispersal rate have higher rate of plasticity in comparison to marine organisms with lower dispersal rate. (Hollander 2008). Y-axis plasticity= magnitude of plasticity measured by Hedges'd whereby Hedges'd is a commonly used to measure the effect size (Buck et al 2022).

For example, research on marine invertebrates shows that species with a higher dispersal rate has higher plasticity compared with those with a low dispersal rate (Hollander 2008) (Figure 1.6). The third aspect is temporal variation whereby greater temporal variation are more favourable for evolution of plasticity. Lastly, is the reliability of environmental cues, where a plasticity response is more readily evolved if such cues are deemed reliable (Reed et al. 2010).

d) Plasticity, colonisation and response towards environmental changes

The fourth question that must be addressed regarding plasticity and evolution is the role of plasticity in aiding colonisation and the extent of plasticity response to environmental changes. In the case of environmental changes, large environmental shifts are typically unfavourable because organisms with an existing phenotype will not be well-adapted to the new conditions. In these cases, it is predicted that organisms will change their phenotype

to adapt to the new environmental conditions, which may result in the persistence or extirpation of the species. This type of phenotypic rescue can be achieved through adaptive genetic changes (evolutionary rescue) or adaptive plasticity (plastic rescue) (Chevin et al. 2010; Barrett and Hendry 2012; Kovach-Orr and Fussmann 2013). The principle of plastic rescue is that the organism will change its phenotype in response to environmental changes, which will subsequently give rise to an increase in mean population fitness and ultimately lead to survival and colonisation in the new environment (the Baldwin effect) (Baldwin 1902; Crispo 2007). Next, the Baldwin effect predicts that genetic changes will slowly follow the direction of plastic response, which will then lead to 'genetic accommodation' (West-Eberhart 2003; Schlichting and Wund 2014) (Figure 1.7).

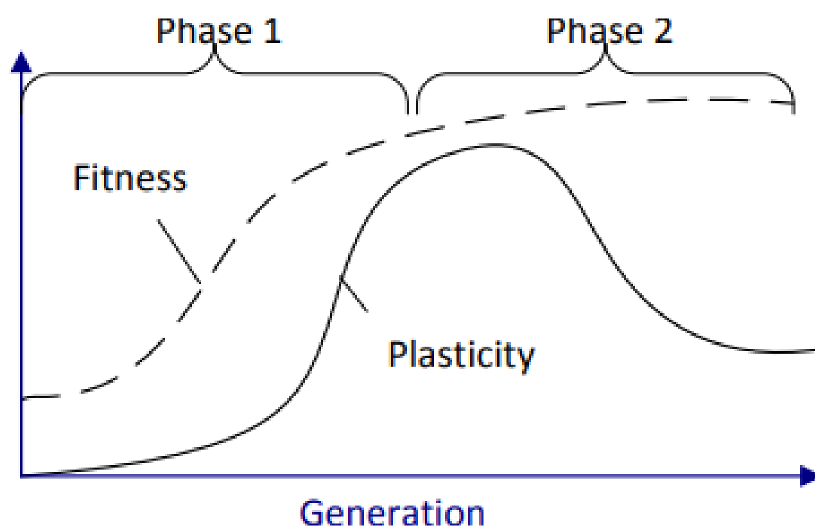


Figure 1.7. Baldwin effect graph. Phase 1 is plastic rescue where the fitness increases as plasticity is selected for when organism exposed to different environment. Phase 2 Genetic accommodation where the organism adapted to the environment changes and the overall plasticity decreases. (Ellefsen 2013).

There is evidence about plants that suggest the importance of plasticity in responding to environmental changes. An example of this is changes in flowering time and the number of plant species in 'Thoreau Woods' in Concord, Massachusetts spanning 150 years. The species that were extirpated in the area had something in common, i.e. a low level of

plasticity during the flowering time in response to changes in temperature while the surviving species showed an increase in flowering time by seven days on average (Willis et al. 2008). In contrast, Both et al. (2006) reported a huge decline (90%) across the nine dutch population of the pied flycatchers (Ficedula hypoleuca) due to insufficient change in plasticity where the food for provisioning nestling peak earlier in the season but the birds were currently mistimed. Together the results suggest that sometimes plasticity is sufficient for phenotypic rescue but in other cases evolutionary changes will be needed.

e) Plasticity and genetic evolution

The fifth question involves determining if plasticity aids or constrains genetic evolution. Some predict that plasticity helps to promote evolution while an opposing view suggests that plasticity constrains genetic evolution. The primary notion of plasticity in promoting genetic evolution is that it aids in the survival and persistence of a species in a new environment, which can increase the selection of the traits (West-Eberhard 2003). A good example of how plasticity changes in one trait leads to a genetic change in other traits is seen in curly-tailed lizards (*Leiocephalus carinatus*). In response to predation, these lizards plastically change their habitat to narrow perches in trees, the change in habitat resulted in the selection of shorter legs in the curly-tailed lizards (Losos et al. 2006).

Contrastingly, the constraint side suggests that plasticity shields the genotype from selection, which results in slow adaptive genetic changes (Linhart and Grant 1996; Ghalambor et al. 2007). An example on plastic changes restrain genetic changes is observed in the cichlid fish *Pseudocrenilabrus multicolor victoriae* (Crispo and Chapman 2010). By collecting cichlid fish from different oxygen environment and raising their offspring under two different oxygen condition (low and high), Crispo found that the variation in the gill size was due to plasticity with no genetic difference observed across population. In the other hand the brain size difference observed in these population has

high genetic variation across population but less plastic. The results from Crispo and Chapman supports the arguments that plasticity constraint genetic evolution with more plastic trait showed lower genetic divergence.

f) Plasticity and ecological speciation

The sixth question pertains to determining the role of plasticity in ecological speciation. There are two opposing perspective on this topic. The first perspective is that plasticity aids ecological speciation. The principle supporting this notion is that plasticity allows the use of new resources and colonization of new environments whereby then natural selection can act on the phenotype resulting in adaptive genetic divergence and ultimately lead to ecological speciation (Skúlason and Smith 1995, Weest-Eberhard 2003, Pfennig et al. 2010) (Figure 1.8).

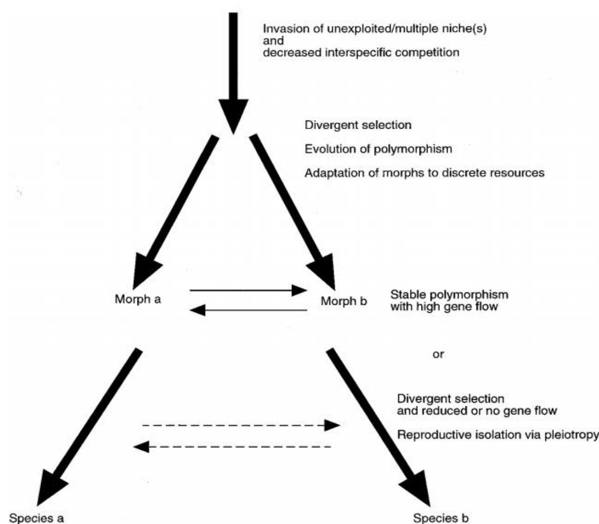


Figure 1.8. Schematic flow for ecological speciation. First the invasion of an unexploited niche where organisms are adapted to new discrete resources, which leads to divergent selection. There are two scenarios of divergent selection (with high gene glow) where the organism shows two different morphs. Second, divergent selection with low gene flow and reproduction isolation via pleiotropy (where a single locus can affect two or more unrelated phenotypic traits). This pleiotropic mutations can then add genetic variation and help establish positive genetic correlations which is requires for speciation. (Adapted from Smith and Skúlason 1996).

A way that plasticity promote ecological speciation is through polyphenism whereby same genotype give rise to different phenotypes under environmental stress which can result in sharpening of reproductive barrier and lead to ecology speciation. An example of this can be observed between the two species of spade toads (*Spea bombifrons* and *Spea multiplicata*), where plasticity in feeding behaviour (being either herbivorous or carnivorous) can be observed in both species. Under the common-garden environment, Pfennig and Murphy (2002) S.multiplicata produce less carnivores when the population of S.bombifrons is high suggesting that the selection has enhanced ancestral polyphenism which then exaggerated species divergence.

Contrasting arguments propose that plasticity prevents ecological speciation. A supporting argument for this idea, as previously discussed, is where plasticity prevents natural selection on targeting on genotype, which results in lower genetic divergence, hence slowing speciation (Svanbäck et al. 2009). Evidence suggesting that plasticity prevents ecological speciation is scarce. An example plasticity constrain speciation is observed in the Misty Lake and Outlet stickleback population which shows high level of plastic differences but only minor reproductive isolation (Roesti et al. 2012). However further rigorous testing is needed to support the idea that plasticity hampers speciation.

g) The rate of plasticity evolution

A seventh aspect concerns plasticity evolution is the need to determine the rate of evolution of plasticity and the factors that influence this rate. An example of study in the rate of plasticity in carried out in the marine mussels *(Mytilus edulis*). During 1988, the Asian shore crab, *Hemigrapus sanguineus*, invaded northern New England and fed on these mussels, which lead to the mussels thickening their shell whereas the southern population has never encounter the Asian shore crab (Freeman and Byers 2006). Freeman and Byers 2006 showed that the southern population induce shell thickening when exposed to

waterborne cues from Hemigrapsus whereby the southern population which previously has never encounter the Asian shore crab showed no response to the waterborne cues. The result suggests a rapid evolution of shell thickening response towards Hemigrapsus within 15 years of introduction to the northern population.

Despite significant supporting evidence, some research also claims that plasticity has not evolved, even after long periods. An example is whereby some organisms retain anti-predator behavior even when predator were no longer present. This observation could possibly due to relaxed selection (where the trait is not expressed in the absence of cues), where the evolution of the trait will take a longer time to develop through drift and mutation (Lahti et al. 2009).

h) The effect of plasticity on communities and ecosystems

Finally, how plasticity impacts evolution at community and ecosystem level must be considered. Since plasticity can change the life-history traits of a species, which will affect their interaction with other organisms in an ecosystem, we predict that plasticity will have an impact at both the community and ecosystem level. However, empirical studies on the effect of plasticity at these levels are limited. As noted above, many fish species exhibit plasticity in terms of foraging traits (jaw size, gill length) in response to changes in diet. For example, fish that adopt a zooplankton diet increase their foraging efficiency by increasing the length of their gill rakers and change their jaw morphology compared to fish with a benthic diet (Day and McPhail 1996). Since plasticity in fish can increase their foraging efficiency and influence aquatic prey communities, we propose that this plasticity will also impose an effect on their prey communities (Lundsgaard-Hansen et al. 2014). That's because improve efficiency in obtaining a given-food type should reduce the availability of the food type (which could induce its evolution) and thereby influencing further plasticity and selection.

1.3 The history of epigenetics

Epigenetics is a term commonly used in the study of development and evolution. The term epigenetics was first used by British embryologist Conrad Waddington (1942) in the 1940s. The definition of epigenetics has progressively changed since the term was first introduced. Epigenetic was first used to describe the interaction between genes that results in phenotypes (Waddington 1968). However, as science has progressed and our understanding of biological processes and development has deepened, the term epigenetics has been redefined. The most common definition of epigenetics in the 20th century was differences in phenotype that are not caused by changes in the DNA sequence (Wu and Morris 2001). Although the term epigenetics was coined in the 1940s, the involvement of epigenetic in regulating gene expression was not described until 1969, when Griffith and Mahler (1969) found that epigenetics could impact long-term memory in mice.

One of the most common models of epigenetics explains that they regulate gene expression through methylation of the cytosine base in DNA (Riggs 1975; Holliday and Pugh 1975). Pugh described how the cytosine base is affected by DNA methylation, which helps regulate gene expression during development and results in phenotype differences. The main element in Pugh's model is specialised enzymes involved in DNA methylation. There is a lot of evidence, especially from Bird, that supports Pugh's model. For example, because DNA methylation usually only targets the sequence CpG, Bird have taken this advantage and introduced the idea that methylation-sensitive enzymes are responsible for detecting the DNA methylation state (Doskocil and Sorm 1962; Bird 1978). Other studies support Bird's proposal and find that endogenous CpG is usually found in two states: completely methylated or completely unmethylated (Bird 1978). Since the discovery of DNA methylation, there has been great interest in understanding the extent that DNA methylation could be preserved through germlines and found that the ability of DNA methylation to be

preserved in the germline is dependent on the CpG sites that survive demethylation events (Felsenfeld 2014). There are many studies that show the importance of epigenetic mechanisms in regulating gene expression. For example, the Igf2/H19 gene is expressed due to methylation in one allele (Kanduri et al. 2000; Hark et al. 2000). Other evidence that DNA methylation is involved in regulating gene expression is seen in the nucleoside analogue 5-azacytidine. The nucleoside 5-azacytidine is incorporated into DNA, which could covalently bind to DNMT and thereby reducing is biological activity and even deactivating DNA methyl transferase and resulting in DNA demethylation (Holliday 2006). By using 5-azacytidine to inhibit DNA methylation, we can observed if there is any changes in the methylation level of and also gene expression of relevant genes that is involved in the phenotypic changes.

There are many differences between traditional Mendelian genetics and epigenetics. Traditional genetics usually revolves around the concepts of cell lineage and clonal inheritance. During gametogenesis, a germ cell will usually undergo meiosis that resulting in haploid cell that will develop into gametes. Then the two haploid cells will fuse together to form a diploid zygote. This suggests that organisms usually start from a single cell, which is cloned, so if the somatic cell carries a mutation or chromosomal change, its descendants will carry the mutation as well. In contrast, epigenetics usually occurs in a specific group of cells (Holliday 2006). For example, induction of the muscle tissue is usually found in the mesoderm cells, because the mesoderm cells contain a receptor that reacts to the specific signals. Another difference is that, in Mendelian genetics, genetic mutations are usually stable and non-reversible. In contrast, epigenetic processes such as genomic imprinting are often reversible. In genomic imprinting, during gametogenesis, DNA changes can be either lost or erased and reset (Paoloni-Giacobino and Chaillet 2004). Additionally, genotypes are usually not affected by environmental conditions aside from mutagens. In epigenetic changes, the environment plays an essential role in determining an organism’s

development, which in this sense encompasses Lamarckian inheritance. Lastly, the difference between epigenetic and genetic is that epigenetic does not involved an alteration on the DNA sequence while genetic usually involves mutation that alters the DNA sequence (Figure 1.9).

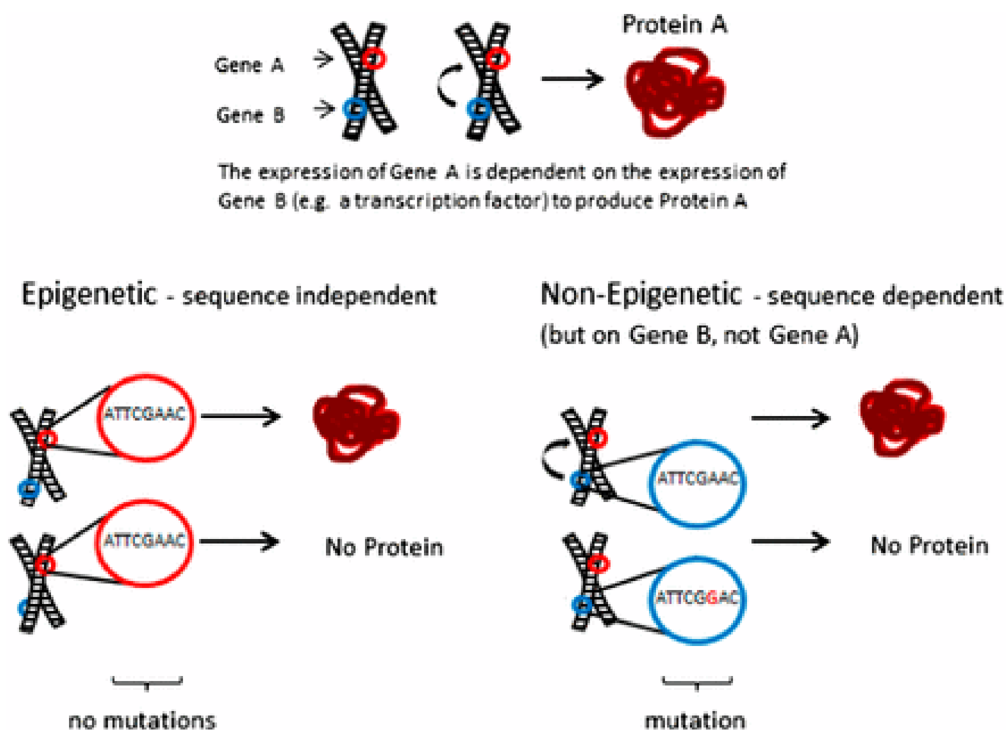


Figure 1.9. Difference between Mendelian genetics and epigenetics. Protein A production is dependent on the expression of gene A and B. When there are no genetic or epigenetic changes in gene A, protein A is produced successfully. When epigenetic changes occur in gene A, the sequence is not altered but protein A is not produced. Protein A is not produced when there are genetic changes (mutation) which alters the sequence of gene B (Adapted from Deans and Maggert 2015).

The difference between genetics and epigenetics led to the introduction of the term dual inheritance by scientist Maynard Smith (1990). The term dual inheritance refers to two types of inheritance: genetic inheritance that involves changes to the DNA sequence,

and epigenetic inheritance that does not modify the DNA sequence. Smith's introduction of this term is based on Jablonka and Lamb's (1995; 1989) proposal that phenotypic changes caused by epigenetic changes due to changes in environmental stimuli could be heritable. There is a lot of evidence supporting Jablonka and Lamb's proposal that many transgenerational effects could be explained by epigenetic modification (Dubrova 2000). There are a few mutations in plants from epigenetic rather than genetic changes. A common example is the *Linaria vulgaris*, which changes between bilateral and radial floral symmetry based on the methylation of the Lcyc gene (Cubas et al. 1999).

1.3.1 Mechanisms involved in epigenetics

There are many mechanisms involved in epigenetics. DNA methylation is the most studied mechanism. The main principle of DNA methylation involves the covalent transfer of a methyl group to the C-5 position of the cytosine ring with the help of specific family of enzymes know as DNA methyltransferases (Dnmt) (Roberston, 2005) (Figure 1.10).

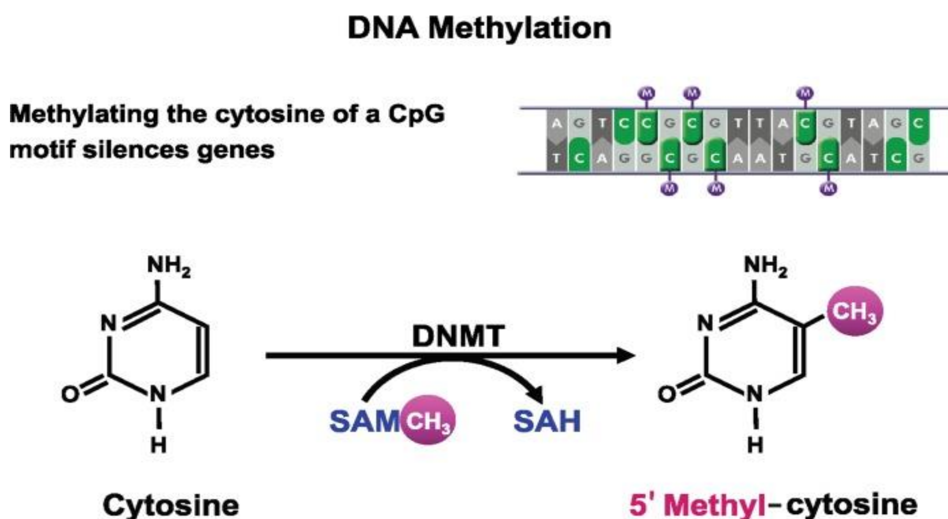


Figure 1.10. Summary of the mechanism of DNA methylation. SAM=S-adenosylmethionine, SAH=S-adenosylhomocysteine (adapted from Zakhari 2013). DNA methyltransferase (DNMT) is a

family of enzymes that are involved in methylation. Specifically DNMT3a/3b transfer the methyl group from S-adenosyl methionine to the C-5 position of the pyrimidine ring.

CpG methylation accounts for almost 98% of methylation. Other methylation includes non-CpG methylation in embryonic stem cells (Lister et al. 2009). The Dnmt family includes enzymes such as Dnmt1, Dnmt2, Dnmt3A, Dnmt3B and Dnmt3L. Each DnmtT has specific role and function (Jeltsch and Gowher 2019; Okano et al. 1999). The presence of Dnmt enzymes varies between species. For example, the honeybee (*Apis mellifera*) has two copies of the DNMT1 enzyme, while fruit flies (*Drosophila melanogaster*) lost Dnmt1 and Dnmt3 during evolution, and the round worm (*Caenorhabditis elegans*) has lost the Dnmt enzyme family entirely (Li et al. 2018). During synthesis phase (S phase), the maintenance methyltransferase Dnmt1 is usually localised in the replication foci (the site where newly synthesised DNA usually accumulates) and shows a tendency to methylate hemimethylated DNA in vitro. This tendency has led to the proposal that the maintenance methyltransferase Dnmt1 is essential for copying DNA methylation patterns to the daughter strands during DNA replication (Probst et al. 2009). The necessity of Dnmt1 in organism development and survival is shown in the mouse model (Li 2002). Li found that mice with both Dnmt1 alleles deleted are lethal at embryonic stage 9. In contrast to Dnmt1, Dnmt2, Dnmt3A and Dnmt3B are classified as de novo methyltransferases. Instead of methylating DNA, Dnmt2 shows a preference for methylating Cytosine 38, located in the anticodon loop of aspartic acid tRNA (Goll et al., 2006), and Dnmt3A and Dnmt3B have a preference for methylating unmethylated CpG during development.

The importance of de novo methyltransferases Dnmt3A and Dnmt3B is shown by knockout mice (Okano et al. 1999). Okona found that mice could survive up to 4 weeks without Dnmt3A, while mice without Dnmt3B were lethal between the embryonic age of 14.5 and 18.5 weeks. Dnmt1, Dnmt3A and Dnmt3B were also found to be involved in controlling telomere integrity. Gonzalo et al. (2006) found that mice with Dnmt1, Dnmt3A and

Dnmt3B knockout had elongated telomeres and increased telomeric recombination, which suggests the importance of Dnmts in maintaining genome stability. Dnmt3L is homologous to Dnmt3A and Dnmt3B and possesses no catalytic activity. Dnmt3L plays an important role in increasing the binding affinity of Dnmt3A and Dnmt3B to S-adenosyl-L-methionine (SAM) and increasing its activity in vitro (Kareta et al. 2006). Bourc'his et al. (2001) found that mice with homozygous DNMT3L survived; however, the heterozygous embryos produced from the homozygous DNMT3L mom oocyte were found to be lethal at embryonic stage 9 with severe impairments in maternal methylation. All this evidence suggests that DNMTs are essential to organism development and survival and each DNMT enzymes has different role (Figure 1.11).

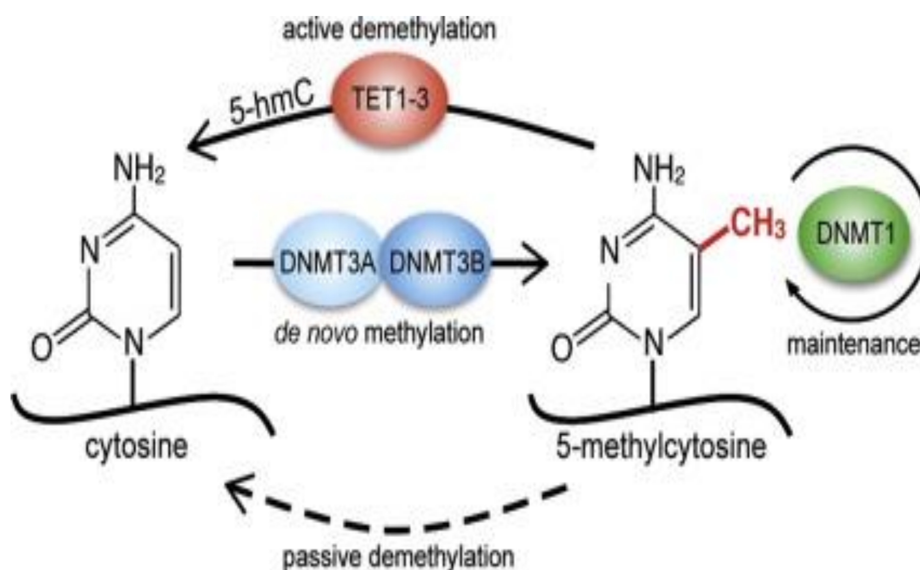


Figure 1.11. Function of the DNMT family. DNMT=DNA methyltransferase, TET=Ten-eleven translocation methylcytosine dioxygenase. DNMT1 involves in maintaining the methylation mark while DNMT3A/3B involves in setting up the methylation pattern (denovo methylation) and TET1-3 is involved in the process of demethylation (Adapted from Ambrosi et al., 2017).

1.3.2 Histone modifications

In eukaryotes, there is a complex known as chromatin comprised of DNA and proteins (Mondal et al. 2010). Chromatins can usually be divided into two categories: heterochromatin and euchromatin. Heterochromatins are usually densely packed and

contain repetitive sequences, low numbers of gene sequences and are mostly inactive for transcription (Tamaru 2010). In contrast, euchromatins have a looser and less dense chromatin structure, contain more genes and are mostly transcriptionally active. The basic unit of chromatins is the nucleosome, a structure containing 147base pair (bp) of DNA that usually wraps around the histone octamer. There are four core histones: H2A, H2B, H3 and H4 (Peterson and Laniel 2004) (Figure 1.12).

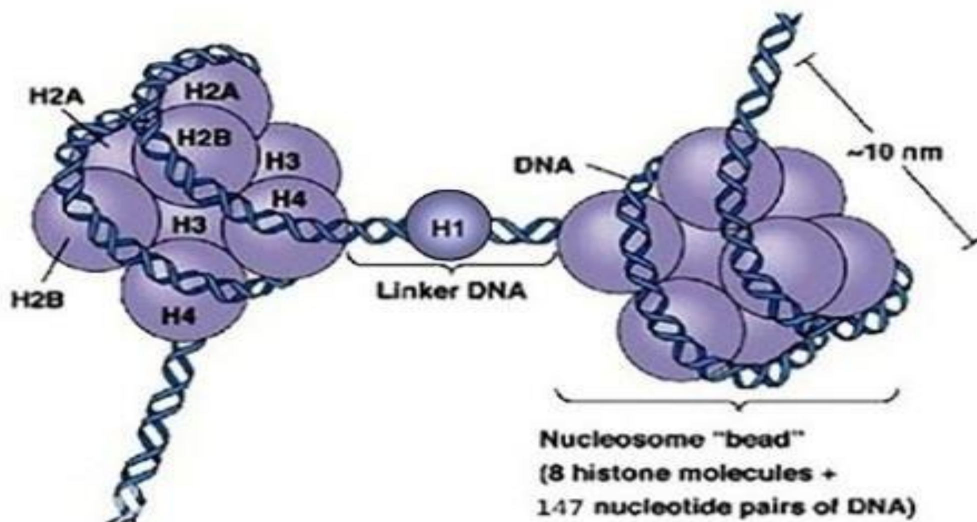


Figure 1.12. Nucleosome structure. Nucleosome bead structure that are formed by histone octamer wrapped around by 147bp of DNA. (Adapted from Lu et al. 2017).

Histone modification, usually found in the core histone tail, can affect cellular process such as gene transcription, DNA replication and DNA repair. There are many types of histone modification, including acetylation, methylation, phosphorylation and ubiquitination (Santos-Rosa and Caldas 2005). Lysine (K) acetylation affects the chromatin structure and accessibility of protein which ultimately affects gene transcription. Lysine methylation could have different effects depending on two factors: the amount of lysine methylated and the number of methyl groups (Kim et al. 2009). Examples of methylation

include trimethylation of H3K9 and H3K27, usually found in inactive genes, and trimethylation of lysine 4 H3K4, involved in active transcription (Pray-Grant et al. 2005). Since the accessibility of euchromatin for protein is essential for gene transcription, high levels of lysine trimethylation, such as H3K4, are usually found in genes that are transcriptionally active (Sims et al. 2003).

Arginine methylation can either activate or repress transcription. The coactivator-associated arginine methyltransferase CARM1/PRMT4 is essential for regulating NF-kappaB. For example, the expression of a subset of NF-kappaB-dependent genes is affected when CARM1 is knocked out of cells (Covic et al. 2005). Another example is how the forced expression of arginine methyltransferase PRMT5 leads to a reduction in the promoter activity of the E1 cyclin (Fabbrizo et al. 2002). Histone lysine methyltransferase G9A shows a preference for methylating histone H3 lysine 9 (H3-K9). Mice studies have shown that embryonic cells that lack H3-K9 methylation exhibit severe growth impairment and early fatality (Tachibana et al. 2002). Tachibana found that embryonic cells that lack the methyltransferase G9A also show a reduction in methylation of H3-K9, which suggests that G9A is important in regulating H3-K9 methylation. The importance of G9A in transcriptional regulation is also supported by research. Yuan et al. (2007) found that knocking out G9A caused the H3-K9 level to decrease, leading to a reduction in pre-rRNA synthesis.

In addition to regulating gene transcription, histone modification can affect cellular process such as DNA damage response. For example, the phosphorylation of H2A.X plays an essential role in marking DNA damage sites. When DNA double strands break, H2A.X responds by providing a signal for the end joining repair pathway (Chowdhury et al. 2005). Methylation of histone 4 lysine 20 (H4K20) plays an important role in healing DNA damage caused by ionising radiation. Botuyan found that methylation of H4K20 interacts with the DNA repair factor 53BP1 and Crb2 to form the 53BP1/H4-

K20me2 complex at the DNA damage site (Botuyan et al. 2006). The above clearly shows that DNA methylation and histone modification play important roles in organism development through affecting several aspects of cellular processes such as transcription, DNA repair and replication (Figure 1.13). It is also clear that the epigenetic mechanisms of DNA methylation and histone modification are closely linked to recruiting specific protein factors into chromatin.

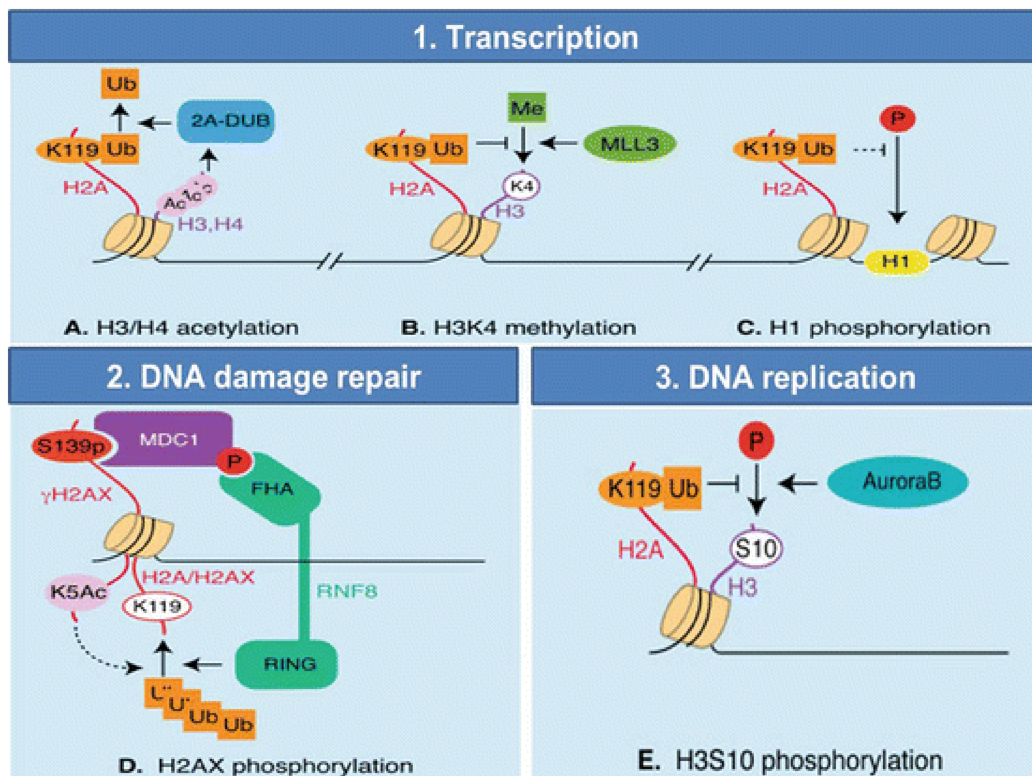


Figure 1.13. Summary of cellular processes affected by histone modification. Ub=mono-ubiqutination, Ac=acetylation, Me=methylation, P=phosphorylation. (A-C) impact of histone modification on transcription, (d) DNA damage repair and (e) DNA replication. (Adapted from Vissers et al., 2008).

1.3.3 Methylation differences between vertebrates and invertebrates and within invertebrates

Although methylation is found in both vertebrates and invertebrates, there are huge differences in methylation between them. One of the biggest differences is the percentage of methylated CpG. In mammals, roughly 60%–90% of CpG is targeted by DNA methylation

(Lister et al. 2009). In comparison, DNA methylation is much lower in invertebrates such as fruit flies (*Drosophila melanogaster*) and round worms (*Caenorhabditis elegant*).

The methylation target region is also different in invertebrates and vertebrates. In vertebrates, most methylation occurs in the promoter region and is associated with gene repression (Cedar and Bergman 2009; Mandrioli 2004). In most invertebrates, such as insects, methylation occurs in the gene body (intron and exon) region and is generally associated with gene activation. Another difference between methylation of invertebrates and vertebrates is that invertebrates have higher levels of methylation in repetitive transposable elements than vertebrates (Schaefer and Lyko 2010; Zemach et al. 2010). The methylation distribution pattern also differs between vertebrates and invertebrates. Methylation distribution can be determined through CpG content analysis. Using this technique, Okamura et al. (2010) showed that insects such as pea aphids (*Acyrthosiphon pisum*) and honeybees (*Apis mellifera*) exhibit a bimodal distribution pattern, while invertebrates exhibit a global methylation pattern (Walsh et al., 2010) (Figure 1.14).

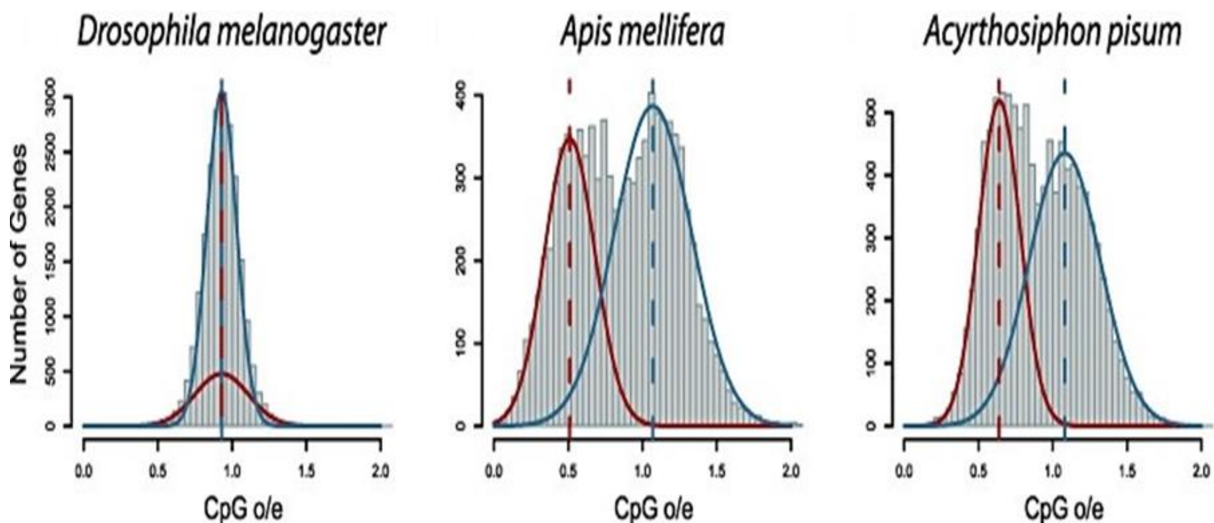


Figure 1.14. Distribution of the methylation pattern of genes across insects. Dashed line= mean of each component (adapted from Glastad et al., 2011). *Drosophila melanogaster* (lacks DNA methylation) shows a unimodal normalized CpG content (CpG o/e) distribution. On the other hand, genes in the honeybee (*Apis mellifera*) and pea aphid (*Acyrthosiphon pisum*) which have DNA

methylation show a bimodal distribution pattern. The distribution with a lower mean represents genes with lower levels of methylation.

The bimodal methylation pattern shows that genes can be classified in two categories: genes with low methylation (high mean normalised CpG content) and genes with high methylation (low mean normalised CpG content; Elango et al. 2009). In addition to showing the methylation distribution pattern, the normalised CpG analysis technique can be used to infer the presence of DNA methylation.

In addition to the methylation differences between invertebrates and vertebrates, there are many differences in methylation patterns between invertebrate species. For example, the presence of DNA methylation Dnmt varies between invertebrates. Research using phylogenetic analysis has shown that Dnmt1 are lost in some insect orders, such as Collembola, Strepsipteran and Diptera (Provataris et al. 2018), and these events are very rare (Figure 1.15).

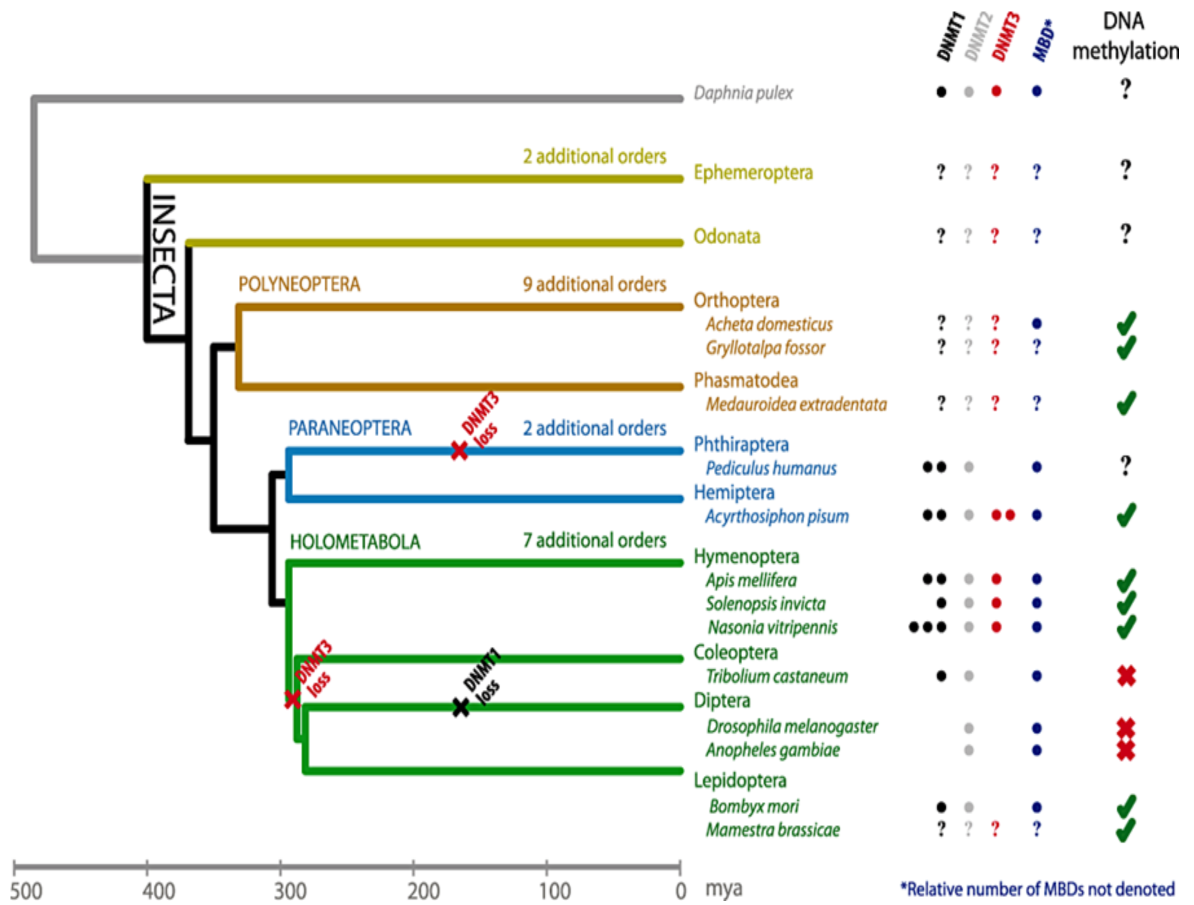


Figure 1.15. DNMT family distribution pattern in different species of insects. DNMT = DNA methyltransferase, ?=not tested/still unknown; MBD=methyl-binding domain, MYA=million years ago the branches show difference insect order with species representing each order. Loss of specific DNMT is indicated on the specific insect order branches. (Adapted from Glastad et al., 2011).

Dnmt3 is also absent in some insect groups, such as Neuropterida, Palaeoptera and Mecopterida. However, the absence of Dnmt3 could be due to it being very lowly expressed. For example, Dnmt3 is not detected in the brown plant hopper (*Nilaparvata lumens*), however, Dnmt3 is lowly expressed across the plant hopper life stages (Zhang et al. 2015). DNA methylation levels vary depending on the insect order (Figure 1.16).

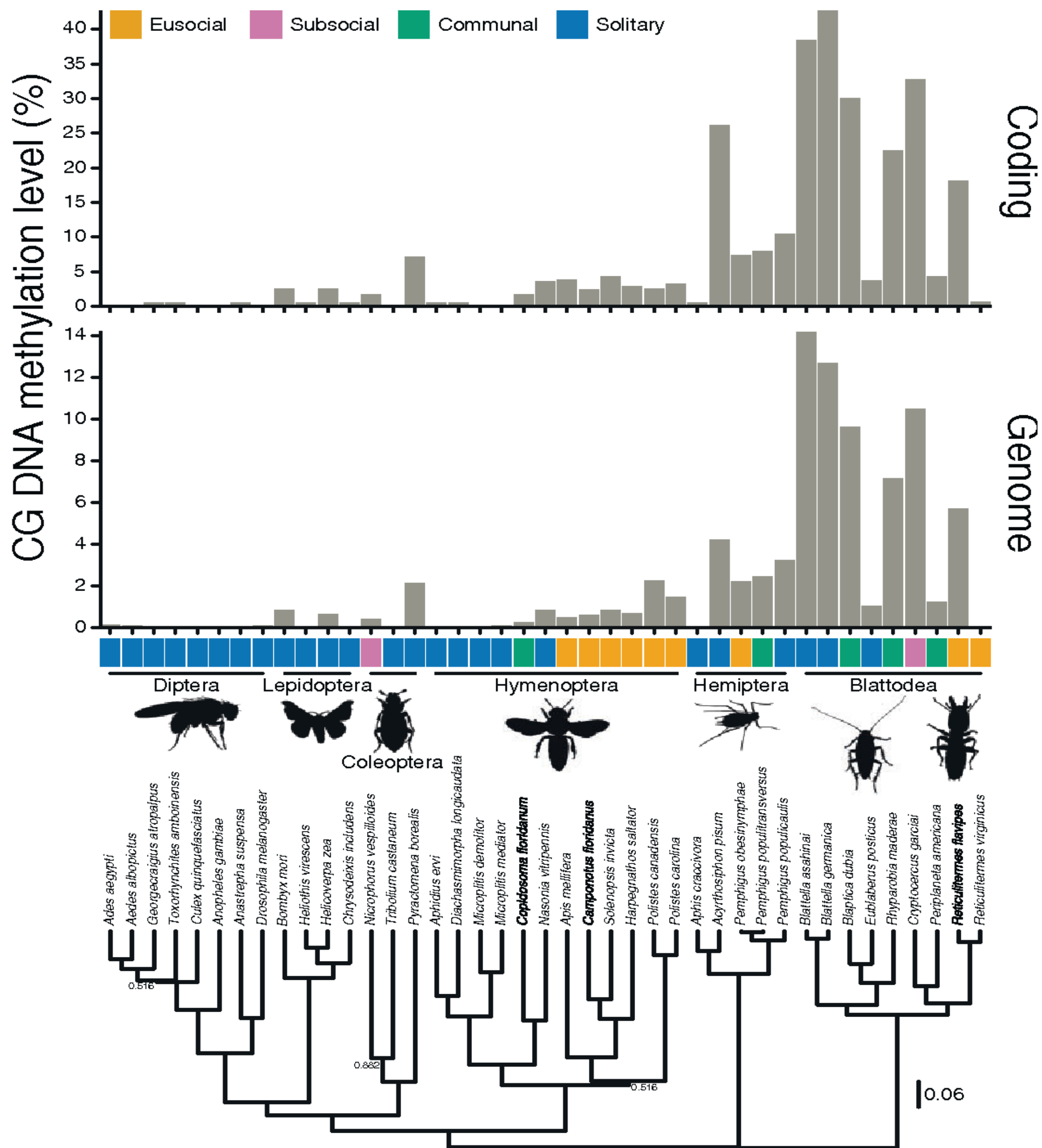


Figure 1.16. DNA methylation level across insect order. Blue=solitary insects, Green=Communal insects, Pink= Subsocial insects, Orange= Eusocial insects. DNA methylation in coding exon insects order ranges from 0% (Diptera), 25%(Hemiptera) and 42% (Blattodea). The genome methylation of insect ranges from 0% (Diptera), 5%(Hemiptera) and 14% (Blattodea). (Adapted from Bewick et al., 2017).

Hymenopteran insects usually have lower DNA methylation levels than insects from orders such as Hemiptera and Blattodea. Additionally, many species within Hymenoptera, such as parasitoid wasps (*Microplitis demolition, Aphids ervi*), show no or extremely low levels of

DNA methylation (Standage et al. 2016). In Blattodea, the asian cockroach (*Blattela asahinai*) has the highest level of methylation (Bewick et al. 2017), and in Hemiptera, the pea aphid (*Acrythosiphon pisum*) has the highest DNA methylation level (Bewick et al., 2017).

1.3.4 The role of epigenetics in insect development

Epigenetic regulation plays an important role in insect development and polyphenism. In social insects such as honeybees, epigenetics plays an essential role in queen and worker caste differentiation (Smith et al. 2008). Most honeybee larvae will develop into workers, while only a few larvae will be selected to develop into queens through environmental stimulus (Wheeler 1986). The role of epigenetic mechanisms in regulating caste differentiation is supported through the knockdown of the Dnmt3 gene in honeybee larvae, which leads to the production of queens (Kucharski et al., 2008). Epigenetics is also important in regulating insect reproduction. In a knockout Dnmt1 in the large milkweed bug (*Oncopeltus fasciatus*), most eggs produced were inviable and the number of eggs produced declined slowly (Bewick et al., 2019). However, the absence of Dnmt1 in milkweed bugs does not lead to any gene expression changes, which suggests that DNA methylation might have an important function in development independent of gene expression. The importance of epigenetics in development can also be seen in the parasitoid wasp (*Nasonia vitripennis*). When Dnmt1A is lowered in the embryo, most embryos show early lethality during the gastrulation stage (Zwier et al. 2012). In addition to their role in development, there is evidence showing that epigenetic mechanisms are important for ensuring insecticide resistance in green peach aphids (*Myzus precise*). Field (2000) found that pesticide-resistant green peach aphids showed high level of methylation on the insecticide-detoxifying esterase E4 gene.

1.3.5 Ecological epigenetics

Since epigenetics was introduced, it has received a lot of attention, as much that epigenetic changes are important to organisms' development. Recently, there has been an increased interest in the study of epigenetics in the context of ecology and evolution. This increase in interest is because that non-genetic phenotypic variation can be transgenerationally inherited and affect the speed of evolution (Bonduriansky 2012; Scoville et al. 2011). There are two pathways for epigenetics to affect the microevolution of a natural population (Figure 1.17).

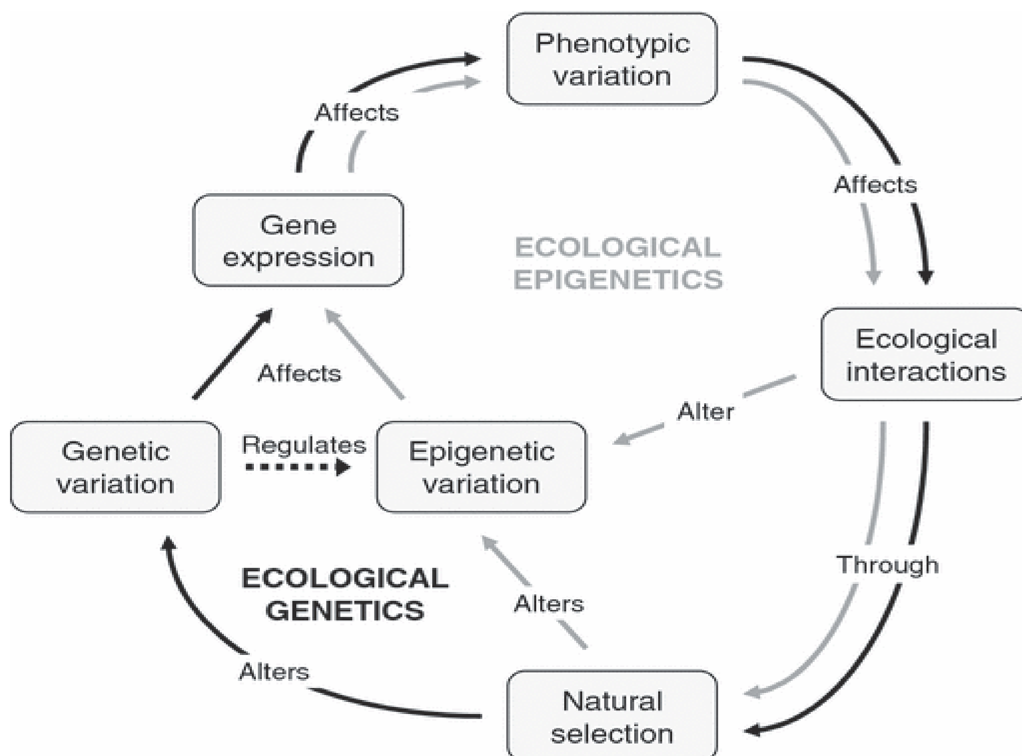


Figure 1.17. Pathway of genetics and epigenetics in ecological context. Pathway of ecological epigenetic (grey arrow) and ecological genetics (black arrow). Epigenetic variation can directly alter by ecological interactions and therefore provide alternative pathway for accelerated evolution. Genetic variation on the other hand only regulates epigenetic variation and evolution only through natural selection. (Adapted from Bossdorf et al., 2008).

Epigenetics could cause a change in gene expression that leads to phenotypic variation and affects the fitness of the individuals (Bossdorf et al. 2008). If epigenetic changes are inherited transgenerationally, they could be targeted by natural selection. The role of

epigenetic inheritance is discussed by Jablonka (2013). However, there are a few fundamental questions in the study of ecological epigenetics, namely the percentage of marked alleles that are transferred to the next generation, the reliability of epigenetic marks in response to environmental stimuli and the number of generations that the epigenetic mark can be inherited before being reset (Becker et al. 2011). Even though there is no empirical evidence of the evolutionary consequences of epigenetic effects, there are similarities in the mechanism of epigenetics and genetics and their evolutionary outcome. These similarities suggest that it is possible to derive some information from the link between epigenetics and evolution through studying genetic accommodation (Herman et al. 2014, Schlichting and Wund 2014).

There are some fundamental questions about ecological epigenetics. The first and most important factor to understand is the distribution pattern and structure of epigenetic variation both within and between natural populations (Richards et al. 2012). Understanding the epigenetic distribution in a natural population could provide insight into the extent of the transgenerational effect of epigenetic variation. The second question in the study of ecological epigenetic is the relationship between phenotypic and epigenetic variation in traits that are ecologically important (Bossdorf et al. 2008). The purpose of this question is to determine if phenotypic traits resulting from epigenetic changes have any effect on organisms' fitness and therefore effects in ecological and evolutionary contexts (Kiltivis et al. 2014). The third question in the study of ecological epigenetics is on the role of epigenetic variation in ecological interactions. It is important to understand the relationship between abiotic factors, such as pollution, salinity and other stimuli, on animals and plants (Cramer et al. 2011). The last question is on the heritability of epigenetic variation induced by biotic and abiotic factors. Transgenerational epigenetic inheritance is commonly known

as ‘soft inheritance’, in contrast to traditional Mendelian inheritance (Chong and Whitelaw 2004). As such, it is crucial to understand the extent to which evolutionary responses triggered by environmental stimuli are mediated by epigenetic inheritance.

In the study of ecological epigenetics, there are many problems faced by researchers. First, many of the ecological and evolutionary consequences found in laboratories cannot be generalised to natural conditions. This is because the fitness in an organism is environment-dependent and epigenetic variation could differ considerably depending on environmental conditions (Ledón-Rettig 2013). For example, in low-stress conditions, adult mice who received high maternal care as pups showed more hippocampal-dependent learning than mice that received low maternal care. However, in high-stress conditions, mice that had lower maternal care as pups showed faster learning (Champagne et al. 2008). Therefore, it is hard to determine if the phenotypes caused by maternal care are adaptive without knowing the environmental conditions the mice will face as adults. Second, in the real world multiple complex cues interact to create the phenotypic variation. However, when studying one environmental effect in laboratory settings, other environmental conditions are maintained to avoid influence (Pigliucci 2005). Therefore, there are many discrepancies between phenotypic variations in laboratory study and field study. In a real-world scenario, multiple environmental factors might cancel each other out or exacerbate the phenotype produced. Third, natural populations are genetically heterogeneous. In laboratory studies, the genetic background is standardised (recombinant inbred line) to allow researchers to attribute phenotypic variation to specific environmental stimuli and assess the interactions of genes and environment across several lines (Churchill 2007; Johannes et al. 2009). However, the number of recombinant inbred lines used in most studies is too small to encompass the amount of variation in the natural population (Churchill 2007). This is particularly

problematic for the study of ecological epigenetics, as the evolutionary implications of epigenetic variation will vary depending on the genetic variation available.

1.3.6 miRNA regulation

Apart from DNA methylation, there are other epigenetic mechanism such as histone modifications, chromatin remodelling and microRNA (miRNA). miRNAs are non-coding RNA and are usually 18-25 nucleotides long. miRNA was considered junk DNA previously but with new technology and techniques in molecular biology, researchers found that miRNAs play a role in regulating gene expression. They do this through binding to the untranslated region (UTR) resulting in decay of mRNA or suppressing protein translation. However, recent research has also found that miRNA not only down-regulate transcripts level but can also up-regulate the transcript level of the target gene (Hussain et al. 2012), and can bind to another region apart from the 3'UTR (Rigoutsos 2009).

miRNAs have been considered as epigenetic modulators because they have the ability to regulate the protein level in target mRNAs without modifying any of the gene sequences. miRNA can contribute to and consider as an important part of epigenetic regulation in three different ways. Firstly, the expression of miRNA is often regulated by the other epigenetic mechansims. Second, miRNA has the ability to repress the expression of other epigenetic factors. Third, some common target genes are modulated through the cooperation of both miRNA and other epigenetic machinery. miRNA variation may affect the heritable variation thorugh its effect on the epigenetics mechanisms mentioned above such as DNA methylation , histone modification through controlling expression of proteins involved in these modification. Next, miRNA usually binds to the 3'UTR and genome wide analysis has showed that gene with high level of miRNA binding at 3'UTR have low level of methylation level. In contrast, gene with low level of miRNA of binding at 3'UTR usually have high level of methylation at promoter region. This suggest that both miRNA and DNA methylation can reciprocally regulate gene expression at the genome level. Further, miRNA is also able to regulate the expression of HDAC enzyme which is a key component of histone modification. Taken together the role of miRNA in regulating gene expression and also involvement in

other epigenetic machinery it is possible that miRNA could be an important epigenetic factors in regulating wing polyphenism in insects.

In insects, mRNA is transcribed by RNA polymeraseII (pol II) which then results in the formation of primary miRNA (pri-miRNA; Lee et al. 2004). pri-miRNA will undergo a series of processing events to produce mature miRNA which then targets mRNAs (Figure 1.18). In the nucleus, pri-miRNA is processed by RNAse III enzyme Drosha together with Pasha resulting in a short hairpin structure around 70 bases, which is known as precursor miRNA(pre-miRNA) (Bartel 2009). Next, Exportin-5 will transport pre-miRNA into the cytoplasm for further processing whereby the terminal loop will be cleaved by the ribonuclease enzyme known as Dicer1 (Dcr-1). This cleavage results in a miRNA:miRNA duplex consisting of two nucleotide overhangs on both ends (Hutvágner et al. 2001). After the cleavage, miRNA duplex will be loaded to the Argonaute (Ago) protein and complete its action on the target sequences aided by the RNA-induced silencing complex (RISC).

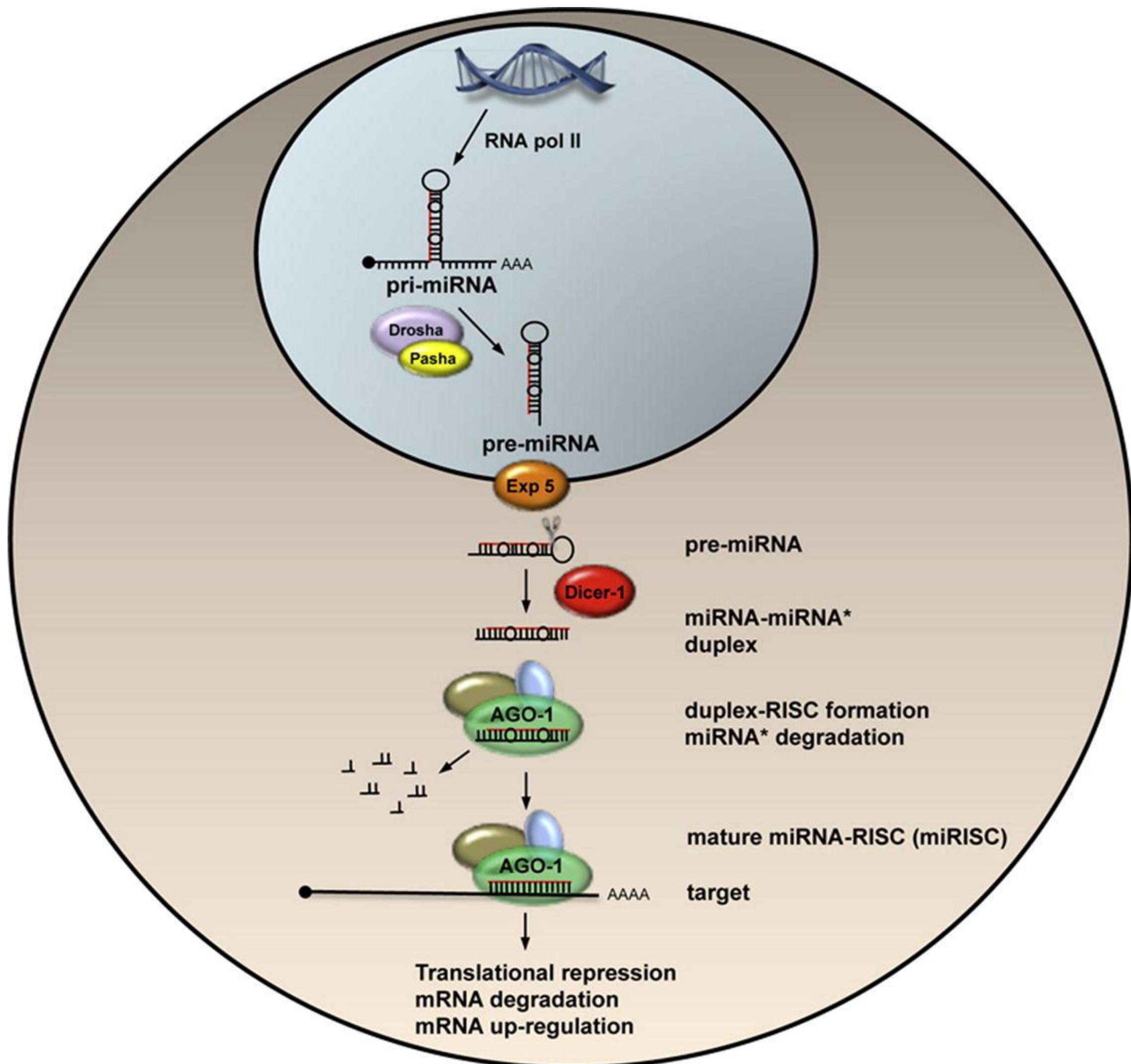


Figure 1.18. miRNA biogenesis pathway. First, RNA poly II will aid in the expression of miRNA gene expression in the nucleus (grey circle) resulting in primary miRNA (pri-miRNA) that consists of one or more stem-loop structures. Next, the RNase III enzyme Drosha together with Pasha will cleave the stem-loop structure producing a hairpin structure that is 70nt, which is known as precursor miRNA (pre-miRNA). Exportin 5 will then transport pre-miRNA into the cytoplasm (brown circle) followed by dicing of the hairpin terminal loop by the Dicer-1 enzyme resulting in a miRNA-miRNA duplex followed by incorporation into the RISC complex. The mature miRNA-RISC complex then interacts with the target sequences of the mRNA target resulting in protein degradation, translational repression or transcription upregulation. Adapted from Asgari 2013.

a) Role of miRNA in insect growth and development

The role of miRNA in insects has gained a lot of attention because of their conserved function in regulating signalling pathways that are essential for animal development. For example, studies have shown that dme-miR-8 can influence cell growth by regulating multiple peptide hormones in the fat body (Jin et al. 2012; Lee et al. 2015). Not only that, in the cockroach *Blattella germanica* miR-8 has been shown to regulate atrophin and thereby influence motor coordination (Rubio et al. 2013). Apart from that miR-14 has also been shown to regulate the gene hedgehog (hh) in the Hg signalling pathway, which is essential for many insects' developmental processes (Kim et al. 2014). Next, in *Drosophila* miR-305 has been shown to interact with both Notch signalling and insulin signalling pathway in intestinal stem cells to regulate self-renewal and differentiation. Next, the 20-hydroxyecdysone (20E) is essential for insect development and metamorphosis. By using *Bombyx mori*, Jiang et al. 2013 showed that the miR-281 was involved in Malpighian tubule regulation through suppressing the EcR-B gene.

b) Role of miRNA in insect behaviour

Apart from its role in growth and development, miRNA plays an important role in regulating insect behaviour. Recently, a study by Yang et al. (2014) using *Locusta migratoria* has shown that miR-133 was responsible for regulating the behavioural changes observed in the social and solitary phase of locusts by targeting key genes that were involved in the dopamine synthesis pathway. Apart from that, another study by Cristino et al. (2014) has shown that absence of miR-932 impacts the learning and memory in the honeybee *Apis mellifera.*

c) Role of miRNA in insect oogenesis and embryogenesis

Next, miRNA also plays an important role in oogenesis and embryogenesis. In *Drosophila*, miR-989 is responsible for the migration of border cells toward oocytes during oogenesis (Kugler et al. 2013). Apart from that, *Drosophila* lacking the miR-124 also showed reduced fertility while flies lacking miR-282 showed a reduction in viability and egg-laying, suggesting the importance of miRNA in embryogenesis (Vilmos et al. 2013). In the mosquito *Aedes albopictus*, lower levels of miR-286b embryos result in delayed hatching rate (Puthiyakunnon et al. 2013). Further, depletion of Argonaute 1 (Ago1) in *Locusta migratoria*, a key gene in regulating miRNA, has shown a huge impact on oocyte maturation and ovarian development (Song et al. 2013). Together, these studies show that miRNA plays an important role in regulating embryogenesis and oogenesis in insects.

d) Role of miRNA in insects wing development

In insects, wing development requires tight regulation of genes involved in cell-cell interactions and cell signalling. In *Drosophila* wing imaginal disc, the bantam miRNA is suppressed by the Notch gene, which would otherwise induce cell proliferation (Becam et al. 2011). Apart from that, bantam also targets Enabled, which then reduces the level of bantam through Notch, therefore limiting the proliferation rates and help to maintain the dorsal-ventral boundary. Next, another miRNA let-7 is also shown to impact wing development in *Drosophila*. By using mutant flies, Caygill and Johnston (2008) showed that flies lacking let-7 produce smaller wings. Apart from that, the wing imaginal disc in *Drosophila* usually stopped dividing after 24hr, but flies lacking let-7 showed continued division of cells in wing imaginal disc resulting in more but smaller cells in wing imaginal disc. In *Drosophila*, another miRNA miR-9a is also shown to be important in regulating wing development whereby mutant flies lacking miR-9a show a substantial loss of wing tissues (Biryukova et al. 2009). *Drosophila* LIM-only (dLMO) is a transcription factor that inhibits

the activity of the wing development gene Apterous that is responsible for dorsal identity of wings. Biryukova et al. (2009) showed that miR-9a regulates wing development by inhibiting wing cell apoptosis by repressing the transcription factor *Drosophila* LIM-only (dLMO).

e) Role of miRNA in insect phenotypic plasticity

Many insects can produce distinct phenotypes when exposed to environmental stress. For example, female aphids usually produce asexually under favourable conditions. However, when exposed to stressful conditions aphids can produce winged offspring. Further, female aphid can also switch to sexual reproduction when expose to low temperature. A study by Legeai et al., 2010 showed that seventeen miRNA were differentially expressed between sexual and asexual morphs in pea aphids, showing a potential role of miRNA in regulating sexual polyphenism in insects. Further, by studying termites *Reticulitermes.speratus*, Matsunami et al. 2019 reported eight miRNAs were differentially expressed between the soldier and workers suggesting that miRNA might play a role in regulating caste differentiation in insects.

1.4 General introduction and history of *Buchnera*

Buchnera was first characterised by the scientist Paul Baumann and was named after the German scientist Paul Buchner. *Buchnera aphidicola* is a primary endosymbiont found in sap-sucking insects, such as the pea aphid (*Acyrthosiphon pisum*) and other aphid species. *Buchnera aphidicola* is usually found in specialised aphid cells called bacteriocytes (Charles et al. 2011). *Buchnera aphidicola* is categorised under the phylum Gammaproteobacteria. All bacteria within this phylum are gram-negative. Many researchers have suggested that the *Buchnera* ancestor is similar to Enterobacteriaceae (Douglas, 1998). There is evidence that supports this argument; a study of 16S rRNA found that the two families have very closely related 16S rRNA structures (Munson et al., 1991). (Figure 1.19).

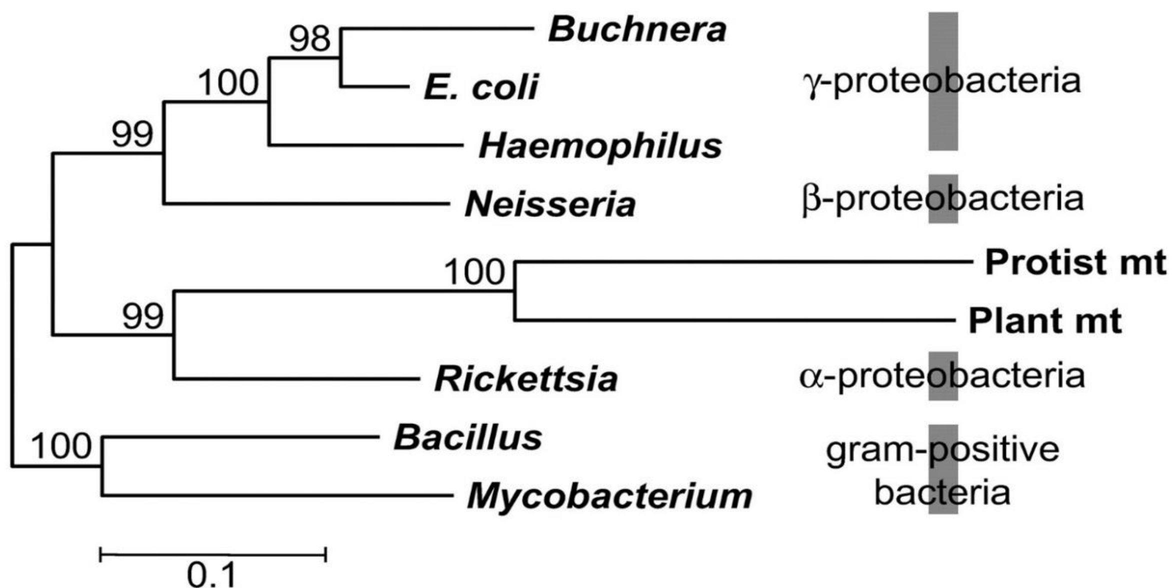


Figure 1.19. Phylogenetic tree from eubacteria and mitochondria using 16S rRNA gene. Number on the nodes represents the bootstrap values on neighbour joining analysis. The Bar 0.1 represents the sequence divergence (Adapted from Itoh et al. 2002).

Buchnera aphidicola is very small in size (around 3μm) and has a gram-negative cell wall, a common characteristic of Enterobacterales. However, a key component in *Buchnera* that

is very different from the majority of gram-negative bacteria is that it lacks the genes required to produce lipopolysaccharides for the outer membrane.

1.4.1 Genome structure of *Buchnera aphidicola*

Buchnera aphidicola has an extremely small genome (<1Mb), one of the smallest known genomes in living organisms. The reason for this small genome could be due to its association with aphids and the transmission mode (vertical transmission) that limits crossover events. This leads to the deletion of many genes, especially those involved in anaerobic respiration and the synthesis of fatty acids and complex carbohydrates. The loss of these genes is predicted to have occurred after a symbiont relationship was established between aphids and *Buchnera*. *Buchnera aphidicola* has one circular chromosome and contains either one or two plasmids (Thomas et al. 2009). The genome structure of *Buchnera* also shows a few unique characteristics, including a high G+C and low A+T ratio in the chromosome. This is due to the substitution of many A+T nucleotides with G+C pairs upon the divergence of the species (Gómez-Valero et al. 2007). As mentioned above, the *Buchnera* genome also lacks genes that produce lipopolysaccharides due to the long association with aphids, which began around two million years ago (MYA). *Buchnera* is a mutualistic endosymbiont that is often found in a host (aphid) and the pattern of gene reduction in *Buchnera* is similar to other small-genome bacteria (Moran and Wernegreen, 2000). However, one difference between the *Buchnera* genome pattern and other small-genome bacteria is that 55 loci (around 10% of the *Buchnera* genome) is related to the synthesis of amino acids required by the host pea aphid (Shigenobu et al. 2000). As well as a reduced number of genes involved in metabolism, *Buchnera* also have fewer genes involved in other cellular processes, such as transcription, translation and replication.

1.4.2 Genome evolution in *Buchnera*

One of the questions surrounding the reduced genome size in *Buchnera* is whether genes were lost gradually or through several bulk losses. Endosymbionts usually live within host tissues that provide a metabolite-rich environment and a low growth rate. These factors often lead to a more relaxed selection on many loci, with genetic drift more frequently affecting the genome. This is clearly reflected in the pattern of sequence evolution in *Buchnera* and many other endosymbionts (Wernegreen and Moran 1999; Spaulding and van Dohlen 2001). This relaxed selection is expected to increase the loss of genes that are both superfluous and beneficial. The small genome size seen in *Buchnera* is predicted to be due to the lack of gene retention through selection, rather than direct selection for a reduced genome size (Mira et al. 2001). The *Buchnera* cell contains roughly 50-200 chromosomes, in which non-functional pseudogenes can exist for a long period of time. This suggests that direct selection for a smaller genome is very unlikely.

By comparing the *Buchnera* genome to a reconstructed ancestral genome, researchers found that the reduction in the *Buchnera* genome involved a large deletion, followed by chromosome rearrangements (Moran and Mira 2001) (Figure 1.20).

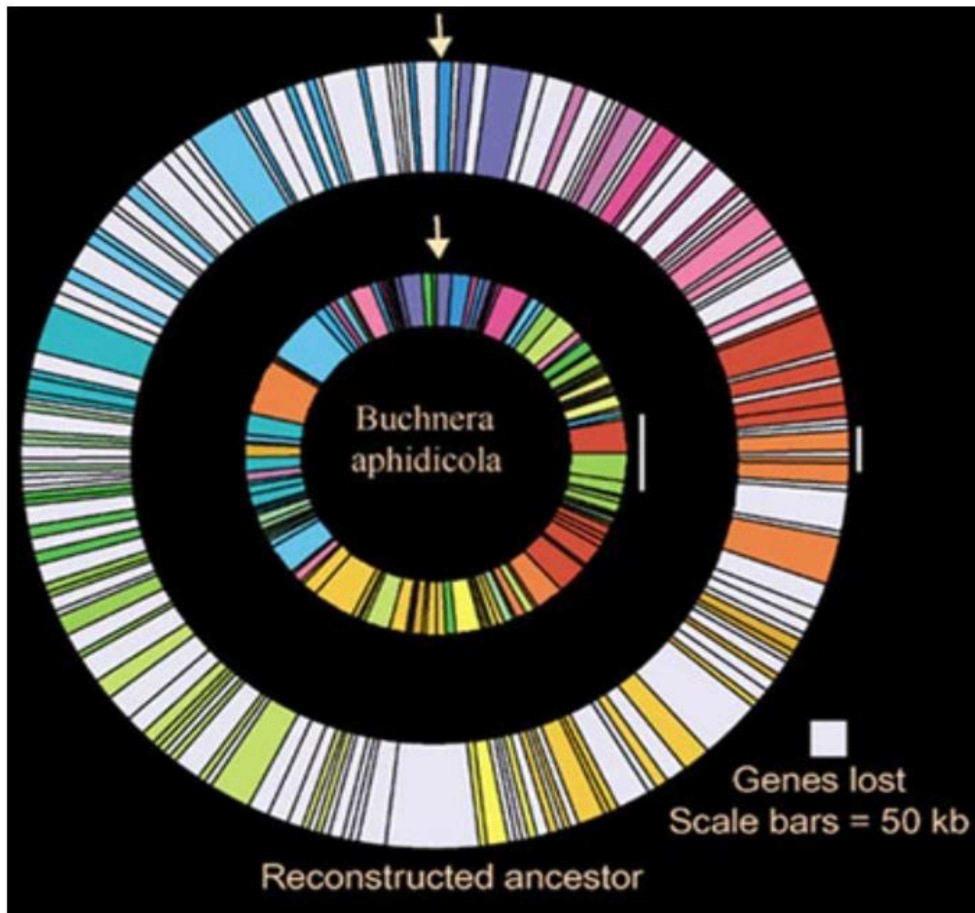


Figure 1.20. Genome of *Buchnera aphidicola* and reconstructed ancestor. White box= gene lost, colour fragments= syntenic fragments, lost fragments between syntenic fragments is represented by grey line (Adapted from Moran and Mira 2001).

This reduction through a large deletion is supported by the distribution of the size region in the ancestor genome. The genes in *Buchnera* are aggregated, rather than a mix of retained and lost genes from the ancestral genome. The absence of a positive association between spacer length and number of genes lost through deletion in the syntenic fragments further supports the fact that the *Buchnera* genome reduction is more likely to have occurred through a large deletion that involved many genes, than through a gradual process. The genome structure of *Buchnera* is also unique in the sense that there is only one copy of each gene that encodes for 16S, 23S and 5S rRNA, instead of the multiple rRNA genes seen in most bacteria (Baumann et al. 1989).

The *Buchnera* genome also shows a significant loss of tRNA (32 tRNAs), with most amino acids only containing one tRNA; in comparison, *Escherichia coli* has 86 tRNAs. As discussed above, many endosymbionts have lost genes involved in DNA repair and this applies to the *Buchnera* genome as well. However, one unique characteristic in the *Buchnera* genome is the loss of the gene RecA (important for recombination and repair; Moran and Wernegreen 2000). However, the *Buchnera* genome retained the gene RecBCD, which was lost among several other small-genome bacteria (Latorre et al. 2005). The retention of RecBCD could have been promoted by the flanking gene argA, which is also retained in *Buchnera* and lost in other small-genome bacteria. The reason for argA retention might be related to the need to synthesise the essential amino acid arginine for the host. *Buchnera* is also unusual in that the genes involved in repairing UV damage to DNA (uvrA, uvrB, uvrC) are absent. The gene uvrA is part of the synthetic region, while uvrB and uvrC fall between the synthetic regions. One explanation for the loss of these genes could be due to an initial large deletion accompanied by a gradual loss of genes involved in the same pathway.

1.4.3 Relationship between *Buchnera* and pea aphids

Endosymbiosis is one of the most common phenomenon seen in insects and plays an important role in the ecological and evolutionary success of many species. One of the most studied endosymbiotic relationships is that of the endosymbiont bacteria *Buchnera aphidicola* and its host, the pea aphid (*Acyrthosiphon pisum*) (Moran et al. 2003). Aphids feed on the phloem sap of plants, which often lacks essential amino acids (these cannot be synthesised by the insect). In order to survive on this diet, the aphid is very dependent on the endosymbiont *Buchnera* to compensate for the lack of the essential amino acids (Gündüz and Douglas 2009). Previous research by Wilson et al., 2010 has shown that the nutritional needs in aphids are compensated by the holosymbiont relationship between the

aphid and the *Buchnera*. This suggest that aphids and the *Buchnera* are dependent on each other for survival. The nutritional-dependent relationship of aphids and *Buchnera* was suggested by Paul Buchner between the 1960s and the 1990s, using artificial diet experimentation. As technology advanced, and with the release of the full genome sequence of pea aphids, the nutritional relationship between aphids and *Buchnera* has been confirmed. There are four central issues involved in aphid-*Buchnera* symbiosis.

a) Lateral gene transfer

First, the large reduction in the *Buchnera* genome and the fact that the aphid genome has undergone transposition of mitochondrial genes to the nuclear genome has led to the prediction that some genes of *Buchnera* origin may be detected in the pea aphid genome. However, recent research has provided evidence in contrast to this and has concluded that the aphid genome does not contain any genes of *Buchnera* origin (Nikoh et al. 2010) (Figure 1.21).

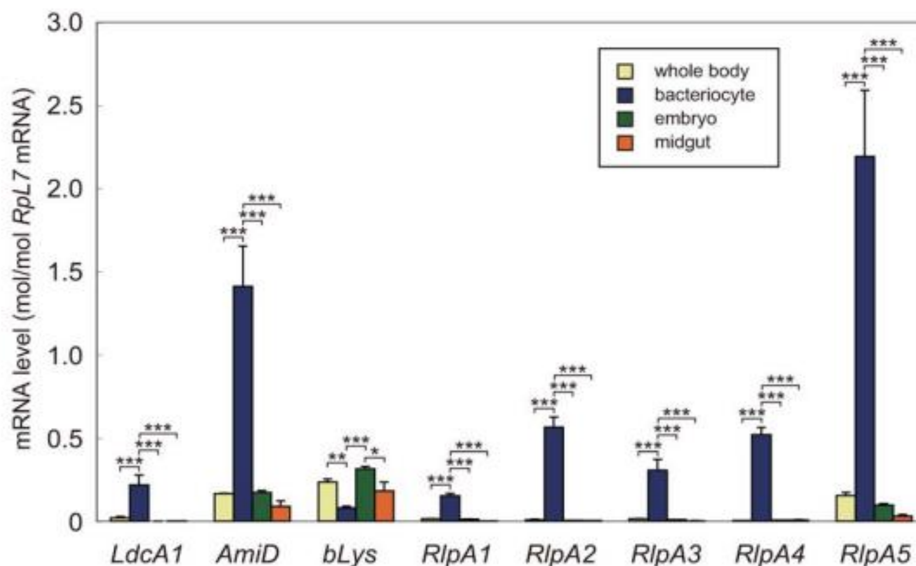


Figure 1.21. Expression profile of genes that are thought to be laterally transferred. Yellow= aphid whole body, blue=bacteriocyte, green=aphid embryo, orange=aphid midgut. The expression profile shows that the expression of these genes that are thought to be laterally transferred are

only highly expressed in bacteriocytes (where Buchnera is found) and not in the aphid itself. (Adapted from Nikoh et al., 2010).

The absence of *Buchnera* genes within the aphid genome rejects the idea that lateral gene transfer is the cause of the genome reduction in *Buchnera*.

Although pea aphids do not contain any genes of *Buchnera* origin, they have acquired some genes through lateral gene transfer from multiple sources. First, the aphid genome contains 56 different mitochondrial genes, but most of these are pseudogenised. Second, the genes in the carotenoid biosynthesis that are required for body colour polymorphism were obtained through lateral gene transfer from fungi. Lastly, 12 genes were obtained from an α-proteobacteria that is not of *Buchnera* origin. Among the groups of genes obtained through lateral gene transfer, the gene from the a-proteobacteria plays an important role in the symbiotic relationship between *Buchnera* and aphid. This includes the gene LdcA (LD-carboxypeptidase), RlpA (rare lipoprotein A) and AmiD (N-acetylmuramoyl-L-alanine amidase). LdcA and AmiD are essential for the production of peptidoglycan, which forms the essential component of the bacterial membrane in *Buchnera* (Houk et al., 1977). Whole genome analysis of the aphid genome suggests that the RlpA gene contains N-terminal eukaryotic-type signal peptides, indicating that the bacterial gene has fused with the aphid signal peptide; these are predicted to target secretory pathway proteins. However, more examination of the structure and function of genes from the α-proteobacteria is required to provide further insight on the importance of these genes in symbiotic regulation.

b) Host immunity

The second important aspect of the *Buchnera*-aphid symbiosis is host immunity. The majority of insects have immune systems that protect them from pathogenic and parasitic organisms. Therefore, the presence of *Buchnera* in an aphid host requires the modification of the aphid’s immune system. Research conducted by wounding aphids with a bacteria-contaminated needle found that aphids show no immune response or

gene expression related to protection against bacteria (Altincicek et al. 2008). In fact, the pea aphid genome has lost many genes that are essential for immunity functions, but retained some, such as the JAK/STAT and JNK signalling pathway (Gerardo et al. 2010). The aphid genome also shows an absence of immunodeficiency (IMD) pathway genes, such as dFADD and Dredd. The peptidoglycan receptor proteins (PGRPs) that are critical for detecting bacterial invasion are also missing in the pea aphid genome. This suggests that the reduced immune response in pea aphids might reflect the success of the symbiotic relationship with *Buchnera* throughout evolution (Shigenobu and Wilson 2011).

c) Symbiotic metabolism (nutritional relationship)

The third aspect is the nutritional relationship between aphids and *Buchnera*. One of the main factors of the symbiosis between *Buchnera* and aphids, this relationship has allowed the successful coevolution of both species. The pea aphid phloem sap diet is not enough to ensure its survival, as phloem sap lacks many amino acids, such as histidine, leucine, methionine, valine and arginine, that are critical for the urea cycle pathway (Sandström and Pettersson. 1994). The amino acid obtained from phloem sap is asparagine, which is processed by the *Buchnera* to provide the essential amino acids. The asparagine ingested by the aphid will first be converted to aspartate by the enzyme asparaginase, followed by a process called transamination with the help of the enzyme aspartame transaminase; this leads to the production of glutamate (Wilson et al. 2010). Both aspartate and glutamate are required for the production of essential (arginine, isoleucine, lysine, threonine, histidine, phenylalanine and valine) and non-essential (serine, tyrosine and alanine) amino acids. The non-essential amino acid serine can be further processed by the Buchnera to produce another two essential amino acids (methionine and tryptophan) (Figure 1.22).

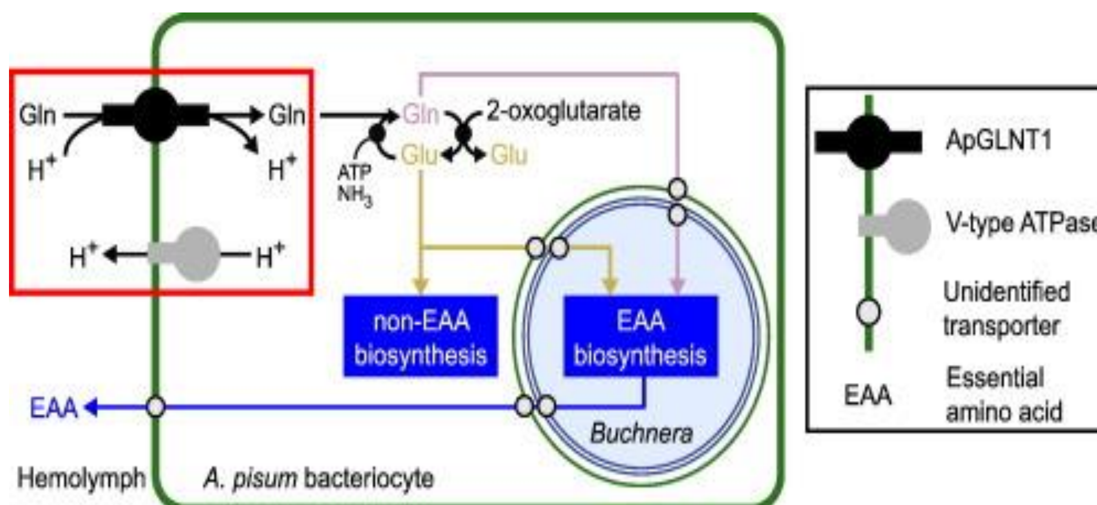


Figure 1.22. Simplified schematic diagram of amino acid production in the *Buchnera*-aphid symbiosis. EAA=essential amino acid, non-EAA=non-essential amino acid, Gln=Glutamine, Glu=Glutamate. Asparagine and glutamine are the most common amino acid that is usually obtain by pea aphid from the phloem sap. Next, both asparagine and glutamine is transferred from the aphid into the bacteriocyte in which Buchnera resides and undergoes a series of reactions to produce EAA and non- EAA, where the EAA will be transferred out to the hemolymph of the host (aphid). (Adapated from Price et al., 2015).

Although the *Buchnera* genome has retained almost all genes necessary for the synthesis of essential amino acids, the genome has lost some genes involved in producing non-essential amino acids, such as IlvE (the branch-chain amino acid transaminase), IlvA (isoleucine pathway synthesis) and MetB/C (methionine biosynthesis pathway; Shigenobu et al. 2000). (Figure 1.23).

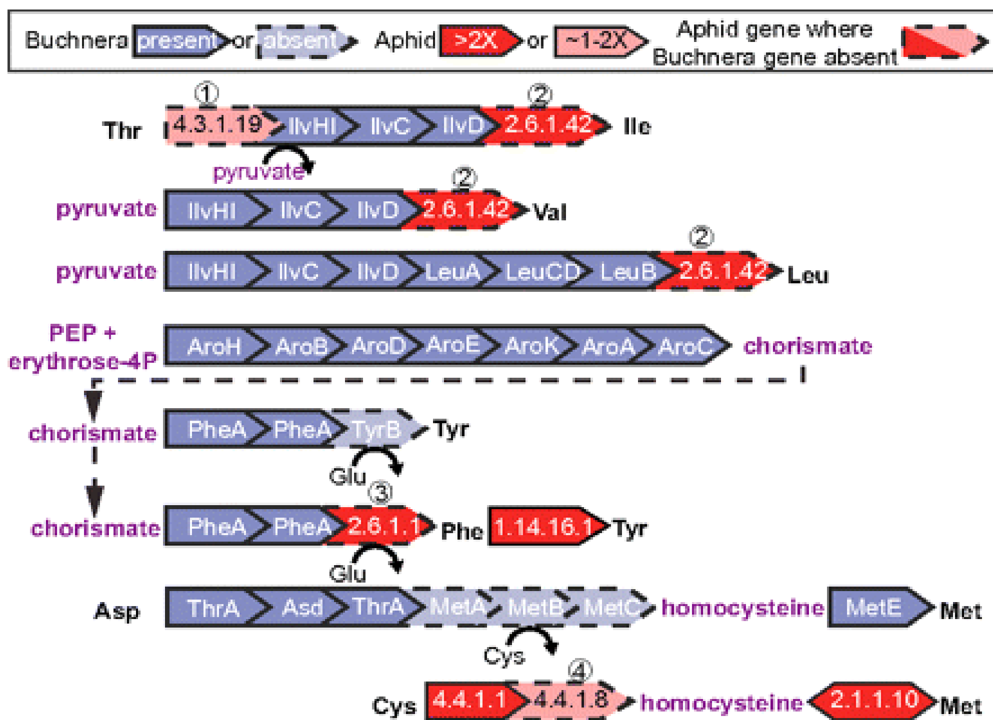


Figure 1.23. Amino acid pathway in *Buchnera* with aphid gene complementation. Light blue=genes that are absent in Buchnera, dark blue=genes that are present in Buchnera, Light red=genes absent in aphid, dark red=genes present in aphid. Number circle=enzymes that are absent in Buchnera that are complemented by increased expression in aphid enzymes. Buchnera is missing genes such as IlvA (1), IlvE (2), TyrB (3), MetC (4) (Adapted from Hansen and Moran 2011).

Further analysis of the *Buchnera* genome has shown that this loss of genes required for non-essential amino acid production is compensated by the functions of other genes. As mentioned above, the main amino acid asparagine is converted to glutamine by the aphid. Glutamine is further processed and converted to glutamate by glutamate-ammonia ligase through Buchnera (Sasaki and Ishikawa 1995). Glutamate-ammonia ligase plays an important role in nitrogen recycling and is one of the five most highly expressed genes in *Buchnera. Buchnera* usually uptake glutamate (that was converted from glutamine in the bacteriocyte) through the symbiosomal membrane which can act as an amino donor for the synthesis of the essential amino acid phenylalanine (Whitehead and Douglas 1993) . Sasaki and Ishikawa found that the symbiosomal membrane in *Buchnera* is the most metabolically

active area, with the lowest synthesis of alanine, aspartate and proline (Sasaki and Ishikawa 1995).

The fact that *Buchnera* is essential for the survival of aphids can be seen by disrupting the endosymbiont dynamic in the aphid. *Buchnera* in aphids can be eliminated through antibiotics, allowing research into the effect of *Buchnera* absence on aphids (Zhang et al. 2015; Machado-Assefh et al. 2015). One aspect of aphids that is affected is the fresh weight, as the weight of antibiotic-treated aphids is significantly lower (around 28%) than control aphids (Adams et al. 1996). This obvious reduction seen in the absence of *Buchnera* is due to protein content, as *Buchnera* are involved in producing essential amino acids. It is not surprising that an absence of *Buchnera* leads to a reduced protein level in aphids. Antibiotic-treated pea aphids also have a lower reproduction rate, as protein content is associated with reproduction ability (Raikhel and Dhadialla 1992).

The soluble sugar and glycogen content is also reduced in antibiotic-treated aphids. Glycogen and soluble sugar are essential for providing energy for pea aphid muscle movement (Hansford and Johnson 1975). Although the effect of low soluble sugar and glycogen levels on pea aphid behaviour has not yet been investigated, some research shown that disruption of the glycogen and soluble sugar content affects green peach aphid feeding behaviour. There is also an increase in neutral fats and lipids as a consequence of transforming extra soluble sugar for lipid synthesis (Febvay et al., 1999).

The evidence above demonstrates that the endosymbiont *Buchnera* and aphids are dependent on each other for survival. The absence of *Buchnera* in aphids could have a huge impact on aphid evolution (Ning et al., 2018), as it can affect multiple aspects, such as reproduction, energy levels and host acceptance. Future strategies for controlling crop damage could focus on endosymbiont disruption in aphids.

d) Regulation of gene expression in *Buchnera* and aphids

The fourth aspect is the regulation of gene expression in both *Buchnera* and aphids. Research on the *Buchnera* genome has found that the endosymbiont lacks genes that encode for regulatory proteins, such those seen in other small-genome bacteria (Fraser et al. 1995). Although *Buchnera* has many genes involved in amino-acid biosynthesis, its genome only retains one amino acid biosynthesis gene, known as MetR (Moran et al. 2005). The *Buchnera* genome also lacks regulatory genes involved in responses to environmental changes (Hoch 2000). In contrast to the endosymbiont, the host aphid genome possesses a great diversity of regulatory genes in comparison to other arthropods. Some of the transcription factors in aphids, such as Distal-less, are expressed in the bacteriocytes; however, the function of the gene in regulating the symbiotic relationship remains unknown (Braendle et al. 2003). Another aspect that needs to be determined is if the lack of regulatory genes in *Buchnera* is associated with reduced regulatory capacities. Analysis of the transcriptional profile of *Buchnera* through microarray has shown that *Buchnera* gene expression does not change significantly when exposed to environmental stress. For example, when exposed to heat stress, only a small number of genes are expressed differently (five of 20; Wilcox et al. 2003). This suggests that the transcriptional dynamics of *Buchnera* are stable.

The evidence discussed above shows that the endosymbiont bacteria and aphids are co-dependent for survival and have evolved together. However, a recent study has shown that the host (aphid) is usually the dominant player in the symbiotic relationship. By reconstructing the metabolic network of aphids and *Buchnera*, Thomas found that carbon and nitrogen intake could affect amino acid production, even though transcriptional regulation is absent (Thomas et al. 2009). This observation suggests that aphids change their feeding behaviour to regulate the amount of carbon and nitrogen provided to the *Buchnera*, controlling the production of essential amino acids

by the *Buchnera*. Therefore, further studies on this symbiotic relationship could focus on feeding behaviours and gene regulation in the pea aphid.

1.4.4 Secondary endosymbionts in aphids

Apart from the primary endosymbiont, some biotypes of pea aphid also contain secondary endosymbionts, such as *Rickettsia, Hamiltonella* and *Wolbachia*.

a) *Rickettsia*

One of the secondary endosymbionts in pea aphids is known as *Rickettsia*, of the genus *Rickettsia*. The endosymbiont *Rickettsia* is commonly found across a wide range of arthropod species and is usually associated with host cells (Dasch and Weiss 1992). The specific cells that host the *Rickettsia* endosymbiont remain elusive. However, some studies have suggested that Rickettsia in pea aphids is commonly found in the haemolymph – whereby an infected pea aphid transfers the *Rickettsia* to an uninfected pea aphid through haemolymph (Chen and Purcell 1997). Some research has also found that the *Rickettsia* endosymbiont can be localised in secondary mycetocytes and sheath cells (Sakurai et al. 2005).

The effect of the *Rickettsia* symbiont on pea aphids has been investigated using antibiotic ampicillin to specifically eliminate the *Rickettsia* bacteria without affecting the *Buchnera* (Tsuchida et al. 2004). Tsuchida shown that *Rickettsia* has a negative impact on pea aphid fitness. Aphids infected with *Rickettsia* have fewer offspring and a lower fresh body weight in comparison to uninfected aphids. Research has also shown that *Buchnera* population dynamics are affected by *Rickettsia*, with infected aphids having a lower density of *Buchnera*. There are two possible explanations for these observations. One is that Rickettsia directly affects *Buchnera* through suppression, which leads to decreased aphid fitness. The second is that *Rickettsia* affects the pea aphid host directly, which

indirectly leads to a decrease in *Buchnera* density. More research is needed to confirm how *Rickettsia* affects pea aphid fitness and *Buchnera* population dynamics.

Rickettsia has been shown to have a positive effect on pea aphid fitness when the host aphids are subject to environmental stress (Montlor et al. 2002). Montlor subjected *Rickettsia*-infected pea aphids and uninfected pea aphids to heat stress and found that the *Rickettsia*-infected pea aphid showed increased reproduction. *Buchnera* are more sensitive to heat stress and, therefore, the increase in reproduction in Rickettsia-infected pea aphids could be due to the rescue of the *Buchnera* by the *Rickettsia* . However, the increase in pea aphid reproduction could also be a direct effect of the Rickettsia on the host. More research is needed to confirm these relationships. There are some observations consistent with *Rickettsia* benefitting pea aphids –pea aphid populations in warmer climate have higher levels of *Rickettsia* in comparison to those in cooler climates. Since *Rickettsia* could have both positive and negative impacts on pea aphid fitness, it is essential to determine the effect of *Rickettsia* in pea aphids in different environmental contexts (Chen et al. 2000).

b) *Hamiltonella defensa*

Another secondary endosymbiont found in pea aphids is *Hamiltonella defensa*. *Hamiltonella* is a type of Gammaproteobacteria that can also be found in other sap-sucking insects, such as whiteflies and psyllids (Russell et al. 2003). The presence of *Hamiltonella defensa* protects the pea aphid in the presence of parasitoid enemies (such as *Aphidius ervi*) by killing the parasitoid larvae. The ability of *Hamiltonella* to decrease parasitoid attack is associated with a lambda-like bacteriophage (APSE). *Hamiltonella* with low level APSE confers lower protection against parasitoids, further supporting the fact that the genes encoded in APSE are essential for protection against parasitoids (Degnan and Moran 2008). The normal functioning of *Hamiltonella defensa* requires essential amino acids that are not found in the phloem diet of the pea aphid,

which suggest that *Hamiltonella defensa* might be dependent on *Buchnera* for the synthesis of these essential amino acids.

Apart from conferring protection against parasitoid wasps, *Hamiltonella defensa* is effective in defending the pea aphid against predators such as ladybirds (*Hippodamia convergens*). Costopoulos et al. (2014) showed that *Hamiltonella defensa* protects pea aphids by affecting several aspects of the predator. Ladybird larvae that preyed on pea aphids with *Hamiltonella defensa* had a decreased rate of survival from egg hatching to pupation. However, ladybird adults that fed on *Hamiltonella defensa*-infected pea aphids had a higher fresh weight in comparison to those that fed on pea aphids without *Hamiltonella defensa*. This observation is particularly interesting as *Hamiltonella defensa* decreases ladybird larvae survival but increases the adult weight. One of the possible explanations for this observation is that ladybird mothers who feed on *Hamiltonella defensa*-infected pea aphids evaluate the environment (diet) as being detrimental to their fitness (Fox and Czesak, 2000). Therefore, the larger size in ladybirds could be a response to changes in the environment, by producing bigger eggs that develop into bigger adults when food resources are of poor quality (Fischer et al., 2011).

c) *Wolbachia pipientis*

Wolbachia pipientis is another secondary symbiont endosymbiont found in pea aphids. *Wolbachia* is commonly transferred through the mother and is found in 65% of arthropod species (Jeyaprakash and Hoy 2000). Wolbachia is usually found in somatic tissues and is usually localised in reproductive tissues and organs. However, *Wolbachia* is very rare in aphid species and studies have shown that this endosymbiont is absent in the pea aphid (Nirgianaki et al. 2003). The role of *Wolbachia* in aphid species remains largely unknown, but *Wolbachia* is usually involved in reproductive alterations in its host. Since a lot of aphid

species have complex lifecycles, such as sexual polyphenism, the impact of *Wolbachia* on these alterations needs to be investigated (Stouthamer et al. 1999).

d) *Regiella insecticola*

The final secondary endosymbiont in aphids discussed here is *Regiella insecticola*. The main role of *Regiella insecticola* is similar to the secondary endosymbiont *Hamiltonella defensa*, which protects pea aphids against parasitoid wasps. A recent study has shown that *Regiella insecticola* protects pea aphids from fungal infection (Łukasik et al. 2013). Fungi are commonly used as a biocontrol against pea aphids and pea aphids usually encounter a wide range of fungal infections in their natural habitat. To determine the role of *Regeilla insecticola*, one study exposed pea aphids to two fungal pathogens, known as *Zoophthora occidentalis* and *Beauveria bassiana* (Parker et al. 2013). *Zoophthora occidentalis* is a type of entomopathogen that is highly specific towards aphids, while *Beauveria bassiana* is a general fungal pathogen that can affect a wide range of insects. Parker et al. found that aphids harbouring *Regeilla insecticola* had higher survival rates upon infection with *Zoophthora occidentalis* (2013). This supports results from previous research, which found that *Regale insect cola* confers protection against another aphid-specialised pathogen, Pandora (Scarborough et al. 2005). In contrast, *Regale insect cola* does not have any effect on aphid survival when exposed to the generalist fungal pathogen *Beaveries basin*.

1.5 Summary

Pea aphids have been widely used as a model to study polytheisms due to their fast reproductive time and easy maintenance. Most studies on trade-offs between reproduction and dispersal in pea aphids have only focused on the winged and wingless morphs. Since pea aphids exist in many different genotypes, with some genotypes changing their body colour instead of producing wings, it is important to understand whether trade-offs also exist in the genotype that cannot produce wings. Trade-offs in dispersal and reproduction could play an important role in animal adaptation to stressful conditions and possibly affect evolutionary trajectories. However, most studies on trade-offs in insects have only been investigated for one generation. Such studies rarely investigate the effect of trade-offs on the subsequent generation. Further, most studies on pea aphid wing polytheism only focused on gene expression differences but not expression regulation. Since pea aphids reproduce asexually (daughters are clones of the mother) under optimal conditions, polytheism in aphids could be predicted to be regulated through epigenetic mechanisms. Furthermore, there is mounting evidence that epigenetic mechanisms play a role in regulating polytheisms in insects. This thesis aims to address the significant gaps in our understanding of trans-generational trade-offs in pea aphids with a particular focus on two different genotypes that respond to environmental stress differently by producing distinct morphs. This thesis also aims to understand the mechanisms in regulating wing polytheism in aphids, the transcriptome and methylene profile of different pea aphid morphs and the relationship between gene expression and DNA methylation in pea aphids. Finally, this thesis also aims to understand the role of other epigenetic mechanisms microRNA in regulating wing polytheism and their relationship with gene expression.

1.6 Hypothesis

Different pea aphid genotypes respond differently to environmental stress and trade-offs of reproduction. The ability to disperse in pea aphids could last more than one generation. I predict that the aphid development genes were differentially expressed between the dispersal and non-dispersal morphs. Next, I predict that epigenetic mechanisms such as DNA methylation, miRNA regulation could play an important role in regulating wing polytheism and development in pea aphids.

1.7 Thesis aims

The principal aim of the experiments presented in this thesis is to understand the role of epigenetic mechanisms in regulating wing polytheism in insects and the trade-offs of reproduction and dispersal in pea aphids. Crowding conditions were used to trigger the different morphs in two aphid genotypes and the subsequent reproduction success of the F0 and F1 generations was measured. Next, we investigated the genes involved in multiple pea aphid development processes at DNA methylation and gene level. Next, we investigated the transcriptome and methylene profiles of different pea aphid morphs. Finally, we investigate the miRNA profile of different pea aphid morphs. The aims of each chapter in this thesis are presented below.

Chapter 3 - Response of pea aphid genotypes to environmental stress and the trade-offs between reproduction and dispersal morphs

- Investigate the degree of alternative morph production between the two pea aphid genotypes
- Investigate the trade-offs between aphid morphs in two genotypes (N116 winged vs N116 wingless and N127 pale vs N127 red)
- Investigate possibility of transgenerational trade-offs in pea aphids

Chapter 4 – Investigate the expression of genes involved in different pea aphid developmental processes (wing development, reproduction, stress response) and the transcriptome profiles of different pea aphid morphs

- Measure the expression level of genes involved in different pea aphid developmental processes such as reproduction, wing formation, carotene production and stress response by qPCR
- Investigate the transcriptome profile of different pea aphid morphs in two genotypes (N116 winged vs N116 wingless and N127 pale vs N127 red) by RNA-seq
- Investigate the differences in GO and KEGG pathways in pea aphid morphs
- Integrate the results of transcriptome profile (RNA-seq) with MBD-seq (data reported in Chapter 5)
- Investigate the protein expression of candidate gene

Chapter 5 Investigate the DNA methylation level of genes involved in different pea aphid developmental processes (wing development, reproduction, stress response), and the methylome profile of different pea aphid morphs

- Measure the DNA methylation level of genes involved in different pea aphid developmental processes such as reproduction, wing formation, carotene production and stress response by pyrosequencing
- Investigate the methylome profile of different pea aphid morphs in two genotypes (N116 winged vs N116 wingless and N127 pale vs N127 red) by MBD-seq
- Integrate the results of transcriptome profile (chapter 5) and methylome profile to provide an insight into the correlation between gene expression and DNA methylation

Chapter 6- The role of miRNA in regulating aphid polyphenism

- Investigate the miRNA profile of different pea aphid morphs in two genotypes (N116 winged vs N116 wingless and N127 pale vs N127 red)

- Determine the possible roles of novel miRNA in pea aphids in pea aphid polyphenism

Chapter 2: Methods

2.1 Materials

Two different genotypes of pea aphid (*Acyrthosiphon pisum*) are used in the experiment because of their phenotypic differences (body colour morphs, virulence level and reproduction rate). The aphid genotype (N116) has green body colour (origin: near Slough, Berkshire, UK) while the N127 genotype has red body colour (origin: near Slough, Berkshire, UK) (Table 2.1). Both aphid genotypes were obtained from Dr Colin Turnbull from Imperial College London since 2019 and has been kept in separate mesocosms at low density for many generations to avoid any alternative morphs formation (Kanvil et al. 2014).

Table 2.1. Differences between life-history traits of two aphid genotypes (N116 vs N127)

N116	N127
Very virulence (aphid that can feed on non-host plant) (Kanvil et al. 2014)	Less virulence
Green colour (1st nymphal- adult stage)	Red colour (1st nymph – adult stage)
Higher reproduction rate	Lower reproduction rate
Respond to environmental stress by producing winged offspring	Respond to environmental stress by changing body colour from pale to red

2.1.1 Aphid stock maintenance

The minor cultivar of faba bean *Vicia faba minor* obtained from Lebanon (Abido.com) is used for stock aphid maintenance. Three faba plants are placed in a plastic container (H25cm x L17cm). Next, the aphids are placed onto the plant and then the top of the container is covered with a polyethylene UV stabilised fine mesh (0.25mm x 0.8mm) (Allotment-garden, UK) to prevent any aphids from escaping (Figure 2.1). The aphid cages were kept in growth chambers (LMS Cooled Incubation, UK) with a 16h/8h light dark cycle, 22°C±0.3 and 80%±3% humidity. The plants are changed every five to six days to ensure the food quality is optimal. Some aphids are transferred to a test tube to prevent

overcrowding, which can lead to wing induction. When changing the food for the aphids, the stems of the previous plants are cut and placed into a beaker. They are then gently shaken to remove the aphids. A paintbrush is used to brush the plant to ensure there are no remaining aphids on the plant. The two different aphid lines are transferred to different sides of the working benches to avoid mixing up the two different lines. Two containers of stock aphids are set up for each aphid biotype with two plants per container.

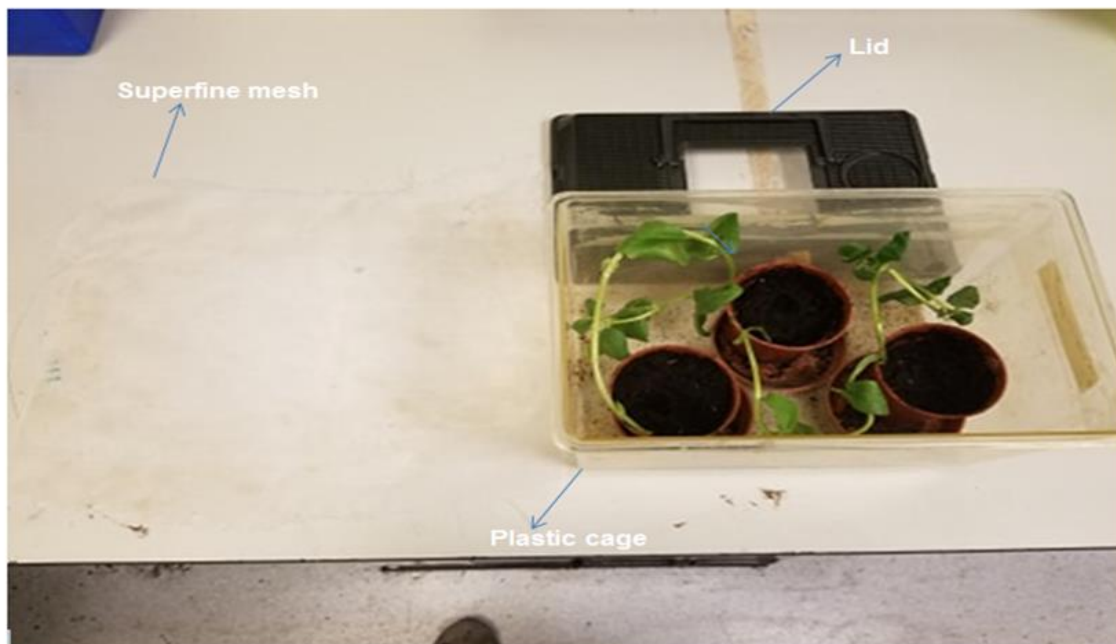


Figure 2.1. Stock aphid maintenance. Two/Three plants are placed in the plastic container and infested with pea aphids. The top is covered with a superfine mesh followed by the lid.

2.1.2 Faba bean sowing (stock plant)

Plastic pots (9cm diameter) were filled up to two thirds with soil (Levington advance seed and modular F2 compost) (ICL Specialty Fertilizers, Ohio, United States) with four seeds placed in the soil. Additional soil is added to cover up the seeds to avoid any exposure of seeds to bacteria and humidity from the environment, which will result in rotten seeds. Next, the pots are placed on a plastic tray and filled with around 500ml of water (Figure 2.2).

Water is added onto the tray every four days to ensure that plants do not dry out. The seeds will take around seven to eight days for germination.

Figure 2.2. Features of stock plant maintenance. Six pots were filled with two-third of soil followed by placing four seeds onto the soil. The seeds were further covered with some soil to prevent exposure to humidity and bacteria from the environment. All pots were placed on a tray and filled with water every three days.

2.1.3 Stock plant and experiment plant maintenance

Upon germination, the plant is left to grow to around 4cm before they are transferred to individual plants to avoid competition for resources in the same pot. Next, the pots are squeezed gently on both sides to gently loosen the soil. The plants are gently shaken to remove the excess soil and plant roots are carefully separated from each other to avoid any damage to the plants. Next, the plant is placed into individual pots and filled with soil followed by watering the soil. Further, the pots are placed on a tray followed by filling the tray with water and placed in a growth chamber 22°C±0.3 to grow. For all the experiments, plants were grown for another three weeks before being infested with aphids.

2.2 Mesocosm experiment

2.2.1 Alternative morph production through crowding condition

Six (three weeks old) fava bean plants were placed into a mesocosm container and the alternative morph inducing crowding conditions were triggered by placing 50 4th instar female pea aphids on the fava plant for 15 days (Figure 2.3A, B). The plants were watered every 3 days followed by randomisation of the mesocosm position in the chamber to eliminate any possible light regime effect on the aphid. Starvation is essential because it increases the mobility of aphids, thereby increasing the chance of aphids coming in contact with one another which ultimately leads to the production of alternative morphs (Sutherland 1969). At the end of the experiment, both dispersal and non-dispersal morphs were counted and flash frozen in liquid nitrogen followed by storing at -80°C for further analysis. For winged morph collection only adult aphid with fully expanded wing were collected, while for pale morphs only adult aphid with >80% body colour becoming pale is collected (Wang et al. 2019).

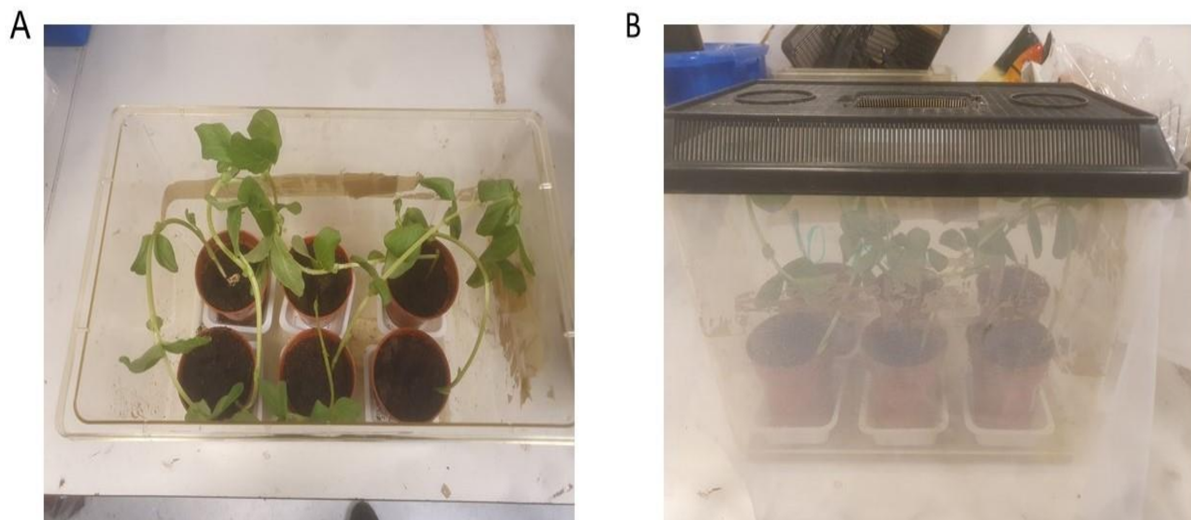


Figure 2.3. Mesocosm container set up for triggering alternative morphs in pea aphids. **A**) from above and **B)** Front view of set up mesocosm cage including plants, superfine mesh and plastic tray for plants.

2.3 Microcosm experiment

2.3.1 Reproductive success of alternative morphs (F0 generation) and the subsequent offspring (F1)

On the last day of the mesocosm experiment, five 4th instar female aphids of each morph (wingless, winged, red, and pale) were transferred to an individual plant and allowed infestation for another 12 days (transfer experiment 1), followed by counting all aphids (total population) on the plants (Figure 2.4). On the 12th day of transfer experiment 1, five 4th instar aphids (F1 generation) were transferred to another individual plant for another 12 days (transfer experiment 2) followed by counting all aphids (total population) on the plants. All aphid samples were collected during the same time of the day (3pm-4pm) to avoid photoperiod effects on the collections. The experiment setup was repeated seven times (n=7 replicates).

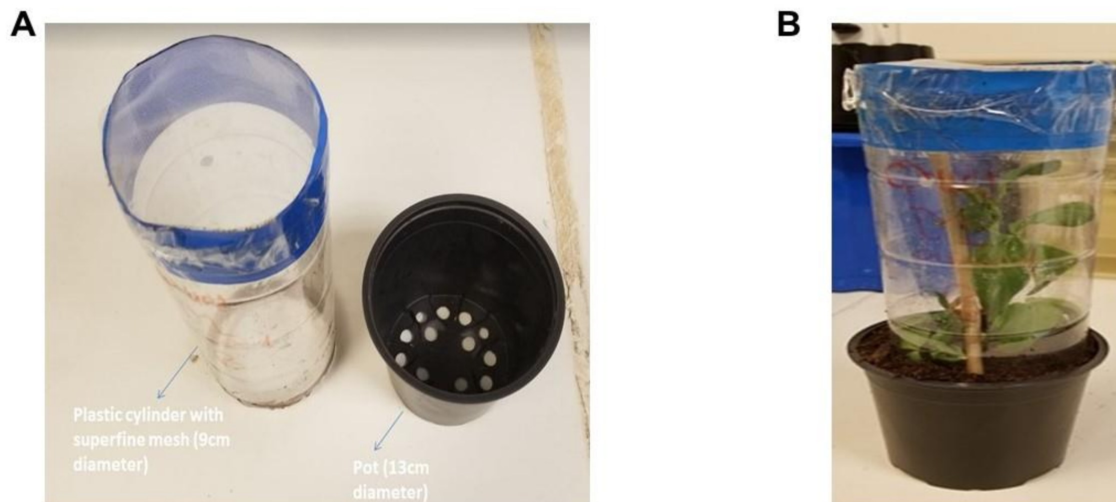


Figure 2.4. Transfer experiment set up. A) Transfer experiment material with a plastic cylinder (9cm diameter) and an open top covered with superfine mesh. **B**) Transfer experiment front view after microcosm set up, a bamboo stick is used to support the plant to prevent the plants from collapsing to the side.

2.4 Molecular analyses

2.4.1 Tissue collection

At the end of the mesocosm experiment, aphids were collected and frozen instantly using liquid nitrogen followed by storing in a -80°C freezer for further analyses.

2.4.2 Quantitative Reverse Transcriptase qPCR

a) RNA extraction

Total RNA from each adult aphid sample (pools of seven individual aphids from each mesocosm) was extracted using the Qiagen RNeasy Kit (Qiagen, UK) according to the manufacturer's instructions, which includes an additional gDNA eliminations step using the RNase-Free DNase kit (Qiagen, UK). Briefly, the frozen aphids were ground into fine powder in liquid nitrogen using a pestle and mortar followed by homogenization in RLT buffer (Qiagen, UK) with 10μL/ml β-mercaptoethanol. gDNA elimination consisted of the addition of DNase I followed by incubation at room temperature (20-30°C) for 15min and centrifugation at 10000 xg. After washing the column, 40μL of RNase-free water was added directly to the column membrane and incubated at room temperature for 1min. RNA was eluted by centrifugation at 10000 xg for 1min. Total RNA samples were transferred into -80°C until further analyses.

b) RNA quantification and purity

RNA purity was assessed using a Nanodrop (Thermo Fisher Scientific, USA). For each biology replicate, 1μL RNA was loaded on the machine and the ratio of absorbance value 260nm/280nm were recorded. Next, the total RNA concentration was measured using the Qubit RNA HS Assay Kits (Qiagen, UK) and Invitrogen Qubit 3.0 Fluorometer (Thermo Fisher, UK) according to the manufacturer's instructions. Briefly, 1μL of sample were added to 199μL of Qubit working solution (RNA HS reagent diluted in 1:200 with RNA HS buffer)

in a thin-walled PCR tube followed by incubation at room temperature for 2min. Total RNA concentration (ng/µL) was calculated using the average of three readings from the Qubit Fluorometer.

c) RNA integrity assessment

RNA integrity was assessed using a non-denaturing RNA agarose gel. A 1.5% agarose gel was prepared in 1x TBE buffer containing 0.01% gel red stain. 4µL of samples were mixed with 1µL of 5x Loading Buffer (Bioline, UK) and loaded into the wells of the gel. Samples were separated by electrophoresis for approximately 1h at 115V until there was a clear separation of the ribosomal RNA (rRNA) bands. Good RNA integrity is signified by presence of two distinctive 28S and 18S rRNA bands with 28S rRNA having twice the fluorescent intensity compared to the 18S band.

d) Complementary DNA (cDNA) synthesis

RNA samples were reversed transcribed into cDNA using QuantiFast Reverse Transcription (Qiagen, UK) according to the manufacturer's instructions which includes an additional gDNA elimination step. 2µg of total RNA were used as in the input and sample volume was then made up to 28µL in RNAase-free water and 4µL of 7x gDNA wipeout buffer. A negative control was included to ensure that there is no contamination: no template control (NTC), where the RNA input is replaced with water; also a no reverse transcriptase (-RT) is included where the reverse transcription enzyme is replaced with (RNase free-H_2O). Samples and negative control were thoroughly mixed in a mini centrifuge and incubated at 42°C for 2min and immediately placed on ice. Next, a final volume reaction volume of 40µL was made up by adding 2µL of Quantiscript Reverse Transcriptase, 8µL of 5x Quantiscript RT buffer and 2µL of RT primer mix into the samples. For the –RT control, the reverse transcriptase enzyme is replaced with water. All the samples were mixed

thoroughly followed by incubation at 42°C for 15min and 95°C for 3min. Samples were quickly mixed using a mini centrifuge and stored at -20°C for further use.

e) Reference gene selection by geNorm analysis

A geNORM analysis was conducted to determine the best reference genes that are stable across different experimental conditions (aphid genotypes, morphs) (Vandesompele et al. 2002). geNORM analysis is commonly used software package to determine most stable (housekeeping genes) from a list of tested candidate reference genes which is used for normalization of qPCR data on the later steps (relative expression analysis). Five genes (SDBH, NADH, 16S, RPL12, 18S, summarised in (Table 2.2) were selected based on a previously published paper by (Yang et al., 2014). All samples were run for each gene on a single 96-well plate to decrease the effect of inter-assay variation on sample reference gene expression. The reaction is prepared using the QuantiFast Reverse Transcription kit (Qiagen, UK). A final reaction volume of 25µL is made up by adding 12.5µL of 2x QuantiFast SYBR Green PCR Master Mix, 5µL of primers, 2.5µL of cDNA template (diluted 1:20) and 5µL of RNase-free water. The sample was mixed vigorously by flicking and picofuged until no bubbles remained. All reactions were run on the AriaMx qPCR machine (Agilent, USA) with the following conditions: activation of HotStar Taq Plus DNA Polymerase with 1 cycle at 95°C for 15min, 40 cycles of cDNA denaturation at 95°C, followed by primers annealing at 59°C for 30s and lastly extension at 60°C for 30s. At the end of the run, a dissociation (melt) curve analysis of the PCR is included with the following conditions: 95°C for 1min, 55°C for 30s and lastly increasing to 95°C in 0.5°C intervals/s. The cycle threshold (Ct) value was automatically calculated by the machine and is defined as the cycle number whereby the reaction generated the fluorescence that crosses the fluorescence threshold. The Ct value obtained from the machine is then entered into the qBase+ software v3.2 (Biogazelle, Belgium), which uses a pair-wise analysis algorithm to determine the most

stable reference genes (Table 2.2). Any genes with the M-value <0.5 is consider stable and is suitable to be used for qPCR data normalization.

Table 2.2. List of pea aphid reference genes used in the geNorm analysis

Gene name	Gene symbol	Gene product function	M-value
16S ribosomal RNA	16S	Initiation of protein synthesis	0.68
succinate dehydrogenase B	SDHB	Plays a role in Krebs cycle and part of the respiratory chain	0.44
NADH dehydrogenase	NADH	Catalyse the transfer of electrons from NADH to ubiquinone	0.35
Ribosomal protein L12	RPL12	Involved in protein synthesis	0.37
18S ribosomal RNA	18S	Activeof protein synthesis in 40S ribosomal subunit	1.040

f) Gene expression analysis using SYBR Green

Candidate genes and references gene initial template values were assayed using QuantiFast SYBR Green Kit with the standard curve constructed using the pooled aphid cDNA. A series of different cDNA dilutions (1:100, 1:50 1:2) were used to establish the most appropriate dilution for qPCR amplification producing Ct values between 15-33. For each qPCR plate, the standard curve was generated using pooled aphid cDNA with an eight-point 2-fold serial dilution with the cDNA input values ranging from 50ng-0.391ng (assuming 100% of the reverse transcription efficiency). A NTC and –RT negative control is included in every plate and all the standard, samples and negatives were run in duplicate.

The two most stable geNORM primers were used as reference genes. Candidate genes for analysis include flightin (fl), Heat shock protein 83 (Hsp83), Acyl-CoA dehydrogenase, (SCAD), vestigial (vg), engrailed (en), apterous (ap1), carotene dehydrogenase (tor), *Acyrthosiphon pisum* densovirus (Apns-1). These genes were selected based on their involvement in stress response, wing development, and metabolic processes as summarised in (Table 2.3). Apart from that, a further eight genes Troponin C (TnC), Krueppel-homolog 1 (Kr-h1), Mothers against decapentaplegic homolog 4 (Mad), Partner of bursicon (Pburs), DNA methyltransferase 1-associated protein 1 (DMAP1), Histone deacetylase Rpd3 (Rpd3), ecdysone 20-monooxygenase (shd), Trehalase (Treh), were selected based on the results from our RNA-seq analysis (Table 2.4). Primers for candidate genes (Thermo Fisher, UK) were reconstituted with RNase-DNase free water according to the manufacturer's instructions. A final reaction volume of 25µL is made up of 12.5µL of 2x QuantiFast SYBR Green PCR Master mix, 5µL of primers, followed by 2.5µL of samples, and the remaining volume with RNase-free H_2O. The initial template values were obtained by interpolating the mean Ct of samples to standard curves followed by normalisation to the reference genes to calculate the relative gene expression. A dissociation (melt) curve was included for each gene to ensure that only a single qPCR product is amplified.

Table 2.3. Candidate genes for qPCR analysis.

Gene name	Gene symbol	Predicted amplicon size (bp)	Gene product function
Flightin	fl	124	Involved in thick filament assembly and sarcomere stability in flight muscle and essential proper flight muscle function
Heat shock protein 83	Hsp83	139	Involved in heat stress, bacterial and fungal infection and involves in aphid lifespan, fecundity and embryogenesis.
Short chain specific acyl-coA dehydrogenase	SCAD	93	Involved glycose biosynthetic process
Vestigial	vg	83	Involved in wing development
Engrailed	en	79	Involved in wing development such as anterior-posterior (A-P) pattern
Apterous	ap1	147	Involved in wing development such as dorsal-ventral (D-V) patterning
carotene dehydrogenase	tor	90	Involved in carotenoid biosynthesis

Table 2.4. Candidate gene selected from Differentially expressed gene (DEG) list that are used for qPCR validation

Gene name	**Gene symbol**	**Amplicon size (bp)**	**Gene product function**
Troponin C	TnC	124	Stretch activation /indirect flight muscle, helps with binding of Ca2+ for flight
Mothers against decapentaplegic homolog 4	Mad4	139	Involved in wing margin morphogenesis (dpp signalling pathway)
Ecdysone 20-monooxygenase isoform X2	Shd	93	Involved in ecdysteroid production
Partner of bursicon	Pburs	83	involved in sclerotization, pigmentation and wing expansion behavior, KO in *Drosophila* shows defects in wing expansion
DNA methyltransferase 1-associated protein 1	DMAP1	79	Involved in DNA methylation (silencing in ladybird causes a reduction in the ovaries numbers)
Krueppel homolog 1 isoform X3	Kr-h1	100	Kr inhibit ecdysone (lower ecdysone= higher winged off springs), Kr downregulated in workers with under-develop bee workers, Kr KO will also reduce fecundity
Histone deacetylase Rpd3	Rpd3	90	Required for resistance against starvation conditions, HDAC knockdown *Drosophila* shows lower survival under starvation

Trehalase	Treh	159	Involved in metabolism during starvation (increase until the end of starvation period)

2.4.3 Pyrosequencing analysis

a) DNA extraction

Total DNA from each adult aphid sample (pools of seven individual aphids from each mesocosm) was extracted using the Qiagen DNeasy Blood & Tissue Kit (Qiagen, UK) according to the manufacturer's instructions, which includes an additional RNA elimination step using the RNase (4mg/mL) (Qiagen, UK). Briefly, the frozen aphids were ground into fine powder in liquid nitrogen using a pestle and mortar followed by homogenization in ALT buffer (Qiagen, UK) with 20µL of proteinase K followed by overnight incubation at 56°C. RNA elimination consisted of the addition of RNase A followed by incubation at room temperature (20-30°C) for 5min. Next, 200µL of buffer AL and ethanol (96-100%) were added to the samples and mix by vortexing followed by centrifugation at 10000 xg for 1min. After washing the column, 100µL of buffer AE was added directly to the column membrane and incubated at room temperature for 1min. DNA was eluted by centrifugation at 10000 xg for 1min. DNA samples were transferred into -80°C until further analyses.

b) DNA quantification and purity

DNA purity was assessed using a Nanodrop, (Thermo Fisher Scientific, USA). For each biological replicate, 1µL DNA was loaded on the machine and the ratio of absorbance value 260nm/280nm was recorded. Next, the total DNA concentration was measured using the Qubit DNA BR Assay Kits (Qiagen, UK) and Invitrogen Qubit 3.0 Fluorometer (Thermo Fisher, UK) according to the manufacturer's instructions. Briefly, 1µL of sample were added to199µL of Qubit working solution (DNA HS reagent diluted in 1:200 with DNA BR buffer) in a thin-walled PCR tube followed by incubation at room temperature for 2min. DNA concentration (ng/µL) was calculated using the average of three readings from the Qubit Fluorometer (Qiagen, UK).

c) DNA integrity assessment

DNA integrity was assessed using a non-denaturing DNA agarose gel. A 1.5% agarose gel was prepared in 1x TBE buffer containing 0.01% gel red stain. 4µL of samples were mixed with 1µL of 5x Loading Buffer (Bioline, UK) and loaded into the wells of the gel. Samples were separated by electrophoresis for approximately 1h at 115V until there was a clear distinctive band. Good DNA integrity is signified by presence of a distinctive band above 10kb using HyperLadder™ 1kb (Bioline, UK).

d) Bisulfite conversion

DNA samples were bisulfite converted using EZ DNA Methylation-Gold Kit (Zymo Research, UK) according to the manufacturer's instructions. 350ng of DNA were used as in the input followed by addition of 130µL of CT conversion reagent and sample volume was then made up to 150µL in RNase-free water. All reactions were run on the 3Primer Thermal Cycler machine (Techne, UK) with the following conditions: 1 cycle at 98°C for 10min, 1 cycle at 64°C for 150min. Next, samples were transferred to a spin column containing 600µL of binding buffer followed by centrifugation at 12500 xg for 30s. After washing the column, 200µL of M-Desulphonation Buffer were added to the column and incubated at room temperature (20-30°C) followed by centrifugation at 12500 xg for 30s. The spin column was washed again one more time and 10µL of M-Elution Buffer were added directly to the column. DNA was eluted by centrifugation at 12500 xg for 30s. DNA samples were transferred into -80°C until further analyses.

Table 2.5. Candidate gene used for pyrosequencing analysis.

Gene	Forward Primer	Reverse primer	Sequencing primer
Heat shock protein 83	GTTATTATGAAGGTAATTGATGATGTTGA G	CCAAATTCCAATTTAAATAAATCAAT CAAC (biotin)	TTTTATTTTAATAAAGAAATTT T
fligthin	GATTTTGATGTATTTGTTTGGATGTT	ACAACCAACTTATAAATTTTCCTCCA TTC (biotin)	GTTGATGTAGGTGGTGAT
vestigial	AGATGGATAATAATTGTAGATTATATAGAT	AGATGGATAATAATTGTAGATTATAT AGAT (biotin)	ATAATTGTAGATTATATAGAT TTAT
engrailed	TAGTTGGTTAGGTTTAAGAAAGAGTTTAT	CAAACCCTAAACCATCAACTACAAAA (biotin)	GAGAATAGGTATTTGATTGAG AGA
apterous1	GTGTGTTTGGAGGTGTTGATTAAG	ACTATTACTATTACTAAACCTACAAC (biotin)	TTTGTTAGTATTATTATTAGTA GTT
Short chain specific acyl-coA dehydrogenase	AATTAGGAAGGAAAAGGGATAGTA	ATTTAATTTTTAATTCCAATAATAACC TAT (biotin)	AAGGAAAAGGGATAGTATT
Insulin-like receptor	GTTGTTTAAGTGATAGTTGTTTTGGTTTG	ACTATACATAAAAAATCAAACTTCCA TACA (biotin)	TTTAAAAGTGAATTTAAAATAT

carotene dehydrogenase	GGGGGAGGATATTAAAAATTATATAGAG	AATCCCTCTTAAAATACTTACTCACT TTAT (biotin)	GTGTTTTATTAATTATTTTGTA TAT
Spalt major	TGGTAGAGTAATGATGGTTAGTTTA	AAAATAACAAATCCTACCATAAACC (biotin)	GATGTTTTTTTATAATAGTAG ATA
Troponin C	TGTATAGTATTAAGTTTTGAGAAAGGGTTT	ATAACATAACCCAACATATCCAAAAT (biotin)	GATATTGAGAAAAATGGTAAA AT
Mothers against decapentaplegic homolog 4	TTGGTTAAATAATTAGTTAAAGAGATGAGT	TCCAAATACATCACTATTCATATTAA AACT (biotin)	AGTTAAAGAGATGAGTATATA GT
Ecdysone 20-monooxygenase	GGTATGGTGTAAAAAAATTTTTGGATTAA	ACTTCAACACCAATTTTACCAAAATT ATAA (biotin)	GGTGTAAAAAAATTTTTGGAT TAAA
Partner of bursicon	TGAAAAGTTTTTGAGGGAGAGAATGA	ACCAATTTATCACTAATCAATCTTAC (biotin)	TGAGGGAGAGAATGAT
Krueppel homolog 1	AAGTTGAATTTAGTATGGATGTGTTTTA (biotin)	ACCAATAACTCCCTACCTAAAAACAA TT	ATAACTCCCTACCTAAAAACA ATTT
Histone deacetylase Rpd3	AGGAAATTATTATTATGGATAAGGTTATT	ATCAAACTTAACTTACATATATTTCC ATTT (biotin)	AAGGTTATTTAATGAAATTAT ATAG

Trehalase	AGTATATGTGAGGGAAAGTGTG	AAAACCAAATAATAACCTACTAAACC (biotin)	GGAAAAGGAGGTTTTGTTG
DNA methyltransferase 1-associated protein 1	AGGTTTTATTTATAAGTTTTAAAAAGTTAG	TTAACAAATAAATACTCTTTACCCTCATC (biotin)	GTTTTAAAAAGTTAGATTGGG
Phytoene desaturase	GGAGTGGTAAAAGGAATAAAATTGTAG	AATCCTCCAATATTTAATCAAAACTTAACT (biotin)	AAAAATTAGGTAAAAAGAAATTGA

e) Pyrosequencing analysis

The PCR product that has been successfully amplified is then used for methylation analysis by pyrosequencing. First, the master mix containing sepharose beads and binding buffer are prepared (Composition: 26.4µL of sepharose beads, 1056µL of binding buffer, 766µL of ddH_20). 10–15µL of the PCR product is mixed with 70µL of the master mix in a 24-well PCR plate. The plate is then shaken for 10–15min at 1400rpm. Next, 0.3µM of sequencing primer for each genes (Table 2.5) mix is prepared and 25µL of the mix is added to each well of a Q24 sequencing plate. The beads are subsequently collected using a vacuum preparation workstation and then sequencing primer is added and heated to 80°C for 2min. Sequencing primer that has annealed to the biotinylated DNA strand is cooled to room temperature. Samples are pyrosequenced using the PyroMarkQ24 machine with PyroMark Q24 Software 2.0 (Qiagen, UK) and CpG methylation is analysed using the PyroMark CpG software.

2.5 Western blot analysis

a) Protein lysate extraction

Protein solutions were obtained by grinding whole aphid sample in a 1.5mL Eppendorf tube with a pestle and mortar. A pooling for seven aphids was used in each extraction with a total of three biological repeats. The protein solution was then sonicated and centrifuged at 10000 xg to remove any debris.

b) Protein loading quantification

Protein lysate quantification was assessed using the Pierce BCA Protein Assay Kit (ThermoFisher Scientific, Massachusetts, USA), according to the manufacturer's instructions. A standard curve was generated using albumin standard with concentrations ranging from 0-2000 (µg/mL). 25µL of each standard and samples were loaded into each

well in a 96-well plate, followed by addition of 200µL of working reagent onto each well and mixed well by placing on a plate shaker for 30s. Next, the plate was incubated for 30min at 37ºC and absorbance at 595nm was read using a Synergy H1 Multi-Mode Microplate Reader (Biotek, USA).

c) Western blot

Samples were separated by SDS-PAGE (45min, 200V), (4-12% Bis-Tris gels, Thermo Fisher Scientific) and transferred to nitrocellulose membrane (10h, 30V, 4°C) (Whatman, Maidstone, UK). Membranes were blocked for 1h at room temperature using casein-blocking buffer (Sigma-Aldrich, Missouri, USA, 1:10 dilution) in TBST (10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.05% [wt/vol] Tween-20). Membranes were incubated with HSP70 3A3 monoclonal mouse primary ab (Thermo Fisher, 1:1000 dilution) on blocking buffer overnight at 4°C. Subsequently, membranes were washed three times for 5min per wash using TBST and then incubated with polyclonal goat anti-mouse secondary Alexa Fluor 680 (Thermo Fisher Scientific, Massachusetts, USA, 1:10000 dilution) on blocking buffer in the dark at room temperature for 2h. Membranes were washed again three times for 5min per wash using TBST and then scanned using Odyssey Infrared Imaging System (LI-COR Biosciences, Nebraska, USA) at 700nm. Band signal intensity was determined based on the median pixel intensity signal of the bands using Odyssey software (LI-COR Biosciences, Nebraska, USA).

2.6 Statistics

All statistics were performed using SPSS v23 (IBM, USA). For the phenotype data (chapter 3), general linear models (GLMs) were used to investigate predictors of alternative morph production in aphids and general linear mixed models (GLMMs) were used investigate predictors of reproductive success of the morphs and subsequent offspring. All models include morph as fixed factor with other relevant fixed effect predictors (e.g., genotype) included where appropriate. For transfer experiments 1 and 2, Parental ID or Grandparental ID were included as covariates where appropriate. For both qPCR (chapter 4) and pyrosequencing analyses (chapter 5), morphs type and genotype, alternative morph vs wild type (AM_WT) were used as fixed factors. Subsequently, a post-hoc Tukey HSD test was applied to examine multiple pair-wise comparisons. All other statistical analysis for RNA-seq, MBD-seq and miRNA-seq will be discussed in the method sections for each chapter, respectively. The F statistic is rounded to two decimal places for all analyses. In all the analyses, the minimal number of significant numbers in a model is analysed used using the exploratory factor analysis (EFA).

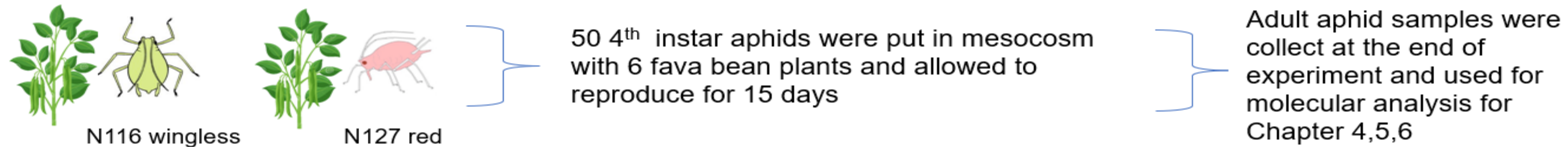
Data are presented with mean ± standard error of the mean (SEM) with errors bars shown in black unless otherwise stated. The aphid genotype N116 is presented in green bars while the N127 genotype is presented in red bars unless otherwise stated. All graphs were produced using GraphPad Prism v.9.4.1 (GraphPad, USA) apart from the results from RNA-seq, MBD-seq and miRNA-seq.

2.7 Summary

The experimental design, timeline and sample collection for phenotype and molecular experiment presented in Chapter 3-6 of the thesis are shown in Figure 2.5.

Experimental design

A) Crowded conditions



B)Transfer experiment 1

B)Transfer experiment 2

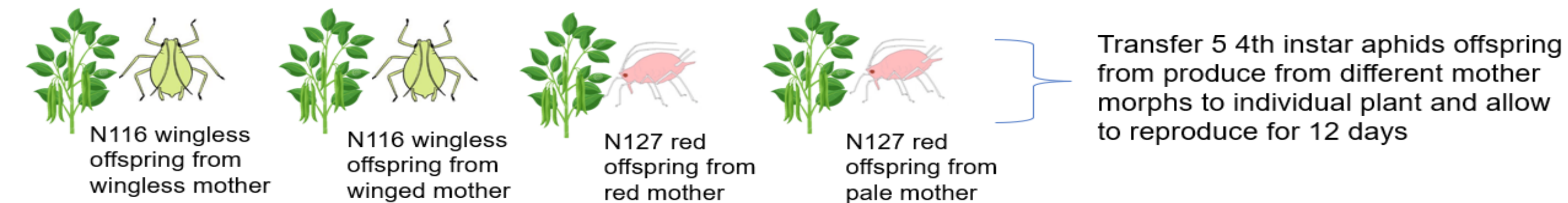


Figure 2.5. Experimental design for crowded conditions and transfer experiments.

Chapter 3: Response of pea aphid genotypes to environmental stress and the trade-offs between reproduction and dispersal morphs

3.1 Introduction

Phenotypic plasticity is the ability of organisms to change their phenotypes when exposed to different environments. Virtually all organisms will exhibit some degree of plasticity (West-Eberhard 2003). One of the most extreme cases of phenotypic plasticity is known as polyphenism. Polyphenism is the phenomenon whereby a single genotype gives rise to multiple discrete phenotypes. Polyphenisms play a key role in the success of insects as they allowed insects to adopt different phenotypes in response to environmental changes without altering their genome. There are many types of polyphenism in insects, for example, seasonal polyphenism in the southern African butterfly *Bicylus anayana* where several prominent marginal eyespots usually appear under the surface of the butterfly's hind winds during the wet season. In contrast, the number of spots will usually decrease and the butterfly changes its body colour to brown during the dry months. The reason behind the phenotypes changes during the dry season is to help them camouflage to the environment and reduced the chances of being preyed upon by predators (Lyytinen et al. 2004). Another widely studied polyphenism observed in insects is known as phase polyphenism. Phase polyphenism is usually observed in a specific species of grasshopper known as locusts such as the migratory locust *Locusta migatoria*. Locusts will respond to the population density around them by shifting between the low-density ‘solitarious’ phase or the high-density, swarm-forming ‘gregarious phase’ (Pener et al. 2009). The changes in this phase usually involved a range of transitions in the locust morphology such as body shape, body colour, metabolic function and reproduction success.

Another example of polyphenism that is commonly observed in insects is known as wing polyphenism. Wing polyphenism is most widely studied in aphids, a type of soft-bodied insects that commonly feed on the sap of plants. Aphids possess a remarkable complexity in their life cycles. In aphids, there are usually two different life cycles known as anholocyclic and holocyclic (Agrawal 2007). In holocyclic cycle (usually during wintertime), aphids

produce sexually with a single sexual generation, producing an overwintering egg that will hatch when conditions become more favourable (Figure 3.1A). In contrast, aphids reproduce asexually in the anholocyclic life cycle (mostly summertime) with only wingless female offspring being produced. During the anholocyclic life cycle, the aphids exhibit wing polyphenism when exposed to environmental stress (crowding, predation) whereby the mother will transmit the environmental signals to the offspring resulting in winged offspring in the next generation (Figure 3.2B, Vellichirammal et al. 2017).

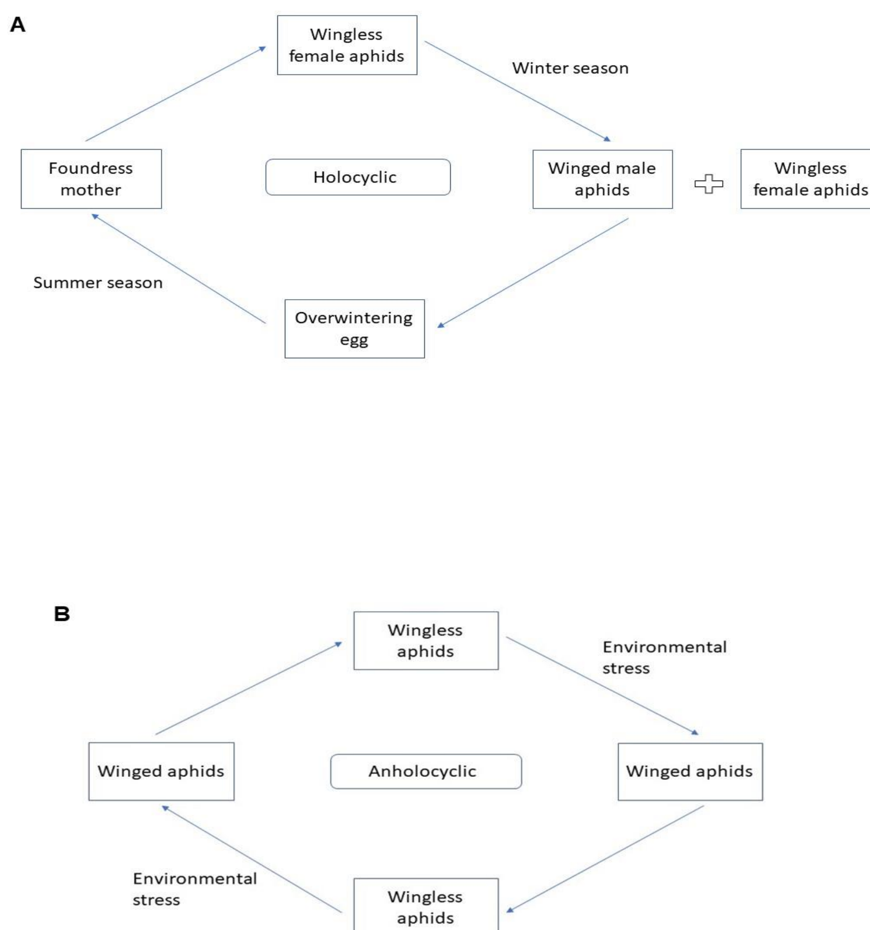


Figure 3.1. Aphid life cycle. A) Holycyclic life cycle consisting of single generation of male aphids and overwintering eggs. **B**) Anholocyclic life cycle consisting of only female aphids and alteration between winged and wingless offspring depending on environmental conditions.

Even though wing polyphenism is common when aphids are exposed to environmental stress, it is important to bear in mind that the aphid response towards environmental stress varies across genotypes with some aphids unable to produce wings but rather changing their body colour morph in response to environmental stress. This made aphids a good model to study the interaction between genotype and environment (Braendle et al. 2005).

The ability to fly in insects has been a major contribution to the evolutionary and ecological success of aphids (Dingle 1996). However, according to the life history theory prediction, when two traits are energetically costly, organisms can only invest energy into one of the traits resulting in a trade-off between the two traits (Zera and Harshman 2001). One of the major trade-offs in winged aphids is observed between the ability to disperse and reproduction. Winged aphids have lower reproductive success in comparison to wingless aphids. The trade-offs in dispersal and reproduction are mainly due to the allocation of energy to ovarian growth versus the maintenance of flight muscle/wing or storage of fuel for flight (Zera and Denno 1997). Most winged morphs possess large functional flight muscles and require large quantities of lipid as flight fuel. In contrast, most flightless morphs have non-functional flight muscles with reduced lipid storage. The differences observed between reproduction and flight are generally large which makes them a good experiment model in investigating physiological trade-offs. For example, Zhang et al. 2008 measured the flight distance in soybean aphid and found a significant decrease in reproduction as the flight distance increases.

In our laboratory, we maintained two different genotypes of pea aphids (Kanvil et al. 2014) with differences in their body colour. The genotype N116 has green body colour while the N127 has red body colour. These two genotypes also differ in other key traits as described in section 2.1.

In this chapter, I investigated the differences in response to environmental (crowding/starvation) between two aphid genotypes. The degree of alternative morph

production and the phenotypes (N116 winged morphs and N127 pale morphs) were measured. Next, we investigated the reproduction rate in different aphid morphs to provide insights into trade-offs. Most trade-offs have only been investigated in one generation; thus to provide insight into the trans-generational effect of trade-offs we investigated the reproductive success of subsequent offspring.

The aims of this chapter were to:

- Investigate the degree of alternative morphs production between the two pea aphid genotypes.
- Investigate the possibility of trade-offs between dispersal and non-dispersal aphid morphs in two genotypes (N116 winged vs N116 wingless and N127 pale vs N127 red).
- Investigate transgenerational trade-offs in pea aphids

3.2 Methods

3.2.1 Aphid colony

Two aphid genotypes N116 and N127 are used in the experiment. The genotypes are kept and raised in separate individual colonies with two colonies for each genotype. The details and differences in traits between these two polymorphic aphid genotypes are described in section 2.1.

3.2.2 Experiment plant maintenance

Faba bean plants (*Vicia faba minor*) were used in the experiment. All plants used in the experiment are three weeks old after germination and details about the plant maintenance and growth is described in Section 2.1.3 and 2.1.4.

3.2.3 Crowded conditions for triggering of alternative morphs

A pilot study conducted in our lab showed that predation failed to trigger alternative morphs in the pea aphids (no data shown). In the predation experiment, we put a ladybird *Adalia bipunctata* on a plant with 10 fourth instar aphid. However, no winged offspring were observed in the next generation. Apart from that, the changes of body colour to pale were not observed in the predation experiment either. Therefore, we conducted a pilot study under crowded conditions and showed that alternative morphs are successfully produced from both genotypes. All aphids used in the experiment are of adult stage and are obtained from the stock colony maintained under low density conditions (15 aphids per plant maximum). The details of the experimental setup and traits measured are described in section 2.2.1.

3.2.4 Reproductive success of alternative morphs (F0 generation) and their subsequent offspring (F1 generation)

At the end of the crowded condition experiment, five 4th star instar aphids from each morph were used in the crowded conditions as described in section 2.3.1. A second transfer experiment is carried out using five 4th star instar aphid offspring produced from each morph as described in section 2.3.1.

3.2.5 Statistics

All statistics were performed using SPSS v23 (IBM, USA) as described in section 2.6. Data are presented as mean ± SEM, with errors bars shown in black unless otherwise stated. The two genotype N116 and N127 are presented as green and red bars, respectively. All graphs were produced using GraphPad Prism v9.1.2 (GraphPad, USA).

GLMs were used for degree of alternative morphs production and total population in the crowded conditions experiments. GLMMs were used for the total population size in both transfer experiment 1 and 2 with either parental ID or grandparental ID as the random effect to account for the variation of reproduction due to the difference in population size the aphid were collected from. For all experiments, fixed factors (morphs, genotype, alternative morph vs wild type) were included in the models where relevant (section 2.6).

3.3 Results

3.3.1 Morph production in different pea aphid genotypes

The two pea aphid genotypes N116 and N127 are able to produce alternative morphs under crowded conditions, with N116 producing winged offspring and N127 changing their body colour from red to pale but rarely producing winged morphs (where no winged morphs were observed during the course of experimental and only one were observed over the period across all experiment in the laboratory. A stepwise regression model (including all the independent variable and start removing those that are not statistically significant) is used to analyse the percentage of morph production in two different pea aphid genotypes. We first analysed the proportion of alternative morphs after 14 days of infestation, which is highly dependent on the genotype with N127 producing a greater percentage of alternative morphs compared to N116 (GLM, $F_{(1,12)}$=74.29, P<0.001; Figure 3.2A).

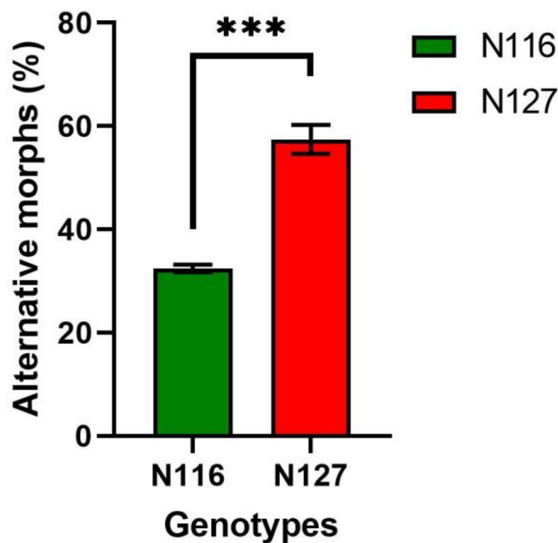


Figure 3.2. Percentage of morph production in two different pea aphid morph. A) Bar represents standard error of means. Green=pea aphid N116, Red=pea aphid N127. Values represent the mean ± SEM of seven replicates. **P-value <0.01, ***P-value <0.001.

Although total population size does not significantly affect the percentage of morphs when the genotype is also a predictor in the model, total population as the only main effect also has a significant effect (GLM, $F_{(1,12)}$=5.930, P=0.031; Figure 3.2B). This is because N116 has a higher population size than N127 (GLM, $F_{(1,12)}$=5.054, P=0.044; Figure 3.3).

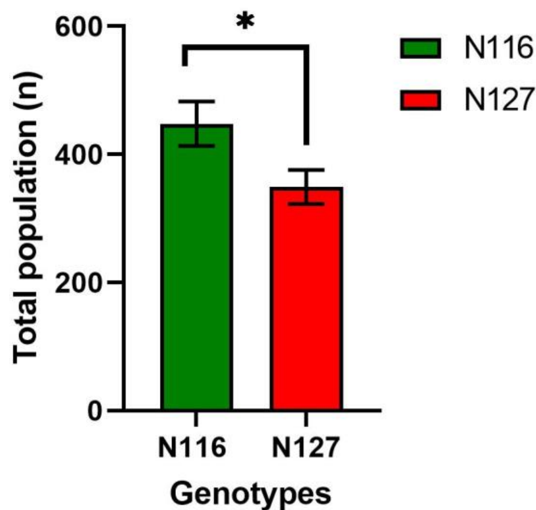


Figure 3.3. Total population size in two different pea aphid morphs and the interaction effect of genotype and population on the percentage of morphs. **A)** Bar represents standard error of means. Green=pea aphid N116, Red=pea aphid N127. Values represent the mean ± SEM of seven replicates. **P-value <0.01, ***P-value <0.001.

3.3.2 Total population size of different pea aphid genotypes

Apart from the percentage of morphs, the total population size of pea aphids was also measured at the end of the experiment. The total population size of aphids was measured by counting all aphids in the mesocosm. A stepwise regression model is then used to analyse the total population size in two different pea aphid genotypes. The total population size of aphids is dependent on the genotype with N116 having a higher population size compared to N127(GLM, $F_{(1,12)}$=5.054, P=0.044; Figure 3.3). Although the percentage of morphs does not significantly affect the total population size when genotype is also a predictor in the model, percentage of morphs as the only main effect also has a significant

effect (GLM, $F_{(1,12)}$=5.930, P=0.031; Figure 3.4). Further, we investigated the effects of three plant traits (plant shoot root ratio, plant dry weight, and plant survival rate) on aphids' total population. None of these factors, either alone or as interaction with aphid genotype and population size, had a significant effect (GLM, $F_{(1,12)}$=1.858, P=0.198).

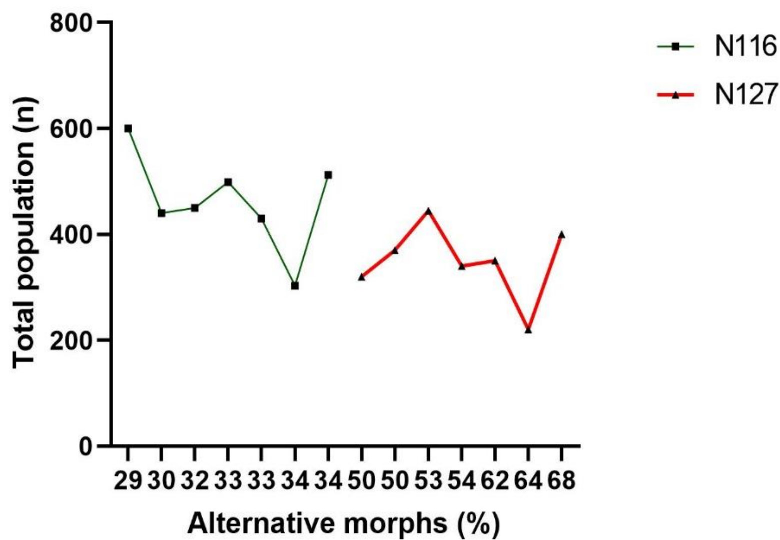


Figure 3.4. Effect of percentage of morphs on aphid total population size. Green line=pea aphid N116, Red line=pea aphid N127. Each point circle for N116 and square for N127 represents a data point.

3.3.3 Plant fitness

Apart from aphid traits, plant traits are also measured to determine the effect of aphid genotypes on plants' fitness since the two aphid genotypes have very different virulence level and come from different host plant (N116 from *Medicago Sativa* and N127 from *Medicago lupulina*). Plant fitness is approximated by the shoot-to-root ratio and is measured by dividing the dry weight of the plant shoot by the dry weight of plant root (Eller and Brix 2012). A stepwise regression model is then used to analyse the plant shoot-to-root ratio (S:R) in two different genotypes. Aphid genotypes, total population size and degree of alternative morphs had no effect on S:R ratio (GLM, $F_{(1,12)}$=1.516, P=0.242, $F_{(1,12)}$=2.456, P=0.143, $F_{(1,12)}$=1.858, P=0.198). The S:R ratio in plants is dependent on the survival rate

of the plant whereby higher plant survival generally results in higher S:R in plant (GLM, $F_{(1,12)}=4.875$, $P=0.03$, Figure 3.5)

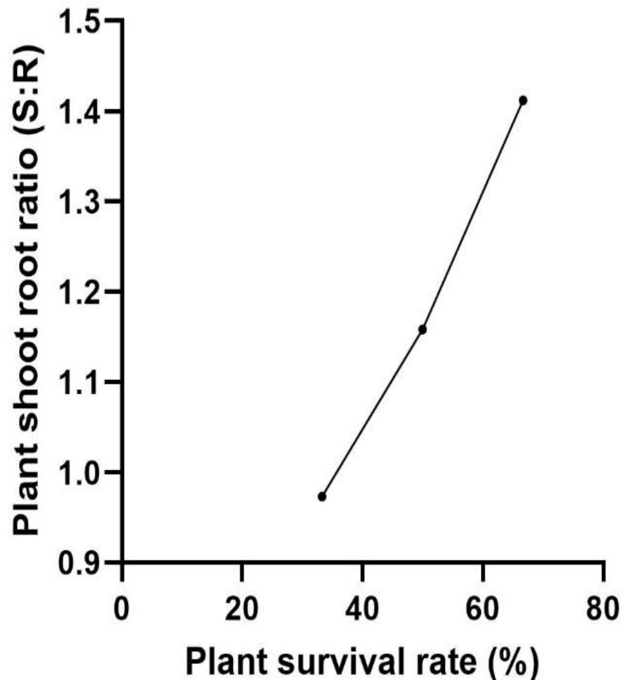


Figure 3.5. Effect of plant survival rate on plant S:R ratio. Black line= plant survival rate

3.3.4 Reproductive success of different pea aphid morphs (T1)

The ability of aphid to disperse has contributed to the major success of aphid evolution as it allows them to escape stressful environmental conditions. According to life-history trade-offs theory, when organisms invest energy in a trait that increases its fitness always comes with a decrease in energy investment in another trait that reduces its fitness. Therefore, a transfer experiment was carried out at the end of the mesocosm experiment to determine the reproductive output (measured by total population after 14 days) in the four pea aphid morphs (N116 winged, N116 wingless, N127 pale and N127 red) to provide insight into possibility of trade-offs between reproduction and ability to disperse. The total population of aphids was measured by counting all aphids in the mesocosm after 12 days of infestation. A mixed model with Parental ID as random effect was used to determine if the population

size from which the aphid morphs come from has any effect on their reproductive output. The reproductive success of aphids are dependent on the interaction of both genotype and alternative morph vs wild type (AM_WT) (GLMM, $F_{(1,12)}$=20.766, P=0.001, Figure 3.6, Table 1) with wild-type morphs (wingless in N116 and red in N127) having a higher reproductive success. Apart from that, the N116 genotype also had a high reduction in reproductive success (measured by total population of wild type / total population of dispersal morphs) (43.4%) compared to the N127 genotype (24.1%). The total population size of the mesocosm from which the aphids originated did not have any effect on the reproductive success of the morphs.

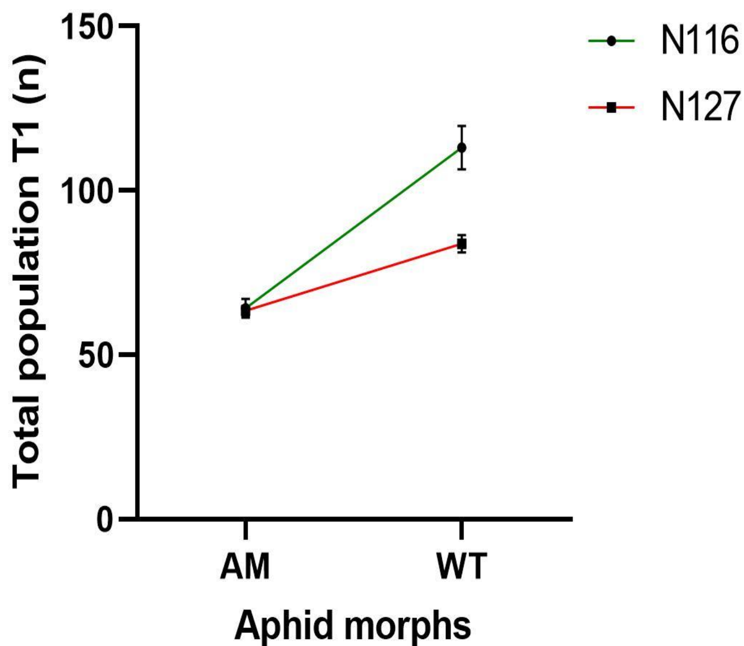


Figure 3.6. Differences in reproduction between offspring from dispersal and non-dispersal mothers in two genotypes and the interactions between alternative morphs vs wild type and genotype. AM= alternative morphs (dispersal), WT=wild type (non-dispersal), green line with circle symbol=N116, red line with square symbol= N127. Values represent the mean ± SEM of seven replicates.

Table 3.1. Summary of statistical model used for the transfer experiment 1 and 2 analysis and the significant predictors for each gene and random effects in each model.

Experiment	Model	Predictors	Random effect	F-value	P-value
Reproductive success of different pea aphid morphs	GLMM	Interaction of morphs and genotype 1AM_WT	Parental ID	24.277	<0.001
Reproductive success of offspring (F1) from different aphid morphs	GLMM	Interaction of morphs and genotype 1AM_WT	GrandParental ID	32.930	<0.001

3.3.5 Reproductive success of offspring (F1) produced from different aphid morphs

A second transfer experiment was carried out at the end of transfer experiment 1 to determine the trans-generational reproductive success (measured by total population) of each of the four pea aphid morphs. The total population of aphids was measured by counting all aphids in the mesocosm after 12 days of infestation. A mixed model with parental ID or grandparent ID as random effect was used to analyse if the total population size the four different pea aphid morphs originated from had any effect on reproductive success. The reproductive success of the F1 generation was dependent on both the genotype and morph type of the mother and their interaction, with wild-type morphs (wingless in N116 and red in N127) having a higher reproductive success compared to N127 (GLMM, F=184.267, P<0.001, Figure 3.7, Table 1). The total population size of the mesocosm from which the aphid originated did not have any effect on the reproductive success of the morphs.

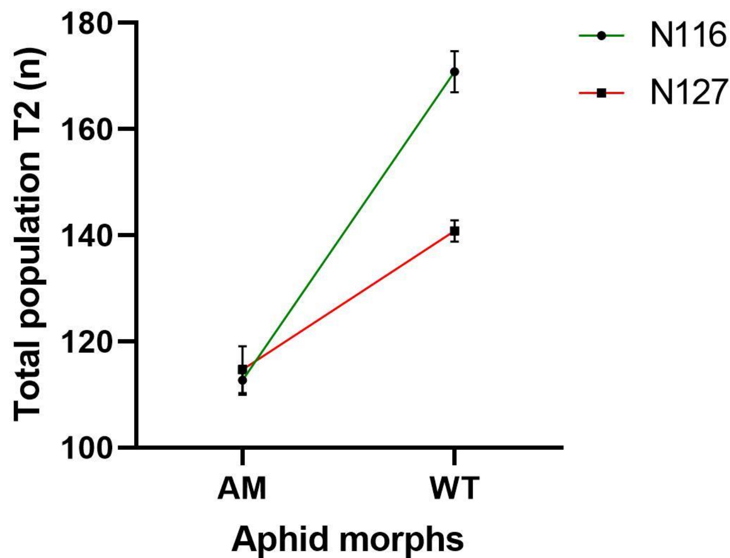


Figure 3.7. Differences in reproduction between offspring (F1) from dispersal and non-dispersal mothers in two genotypes. AM= alternative morphs (dispersal), WT=wild type (non-dispersal), green line with circle symbol=N116, red line with square symbol= N127. Values represent the mean ± SEM of seven replicates.

3.4 Discussion

In this chapter, the first aim was to investigate the degree of alternative morph production between the two pea aphid genotypes. Our results showed that the N127 red genotype produced a higher degree of alternative morphs in comparison to the N116 genotype. The second aim was to investigate the possibility of trade-offs between the dispersal and non-dispersal morphs. Our results showed a lower reproduction output between the dispersal morphs across two genotypes (N116 winged vs N116 wingless and N127 pale vs N127 red). However, due to some limitation of the study such as not measuring the actual metabolites level in aphids or life time reproduction in the aphid morphs (further details discussed in Chapter 7) we can only suggest a possibility of trade-off but not a definite conclusion. Lastly, I sought to investigate the possibility of transgenerational trade-offs in pea aphids. Our results suggest that the reproductive output of the F1 generation (offspring

from different mother morphs) were also reduced. However, due to the same limitation in the transfer experiment we can only suggest that possibility of trade-off lasting more than one generation.

3.4.1 Aphid genotype response to environmental stimulus and the degree of alternative morphs production

This chapter investigated the response of two different pea aphid genotypes to environmental stress (crowding). Organisms usually react to environmental stimulus by altering some of their traits. Sometimes exposure of different genotypes to the same environment stress will result in different phenotypes suggesting that there might a role of genetic variation in these traits. In our study, we found that the aphid genotype N116 genotype and N127 react differently when exposed to environmental stress. N116 usually produced winged offspring while N127 rarely produced any winged offspring but changed their body colour from red to pale. Previous studies have described differences in herbivore reactions to predation stress, for example the production of winged offspring in aphids is dependent on genotype (Braendle et al. 2005) with some genotypes able to produce winged offspring while others were unable to produce wings. One explanation for the differences in the response of genotype to crowding is due to population growth rate differences, whereby fast developing populations usually crowd faster in comparison to slower developing populations (Grantham et al. 2016). Another explanation for the different responses may be due to differences in virulence between the two aphid genotypes. N116 is more virulent than N127 which means they have a higher resistance (measured by survival fecundity and growth rate) on different host genotypes (Kanvil et al. 2014). In our experiment, both N116 and N127 were raised on *Vicia faba*, which is a common host plant for both genotypes. N127 is one of the least virulent genotypes, which means they might have lower developmental success on different plant genotype (Kanvil et al. 2014).

Together with the crowding conditions, it might be more beneficial for N127 to respond to environmental stress by changing body colour instead of producing winged offspring (changing body colour does not require trans-generational signal) to maximize their chance of survival (Grantham et al. 2016).

Another possible explanation for the difference in response to the same environmental stimulus is the presence of different densoviruses and endosymbiont strains in pea aphids. In crowded conditions, Parker et al. (2019) reported that aphid genotypes that produced lower winged offspring has lower expression of the densovirus. Next by, knocking out the densovirus gene in highly wing inducible aphid genotypes leads to a lower percentage of winged offspring production. Apart from densovirus, aphids usually harbour the primary endosymbiont on their gut. Some recent study shows that *Buchnera* plays a role in mediating aphid polyphenism. For example, Zhang et al. (2015) showed that aphids with *Buchnera* eliminated produced a lower proportion of winged offspring in comparison to those with harbour *Buchnera*. The reduction of winged aphid production in aphid could be due to the role of *Buchnera* in regulating nutrients in aphids. Further, Hardie and Leckstein (2007) also showed that the deletion of a few essential amino acids such as isoleucine, and histidine, greatly impacts the number of winged offspring production in aphids. Further, the aphid N127 also has lower *Buchnera* density compared to N116. The results together suggest that instead of a direct effect of *Buchnera* density on aphid wing polyphenism, it is more likely that the lower *Buchnera* density results in a reduction in the essential amino acid that aphids could obtain and therefore affects the proportion of winged offspring produced.

A further possible explanation for the difference in winged plasticity is that the wing trait is considered a bet-hedging trait. The difference between bet-hedging and adaptive phenotypic plasticity is that in bet-hedging traits there are no predictive cues whereby different phenotypes are always produced (Seger and Brockmann 1987). Simons (2011)

suggested a number of characteristics for a trait to be categorised as bet-hedging. First is the identification of bet-hedging traits, which in our study will be the production of winged and wingless offspring. The second is the identification of unpredictable environmental factors. Winged offspring are usually induced by crowding, starvation and interspecific interactions. However, aphids do not fully expand their wing until they reached the adult stage which suggests that winged offspring are not capable of flight until they reach adulthood. The time lag between the environmental signal and the ability to fly suggests that the environment is not 100% predictive of the future whereby the environment stress could change such that a predator can leave or aphid could walk to nearby plants with better quality. The next evidence is that there is genotype-level candidate bet-hedging. By using multiple genotypes of aphids, Grantham et al. (2016) showed that there is significant variability in winged production between aphid genotypes suggesting that there is some genetic variability in winged polyphenism in aphids. These three levels of evidence have shown to be sufficient to categorise a trait as bet-hedging in other systems (Bradford et al. 1993). The other remaining category of evidence for bet hedging required further studies and includes demonstrating variable fitness consequences of the trait, indication of the advantage of the bet-hedging trait under fluctuation selection, and, lastly, quantifying the fitness of bet-hedging trait to the degree of fluctuating selection.

Next, I measured the degree of alternative morph production in the two aphid genotypes. Our results show that the red genotype N127 produced a greater proportion of alternative morphs compared to the green genotype N116. A study by Wang et al. (2019) has shown that it only takes 10h of starvation for the red pea aphid to start changing its body colour. In contrast, winged production in aphids is usually trans-generation in that the mother experiences the environmental signal but it is her offspring that develop into winged morphs (Mehrparvar et al. 2013). The ability to change its body colour rapidly (10h starvation is enough to start triggering body colour change) and without the need for trans-generational

signal transmission to offspring could explain the higher alternative morph production in N127. Migration or dispersal is an essential part of the life cycle of aphids as they need to escape from various environmental changes such as decreasing plant quality, crowding, and the presence of natural enemies to maximize their reproductive success (Weisser and Stadler 1994; Weisser et al.1999). Since N127 is able to react to environmental stimuli more quickly by changing body colour with increased locomotion compared to N116, it might increase their chance of survival (e.g. dispersal more quickly to look for food resources) when exposed to this abiotic stress. Another possible explanation for the difference in the proportion of alternative morph production is that N116 might be a genotype with lower production of winged offspring since studies have shown that variability of winged production in aphids is dependent on genotype possibly due to the presents of different expression level of densovirus (Parker et al. 2019). In contrast, in N127 the change in body colour is reversible in any stage, therefore it might be beneficial to quickly change a higher proportion of the population to pale with increased locomotion so they can quickly disperse to another plant.

3.4.2 Reproductive success in different pea aphid morphs

In this chapter, I found that the dispersal morph in both genotypes has significantly reduced reproduction (as measured by total population) compared to the non-dispersal (wild type) suggesting that a trade-off might present between dispersal and reproduction. Our results are in line with the prediction of life history trade-off theory whereby an increase in one life history trait (improving fitness) is usually accompanied by a reduction in another life history trait (reducing fitness). Some of the main evidence that supports the trade-off hypothesis in wing polymorphic species is that there is a negative correlation between flight muscle mass and ovarian mass (Roff 1986). For example, Lorenz (2007) showed that in the two-spotted cricket, *Gryllus bimaculatus*has, the flight activity peaked at the second day after adult wing

formation accompanied by increase flight muscle mass. After flight, the flight muscle usually undergoes histolyse resulting in an increase in oocyte development. This result suggests a competition in allocating nutrients to the maintenance of the flight apparatus and egg production in insects. Another study by Zhang et al. (2008) using soybean aphid, *Aphis glycines,* also further supports the presence of trade-offs between flight ability and reproduction. By using a flight mill system (a tool commonly used in measure different aspect of insect flight such as speed, distance), Zhang showed a positive correlation between the amount of flight and reduction in reproduction with longer flight resulting in lower reproduction. I also found a difference in reproductive success between the alternative morphs in N127 with the pale morphs having lower reproduction compared to the red morphs suggesting that trade-offs might be present between increased locomotion and reproduction. Instead of producing winged offspring, N127 usually changes its body colour from red to pale (with increased locomotion) (Tabadkani et al. 2013). Our results suggest a possibility of trade-off in aphids between increased locomotion and reproduction as most trade-off studies on aphids usually focus on flight and reproduction. It is possible that the trade-off observed between increased locomotion and reproduction is due to the energy allocation as observed in those with flight capability (Tabadkani et al. 2013). For example, Tabadkani showed that pale aphids have a lower content of lipids and carbohydrates in comparison to the red aphid. This suggests that the trade-off in reproduction and increased locomotion could be due to the utilization of a higher amount of energy reserve to stimulate the aphid to walk more actively in search of a better host quality plant. Another study by Samietz and Köhler (2012) investigated the effect of mobility and female fitness (egg-laying) in a non-migratory, wing monomorphic grasshopper, *Stenobothurs lineatus,* and reported a negative relationship between daily activity radius with egg-laying number. Together with the studies above, our data is in agreement with the

fact that limited internal energy resources are the main explanation for such trade-offs (Harshman and Zera 2006).

Apart from lower reproduction in dispersal morph in both genotypes, our results showed that the reduction of reproductive success in the N116 genotype (43.4%) is significantly higher compared to the N127 genotype (24.1%). One of the possible explanations for such a big decrease in reproduction between dispersal and non-dispersal morphs in N116 compared to N127 might be the results of higher energy requirement in flight compared to increased walking. Heinrich (1975) has shown that in pollinator insects the metabolic cost rate of walking per unit time is much lower compared to flight. However, the metabolic rate of flight is also affected by the type of flight such as hovering and gliding and the distance of the flight. Therefore, future studies are needed to determine if the higher reduction is due to lower metabolic rate cost in increased walking or if such observation is due to the distance of flight and type of flights in aphids.

3.4.3 Reproductive success of offspring from different aphid morphs

In this Chapter, the reproductive success of offspring is dependent on the mother morphs with offspring coming from dispersal morphs having lower reproductive output compared to those from a non-dispersal mother. Our study is the first to report such a difference in aphids suggesting that trade-offs in aphids might be trans-generational and last more than one generation. One explanation for the reduced reproduction in the offspring from dispersal mothers could be due to maternal or grand-maternal effects. For example, Pers and Hansen et al. (2019) showed that stressed granddaughter aphids took a significantly longer time to complete nymphal development in comparison to the non-stressed granddaughter. Therefore, the lower reproduction in offspring from dispersal mothers (reared in good conditions) observed in our study might be due to the high level of crowding and starvation

conditions experienced by their grandmother. Next, Kangassalo et al. (2020) also reported trans-generational effects of larval diet on body size and developmental timing in the greater wax moth, *Galleria mellonella*. Kangassalo showed that both male and female offspring from low nutrition parents showed a longer development time. Apart from that, the host plant quality might play a role in this observation. Leather (1989) reported differences in offspring reproduction in bird-cherry oat aphids from different mother morphs depending on the quality of the host plant. In the normal host plant, the offspring from the winged mother has higher fecundity and faster development time. In contrast, lower fecundity and prolonged development in offspring from winged mothers were observed on poor host plants. Therefore, the difference in fecundity of offspring from different morph mothers observed in our experiment might be due to the host plant used.

3.5 Summary

In this Chapter, the response of two different aphid genotypes to crowding was investigated with a focus on the degree of alternative morph production and the type of alternative morph produced. The two aphid genotypes N116 and N127, obtained from Kanvil et al., 2014, has a high variability in the degree of alternative morph production as well as the type of alternative morph produced with the former producing winged offspring while the latter changing their body colour. A reduction in the offspring produced by different aphid morphs was observed in the transfer experiment with dispersal morphs in both genotypes having lower reproductive output compared to the non-dispersal morphs regardless of the population size from which the morphs were obtained. Finally, the reproductive output of the offspring produced by the dispersal mother was also found to be lower than those offspring that were produced by the non-dispersal mother.

The results from this Chapter strongly support the idea that aphid plasticity varies when exposed to the same environmental stress and the type of plasticity is dependent on the aphid genotype with some aphids changing their body colour instead of producing wings. By understanding how a given aphid genotype reacts to stress we may potentially understand the prevalence of different aphid genotypes in the wild. Apart from that, the results of this chapter also suggest a possibility of trade-offs between dispersal and non-dispersal regardless of the type of alternative morph (winged production or changing body colour with increased locomotion). Our result is the first to report such possibility of trade-off in an aphid genotype that does not produce wings, suggesting that trade-off in insects not only occur between flight and reproduction but also potentially between increased locomotion and reproduction as well. This is important and can potentially explain the prevalence of both genotypes in wild even though the changing of body colour morph (with increased locomotion) requires less time compared to winged production.

Lastly, our results showed the reproductive output of the offspring produced from different aphid mothers is also affected, suggesting trans-generational effects of plasticity, possibly caused by grandmaternal/maternal effects and the host plant genotype. Our result are the first to reveal such a decrease in reproductive output in the offspring from different aphid morphs. Since all aphids have telescopic generations, therefore understanding the potential grandmaternal effect on the future offspring reproduction can help provide insight into the evolutionary success of aphid under different environmental stress. For the remainder of this thesis, the different morphs induced in this experiment are used to investigate the candidate genes involved in aphid development, the transcriptome and methylome profile of the different pea aphid morphs and the role of epigenetic mechanisms in regulating aphid polyphenism.

Chapter 4: Expression of developmental genes and transcriptome profiles of dispersal and non-dispersal aphid in two aphid genotypes

4.1 Introduction

The ability to disperse (e.g. flight) is one of the major reasons behind the unparalleled diversity of insects and their evolutionary success. The advantages of wing evolution in insects are beneficial, for example due to increasing their ability to access novel resources and ecosystems and escape from predators (Wagner and Liebherr. 1992). The underlying developmental process of adult wing production has diverged significantly between insect orders. In holometabolous development, pupal reorganization usually occurs between the juvenile to adult stage, and the insect's wings are developed directly from the ectoderm (Jockusch and Ober 2004). One of the most used models to study wing development in holometabolous insects is *Drosophila melanogaster* (Serrano and O'Farrell 1997). During the embryonic development of *Drosophila*, the primordia cell for wing formation are set aside and the patterning event of these cells in the wing disc does not occur until they reach the larval instars (Bryant 1978; Cohen 1993). In the pupation stage, the wing appears similar to the adult stage but is usually found in a compacted form (Turner and Adler 1995). On the contrary, during hemimetabolous development or otherwise known as incomplete metamorphosis, the insects usually go through a sequence of juvenile stages that resemble the adult version without any pupal reorganization process. In hemimetabolous insects, the nymph possesses wing buds that extend from the thoracic body wall. The wing buds will slowly develop as the nymph grows followed by fully extended wings when they reach adulthood. As opposed to the widely studied holometabolous insects, the underlying molecular biology of wing development in hemimetabolous insects remains widely unknown (Jockusch and Ober 2004). Aphids, an insect in the order Hemiptera, became one of the best models to study wing development in hemimetabolous insects as they exhibit wing polyphenism. Aphids reproduce through parthenogenesis; therefore, all the daughters are essentially clones of their mother. A single aphid genotype can give rise to winged or wingless offspring depending on the environmental conditions experienced by the mother

(Müller et al. 2001, Braendle et al. 2006). The winged morphs differ from the unwinged morph not only in the presence of wings but also in other life-history traits such as reproductive output, immunity, developmental timing and others (Dixon and Howard 1986). In aphids, all nymphs usually possess the wing primordia (Johnson and Birks 1960; Tsuji and Kawada 1987). The nymph that will become winged and unwinged nymph cannot be distinguished by investigating the outer morphology during the 1st and 2nd instar stage. It is only at the 3rd instar stage that the two can be distinguished as the winged destined nymph will maintain and continue to grow their wing bud while the wingless nymph wing bud will be completely degenerated (Ganasi et al. 2005; Ishikawa et al. 2008). With the availability of the full genome sequence of the pea aphid, Brisson et al. (2010) have deduced the wing-patterning gene network and gene homologs for wing development in pea aphids based on the wing development pathways in *Drosophila melanogaster* (Figure 4.1).

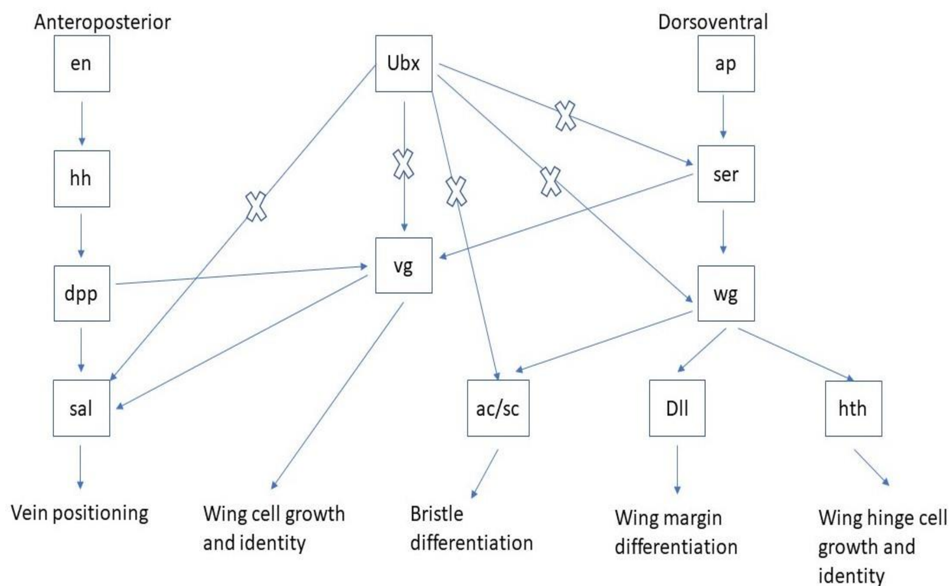


Figure 4.1. The wing-formation gene network of pea aphids deduced from *Drosophila melanogaster*. (Adapted from Brisson et al., 2010). Abbreviations: en, engrailed; hh, hedgehog; dpp, decapentaplegic; sal, spalt major; Ubx, Ultrabithorax; vg, vestigial; hth, homothorax; ap, apterous; Ser, Serrate; wg, wingless; Dll, Distalless; ac/sc, achaete/scute. Arrowheads indicate activation of genes, while arrowheads with a cross symbol indicate the repression of the gene. (Adapt and edited from Brisson et al. 2010).

The winged morph production in pea aphid is also highly dependent on the genotypes with some triggering a high proportion of winged offspring while others producing none, and some genotypes only change their body colour and increase locomotion as seen in previous study by (Tabadkani et al., 2013)

Apart from differences in their response to environmental stimulus, the two genotypes N116 and N127 are also quite different morphologically and physiologically (Section 2.1). For example, the genotype N116 usually exists in green body colour morphs while the N127 has a bright red/pink body colour. Next, the N116 is also more virulent and has higher reproductive output in comparison to the N127 (Kanvil et al. 2014). The molecular mechanisms that regulate this variation in life history traits between aphid genotypes remain poorly studied. Apart from less understood variation of molecular mechanism regulating life history traits between aphid genotypes, most studies on aphid polyphenism have been focused on just wingless and winged development. However, our results from Chapter 3 show that the N127 genotype rarely produces any winged offspring but changes their body colour with increased locomotion (Tabadkani et al. 2013). The mechanisms regulating locomotion and any difference in transcriptome profile between the two different dispersal morphs remain largely unknown.

This chapter investigates the candidate genes involved in aphid development including wing development, stress response, metabolite regulation and hormone regulation building upon the data from previous Chapter and previous studies (Brisson et al. 2010). The overall aim of this chapter was to investigate whether the candidate genes involved in aphid development are differentially expressed in the dispersal and non-dispersal morphs and between genotypes. Recent work has also focused on understanding the difference in aphid genotype and the differences in degree of phenotypic plasticity between aphid genotype (Kanvil et al. 2014; Parker et al. 2019; Sentis et al. 2019). Because of the critical role of the ability to disperse in aphid evolutionary success, this chapter also aimed to assess the

transcriptome profile between the aphid morphs and genotypes to provide insight into genes that could possibly regulate the difference in life-history traits between the morphs and genotypes. Specifically, the aims of this Chapter were to:

- Measure the expression level of genes involved in different pea aphid developmental processes such as reproduction, wing formation, carotene production and stress response by qPCR
- Investigate the transcriptome profile of different pea aphid morphs in two genotypes (N116 winged vs N116 wingless and N127 pale vs N127 red) by RNA-seq
- Investigate the differences in GO and KEGG pathways in pea aphid morphs
- Integrate the results of transcriptome profile (RNA-seq) with MBD-seq (data reported in Chapter 5)
- Investigate the protein expression of a candidate gene (Hsp70)

4.2 Methods

4.2.1. Materials

All samples used in the experiment were obtained from the mesocosm experiment described in section 2.2.1. The number of samples used in the qPCR analysis is outlined in (Table 4.1), stratified by genotype and morphs. The number of biological replicates used in the experiment is determine based on the minimum requirement of three biological replicates commonly used in qPCR analysis (Taylor et al. 2019).

Table 4.1. Summary of morph numbers used for molecular analysis in Chapter 4.

Assay	Genotype	Morphs	Replicates
qPCR	N116	wingless	x5 (pooled samples)
	N116	winged	x5 (pooled samples)
	N127	red	x5 (pooled samples)
	N127	pale	x5 (pooled samples)

4.2.2 Candidate gene expression by qPCR

Total RNA was extracted from whole adult aphids obtained as described in in Chapter 2 (section 2.4.2), quantified (section 2.4.2b), quality checked for integrity and quality (section 2.4.2c) and converted into cDNA (section 2.4.2d). Reference gene(s) used for data normalisation were selected based on a the top 5 most stable genes that were suitable to be use as endogenous control as determine by multiple analytical tool such as (geNorm, Normfinder, BestKeeper) as published in the study by Yang et al. (2014). geNorm calculate the stability of reference genes using pairwise variation between each reference genes and other reference with stability indicate by M-value. Further, NormFinder calculate the stability value (SV) among samples in the given group and select genes with lower SV. Then, BestKeeper calculate the stability of the reference genes based on a few parameters such as standard deviation of the Cq values, correlation coefficient (r), coefficient of

variance (CV) and also the p-value. Next, the reference genes selected were validated using our sample and analysed using qBase+ software v3.2 (Biogazelle, Belgium; section 2.4.2e). Candidate genes were quantified using a 1:20 cDNA dilution assay while the reference gene was quantified using 1:2 dilution and SYBR green as the fluorescent dye (both section 2.4.2f) with a melting curve analysis included to check for a single amplicon.

4.2.3 Transcriptome sequencing

Total RNA was submitted to the Genomic Technologies Core Facility (GTCF). The quality and integrity of the RNA samples was assessed using a 2200 TapeStation (Agilent Technologies) and then libraries were generated using the TruSeq® Stranded mRNA assay (Illumina, Inc.) according to the manufacturer's protocol. Briefly, total RNA (0.1-4μg) was used as input material from which polyadenylated mRNA was purified using poly-T, oligo-attached, magnetic beads. The mRNA was then fragmented using divalent cations under elevated temperature (3 minutes at 94°C) and then reverse transcribed into first-strand cDNA using random primers. The second strand of cDNA was then synthesised using DNA Polymerase I and RNase H. Following a single 'A' base addition, adapters were ligated to the cDNA fragments, and the products were then purified and enriched by PCR to create the final cDNA library. Adapter indices were used to multiplex libraries, which were pooled prior to cluster generation using a cBot instrument. The loaded flow cell was then paired end sequenced (76 + 76 cycles, plus indices) on an Illumina HiSeq4000 instrument. Finally, the output data was demultiplexed (allowing one mismatch) and BCL-to-Fastq conversion was performed using Illumina's bcl2fastq software, version 2.20.0.422.

4.2.4 Sequence assembly and annotation

Unmapped paired-end sequences from an Illumina HiSeq 4000 sequencer were assessed by FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequence

adapters were removed, and reads were quality trimmed using Trimmomatic_0.36 (Bolger et al. 2014). The reads were mapped against the reference *Acyrthosiphon pisum* genome v3.0 obtained from Bioinformatics Platform for Agroecosystem Arthropods (BIPAA) and counts per gene were calculated using Structural Annotation OGS3.0 using STAR_2.7.2b (Dobin et al. 2013). The genome assembly and structural annotation were obtained from the Bioinformatics Platform for Agroecosystem Arthropods (BIPAA). Normalisation, Principal Components Analysis, and differential expression were calculated in DESeq2_1.20.0 using default settings (Love et al. 2014). Blast2GO software (http://www.geneontology.org) was used for gene ontology (GO) annotations. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was used to identify significantly enriched metabolic pathways or signal transduction pathways in DEGs based on the database with the criteria of p-value < 0.1 and $\log_2$ fold change>1 (Corchete et al. 2020).

4.2.5 Protein expression by Western blot

Tissue lysates were prepared from different adult aphid morphs from Chapter 2 (section 2.2.1). Lysate protein concentration was quantified using Pierce BCA Protein Assay Kit (ThermoFisher Scientific, Massachusetts, USA). Samples were separated using SDS-PAGE followed by transferring to nitrocellulose member. Next, Membranes were incubated with HSP70 3A3 monoclonal mouse primary ab (Thermo Fisher, 1:1000 dilution) followed by polyclonal goat anti-mouse secondary Alexa Fluor 680 (Thermo Fisher Scientific, Massachusetts, USA, 1:10000 dilution), Lastly, the membrane is read using Odyssey Infrared Imaging System (LI-COR Biosciences, Nebraska, USA) at 700 nm and analysed using Odyssey software (LI-COR Biosciences, Nebraska, USA).

4.3 Results

4.3.1 RNA-seq analysis

Four different morphs from two genotypes were selected for whole transcriptome profiling: genotype N116 (wingless and winged) and genotype N127 (red and pale). The two genotypes have very different body morph colours: N116 is usually green, while N127 is usually red, and they respond differently to crowded conditions, with N116 producing winged offspring and N127 changing their body colour from red to pale.

4.3.2 Differentially expressed genes, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis between genotype and morphs

The transcriptomes of four different pea aphid morphs were compared based on the gene expression results. Many genes and pathways were identified that could play a key role in wing development, body colour, the trade-off between reproduction and flight, cuticle synthesis, and metabolism in the pea aphid. 79 million reads were obtained for N116 (winged), 81 million reads for N116 (wingless), 97 million reads for N127 (pale), and 82 million reads for N127 (red) (Table 4.2). After filtering (includes trimming and mapping to the reference genome), 38 million, 53 million, 80 million, and 68 million high-quality reads were obtained for each morph, respectively. Apart from that, principal analysis plot was carried out to visualise the overall patterns of gene expression across morphs (Figure 4.2).

Table 4.2. Number of raw reads and unique mapped reads for obtained from RNA seq.

Strain	Number of input reads	Uniquely mapped read
N116 winged	78,611,329	37,785,919
N116 wingless	80,283,756	52,690,343
N127 pale	96,893,751	79,957,269
N127 red	82,289,189	67,687,920

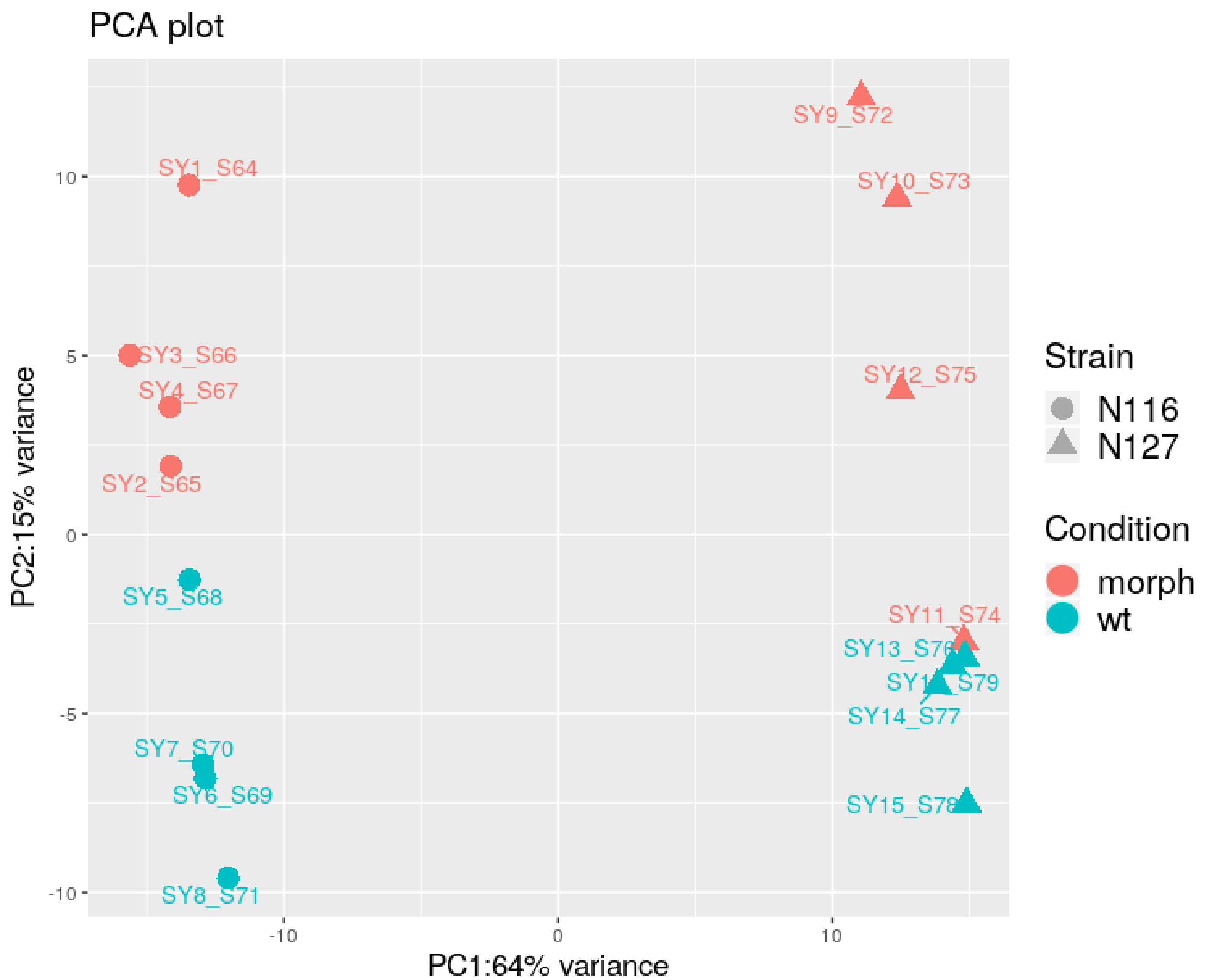


Figure 4.2. Principal component analysis of gene expression of five pea aphid morphs. Each dot represents one biological replicate from the RNA-seq experiment.

For gene expression measurements, a range of 35%–88% of the total reads were uniquely mapped to the genes in the reference database. The significantly DEGs for all groups was then further annotated for GO by using the Blast2Go parameters. The functions of the DEG by Blast2Go (P-adjust < 0.1) were then classified using GO assignments. For further functional categorisation, KEGG pathway analysis was performed using the pea aphid KEGG database. The DEG was then categorized into different KEGG Ontology (KO) terms.

a) Differentially expressed genes between different group comparison

Based on the results of RNA-seq we found 2308 genes that were differentially expressed between the N116 winged aphid vs N116wingless aphid, 2076 genes between N127 pale aphid vs N127 red aphid, 5108 genes between N127 red aphid vs N116 wingless aphid and 3153 genes between N127 pale vs N116 winged aphid (Figure 4.3). Then, we only found 533 overlapping genes between the N116 winged aphid vs N116wingless aphid group with N127 pale vs N127red aphid group. Further, we found 1102 genes overlap between N127pale vs N127red aphid group vs N127red vs N116wingless aphid group. Most overlapping genes 1214 were observed between the N127red vs N116wingless aphid group vs N127pale vs N116 winged aphid group. Then, 409 overlapping genes were found between N116winged vs N116 wingless group vs N127pale vs N116winged aphid group. Finally, 59 overlapping genes were found across all the group comparisons (Figure 4.3). All of the differentially expressed genes were obtained based on the parameter of (P-value <0.1 and -0.3<log2fold>0.3).

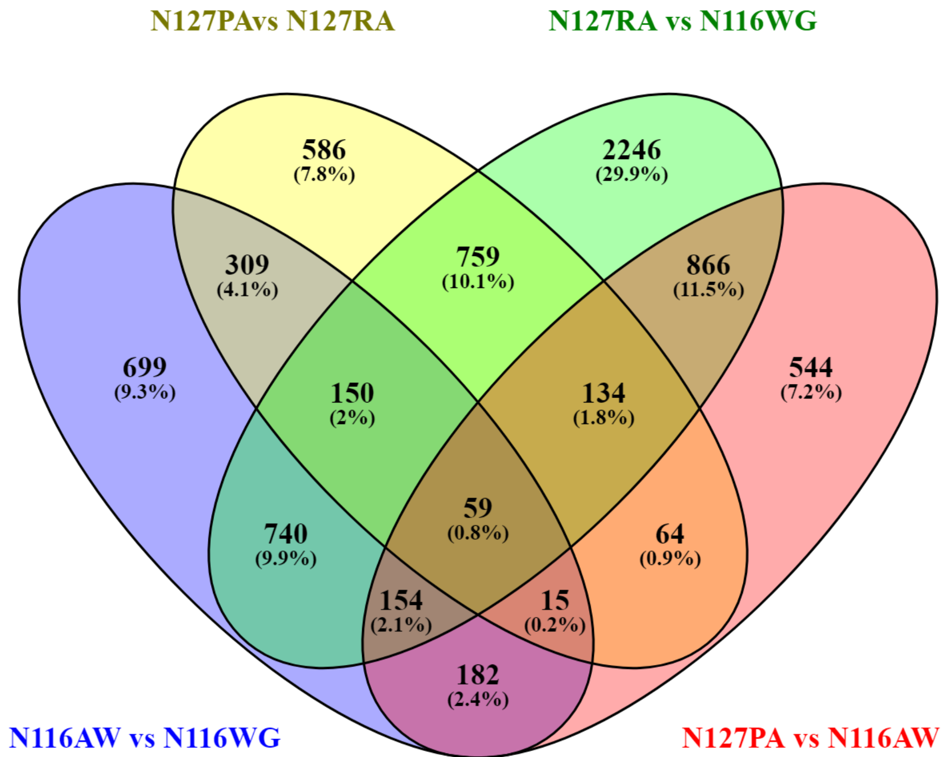


Figure 4.3. Venn diagram comparing DEG identified by RNA-seq between different aphid morphs. Blue=N116 winged +N116wingless morphs, Green=N127red +N116winglesss, Yellow= N127 pale + N127 red morphs, Red=N127 pale+ N116 winged morphs.**.** Each sub-category represents the overlapping genes that were found between the two or more different group comparisons.

b) Differentially expressed genes, GO and KEGG between N116 winged and N116 wingless aphids

44 differentially expressed genes were filtered out from the list of 2308 genes based on their potential role in aphid development. These genes were involved in wing development, reproduction, metabolism, longevity and cuticle formation. 19 genes involved in the formation of wing muscles, lipid metabolism, the behaviour and stability of wing expansion, cuticle formation, longevity, pigmentation, DNA methylation, insect hormone synthesis, and

trehalose metabolism were upregulated in the winged morphs. In contrast, 25 genes that were involved in fatty acid metabolism, insecticide resistance, carotene production, wing expansion, stress response, reproduction, glycogen metabolism, and immune response were downregulated in the winged morph (Table 4.3).

Next, 382 genes with (P-adjust<0.1 and log2fold>1) from the DEGS gene list were subjected to GO analysis. Significant enrichment was obtained for 179 genes that fall under 15 functional groups (Figure 4.4). The 179 genes were classified into three categories: 85 genes in biological process (BP), 89 in molecular function (MF) and five in cellular components (CC). Of the genes in the biological processes category, 40% were involved in transmembrane transport; 34% in proteolysis processes; and the remaining genes in ion transport, receptor signalling, and metabolic processes. In molecular function category, 34% of genes were involved in the structural integrity of the cuticle, 25% in transmembrane transporter activity; and the remaining genes in G-protein coupled activity, vitamin transportation and other function. Lastly, all genes in cellular components were involved in nucleosome activity.

In addition to the GO analysis, the 382 genes with (P-adjust<0.1 and log2fold>1) from DEG gene list were subjected to the Kyoto Encyclopedia of Genes and Genomes analysis. Significant enrichments were obtained for 36 genes that fall under 5 different KEGG function description (Figure 4.5). The top four KEGG functional description group with the highest gene count is lysosome, autophagy (animal), arginine and proline metabolism, and ABC transporter. The KEGG function description can be further categorised under three wider categories with 67% of genes in metabolism, 22% in environmental information processing and 11%in cellular processes.

Table 4.3. Differentially expressed genes between N116 winged and N116 wingless aphids. 38 genes has been filtered out from the 2308 genes based on their functions such as (metabolites, flight development, ecdysteroid pathway signalling)

Gene ID	Protein Information from NCBI	log2fold	p-value adjusted
LOC100570344	esterase E4	-2.11	0.0003
LOC100160034	phenoloxidase 1	-1.18	0.0008
LOC100159050	phytoene desaturase	-1.15	0.0135
LOC100160397	heat shock protein 83	-0.90	0.0034
LOC100167672	protein dachsous	-0.79	0.0184
LOC100168026	heat shock protein 70 B2	-0.74	0.0068
LOC100162796	homeotic protein spalt-major isoform X2	-0.63	0.0166
LOC100159778	glycogen phosphorylase isoform X1	-0.59	0.0085
LOC100164739	acetyl-CoA carboxylase	-0.56	0.0210
LOC100158806	protein held out wings isoform X1	-0.53	0.0149
LOC100165557	trithorax group protein osa isoform X4	-0.50	0.0288
LOC100167980	homeobox protein homothorax isoform X2	-0.50	0.0342
LOC100163455	heat shock 70 kDa protein 4	-0.47	0.0018
LOC100158703	optomotor-blind protein	-0.44	0.0058
LOC100574903	histone-lysine N-methyltransferase eggless	-0.41	0.0337
LOC100162023	histone deacetylase Rpd3	-0.40	0.0320
LOC100160060	protein gustavus isoform X4	-0.36	0.0148
LOC100159933	isocitrate dehydrogenase	-0.35	0.0027
Atpcl	ATP citrate lyase	-0.34	0.0413

Awd2	abnormal wing discs 2 isoform X1	0.37	0.0014
LOC100165833	ecdysone 20-monooxygenase isoform X2	0.41	0.0155
LOC100169493	mitogen-activated protein kinase p38b isoform X1	0.42	0.0191
LOC103309827	DNA methyltransferase 1-associated protein 1	0.42	0.0776
LOC100159349	G-protein coupled receptor Mth2	0.43	0.0081
LOC100161983	apolipoprotein D isoform X1	0.46	0.0279
LOC100162614	mothers against decapentaplegic homolog 4	0.49	0.0073
LOC100161202	lipase 3	0.51	0.0018
ORF2	chemosensory protein-like precursor	0.54	0.0293
LOC100168129	histone-lysine N-methyltransferase, H3 lysine-79 specific	0.59	0.0207
ORF4	chemosensory protein-like	0.65	0.0461
LOC100166282	partner of bursicon	0.67	0.0222
LOC100164834	Krueppel homolog 1 isoform X3	0.69	0.0303
LOC100166489	eclosion hormone-like	0.93	0.0014
LOC100158736	facilitated trehalose transporter Tret1 isoform X1	1.04	0.0191
LOC100570184	nose resistant to fluoxetine protein 6 isoform X1	1.13	0.0002
LOC100160700	phosphoenolpyruvate carboxykinase	1.60	1.87E-05
LOC100168346	troponin C	3.25	7.75E-06
LOC100161369	flightin	5.28	0.0001

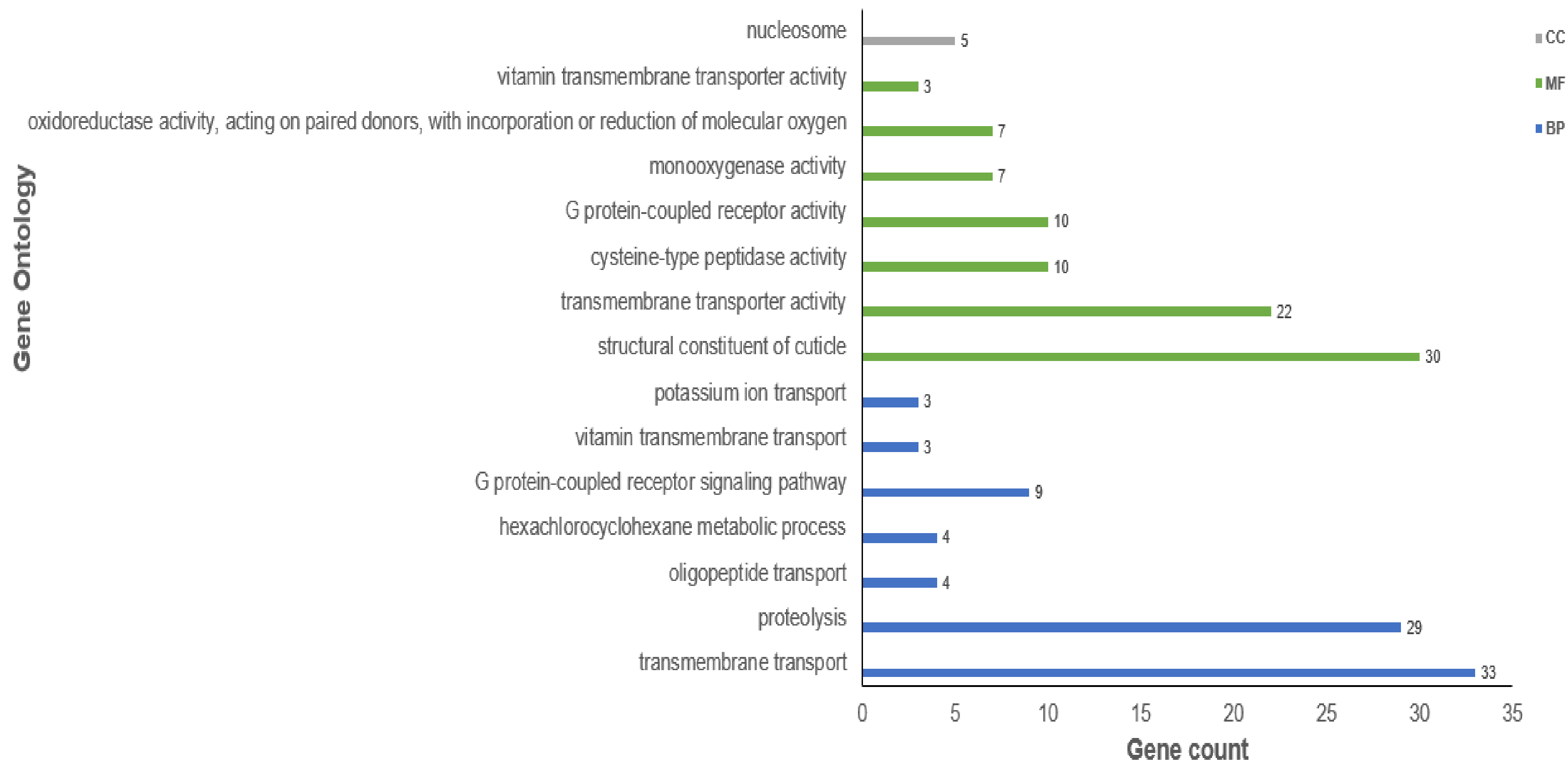


Figure 4.4. GO enrichments between N116 winged and N116 wingless aphids. Gene ontology was classified into three main categories: BP=Biological process (blue), MF=Molecular Function (green), CC=Cellular Components (grey). There are 85 genes in the biological process, 89 genes in molecular function and 5 genes in cellular component.

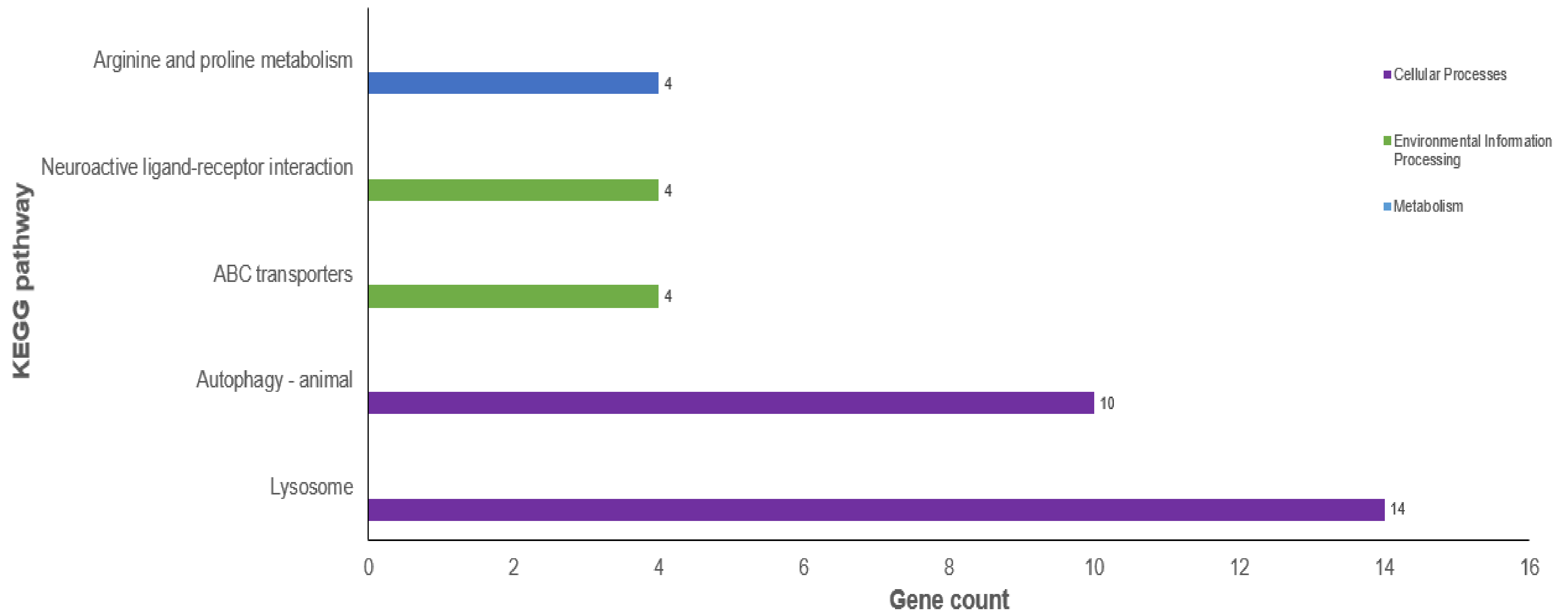


Figure 4.5. KEGG classification between N116 winged and N116 wingless aphids. X-axis represents the number of unigenes in the corresponding functional class. The genes were classified into three different wider categories, blue=metabolism, green=environmental information processing, purple=cellular processes. There are 24 genes in the cellular processes, 8 genes in environmental information processing and 4 genes in metabolism.

c) Differentially expressed genes, GO and KEGG between N127 pale and N127 red aphids

40 differentially expressed genes were filtered out from the list of 2076 based on their potential role in aphid development. These genes were involved in carotene production, reproduction, metabolism, longevity and cuticle formation. 23 genes involved in juvenile hormone production, lipid metabolism, fatty acid metabolism, response to starvation, longevity, pigmentation, DNA methylation, locomotion, and trehalose metabolism were upregulated in the pale morphs. In contrast, 17 genes involved in fatty acid metabolism, insecticide resistance, carotene production, stress response, reproduction, glycogen metabolism, and immune response were downregulated in the pale morphs (Table 4.4).

Next, 332 genes with (P-adjust<0.1 and log2fold>1) from the DEG gene list were subjected to the GO analysis. Significant enrichment was obtained for 100 genes that fall under 11 functional groups (Figure 4.6). The 100 genes were classified into three categories: 34 genes in biological process (BP), 66 molecular function (MF) and 0 cellular components (CC). Of the genes in the biological process, 74% were involved in transmembrane transport with the remaining genes involved in different amino acid catabolic process. In molecular function, 24% of genes were involved in transmembrane transporter activity, 15% in both glucuronosyltransferase activity and iron binding activity; and the remaining genes in acyl group transfer activity, monooxygenase activity and heme binding.

In addition to the GO analysis, the 332 genes with (P-adjust<0.1 and log2fold>1) from DEG gene list were subjected to the Kyoto Encyclopedia of Genes and Genomes analysis. Significant enrichments were obtained for 118 genes that fall under 16 different KEGG function descriptions (Figure 4.7). The top four KEGG functional description group with the highest gene count in is the biosynthesis of cofactors, drug metabolism - cytochrome P450, metabolism of xenobiotics by cytochrome P450, drug metabolism - other enzymes. The

KEGG function description can be further categorised under two wider categories with 94% of genes in metabolism pathways and 6% in cellular processes.

Table 4.4. Differentially expressed genes between N127 pale and N127 red aphids. 40 genes were filtered out from the 2076 genes based on function such as (metabolites, flight development, ecdysteroid pathway signalling).

Gene ID	Protein Information from NCBI	log2fold	p-value adjusted
LOC100159914	cuticular protein-like precursor	-5.40	0.0017
LOC100570344	esterase E4	-1.70	0.0045
LOC100574272	actin-1, partial	-1.25	0.0054
LOC100574964	bifunctional lycopene cyclase/phytoene synthase	-1.19	0.0060
LOC100163179	protein takeout isoform X2	-0.82	0.0521
Y-y	yellow-y precursor	-0.74	0.0372
LOC100166877	UDP-glucose 6-dehydrogenase	-0.59	0.0002
LOC100159795	protein phosphatase 1 regulatory subunit 3B isoform X1	-0.50	0.0421
LOC100574903	histone-lysine N-methyltransferase eggless	-0.44	0.0018
cprr1-4	RR1 cuticle protein 4 precursor	-0.42	0.0507
LOC100574964	juvenile hormone acid O-methyltransferase isoform X2	-0.40	0.0059
LOC100162429	dihydrolipoyl dehydrogenase, mitochondrial	-0.37	0.0013
LOC100162792	isocitrate dehydrogenase	-0.35	0.0226
LOC100166213	succinate dehydrogenase	-0.32	0.0111
LOC100167046	probable phosphoglycerate kinase	-0.32	0.0138
LOC100164251	pyruvate dehydrogenase E1 component subunit beta, mitochondrial	-0.32	0.0157
LOC100159778	glycogen phosphorylase isoform X1	-0.31	0.0324
Usp	ultraspiracle isoform X2	0.34	0.0068
LOC103309827	DNA methyltransferase 1-associated protein 1	0.34	0.0099
LOC100168659	insulin-like receptor	0.36	0.0214
LOC100169464	insulin-like peptide receptor isoform X1	0.39	9.26E-05
LOC100168097	forkhead box protein O	0.40	0.0067
LOC100161983	apolipoprotein D isoform X1	0.43	0.0016
LOC100160300	juvenile hormone epoxide hydrolase 1	0.50	0.0001
LOC100160300	juvenile hormone epoxide hydrolase 1	0.50	0.0001
LOC100161832	Insulin	0.51	0.0415
LOC100165833	ecdysone 20-monooxygenase isoform X2	0.53	0.0003
LOC100161380	phytoene desaturase	0.53	0.0464

LOC100159065	heat shock 70 kDa protein cognate 4	0.57	8.94E-07
LOC100167145	heat shock protein 68-like	0.64	0.0481
LOC100159349	G-protein coupled receptor Mth2	0.68	7.67E-06
LOC100168129	histone-lysine N-methyltransferase, H3 lysine-79 specific	0.75	0.0003
LOC100569479	insulin receptor substrate 1 isoform X1	0.76	4.97E-06
LOC100160293	Phosphoglucomutase	0.86	2.84E-05
LOC100166514	octopamine receptor	0.88	0.0205
LOC100161594	short-chain specific acyl-CoA dehydrogenase, mitochondrial	1.09	4.19E-07
LOC100574398	trehalase-like	1.10	0.0136
LOC100165228	Galactokinase	1.60	2.41E-06
LOC100167084	lipid storage droplets surface-binding protein 1 isoform X2	1.69	1.22E-15
LOC100169576	facilitated trehalose transporter Tret1 isoform X1	3.43	1.07E-08

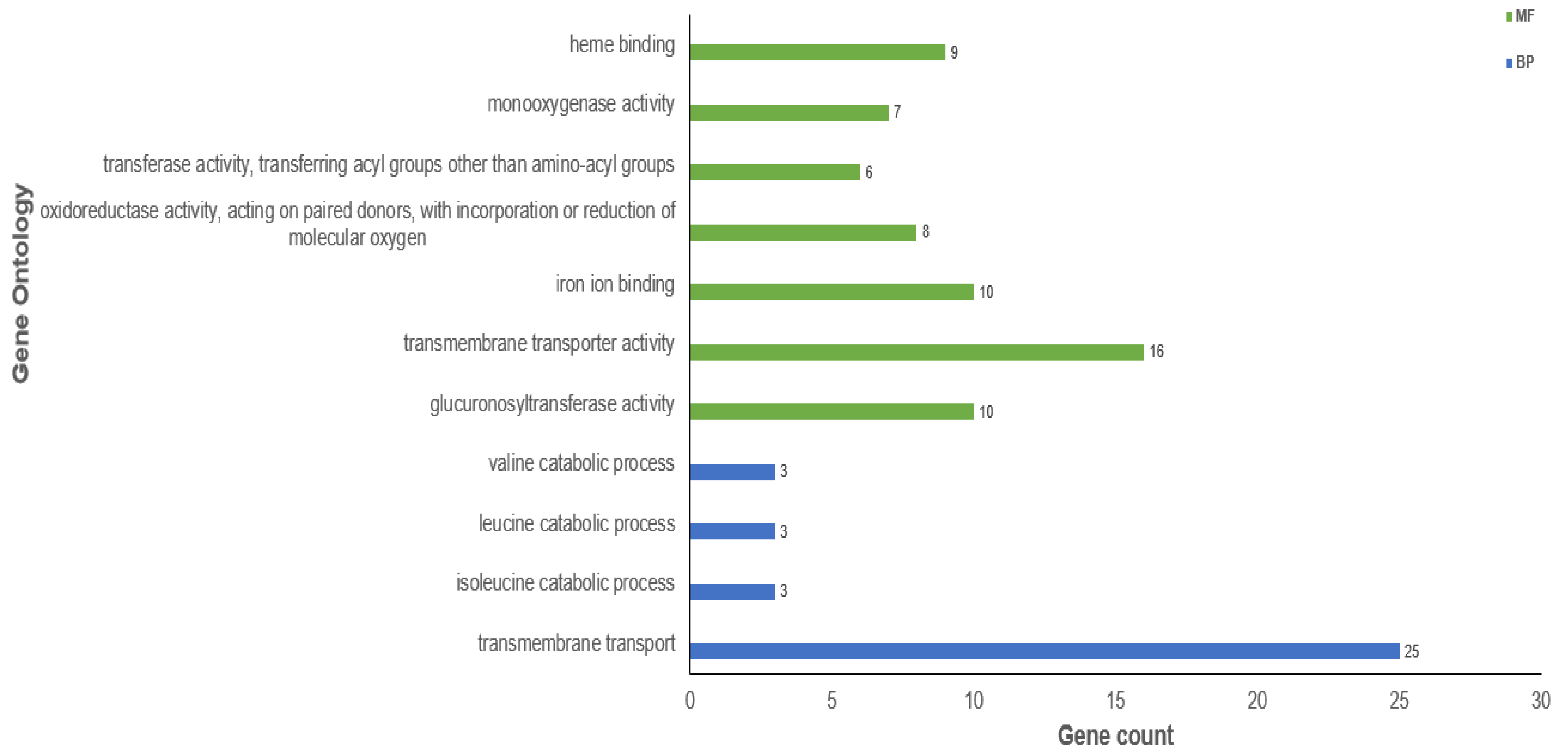


Figure 4.6. GO enrichment between N127 pale and N127 red aphids. The gene ontology was classified into three main categories: BP=Biological process (blue), MF=Molecular Function (green), CC=Cellular Components (grey). There are 34 genes in the biological process, 66 genes in molecular function.

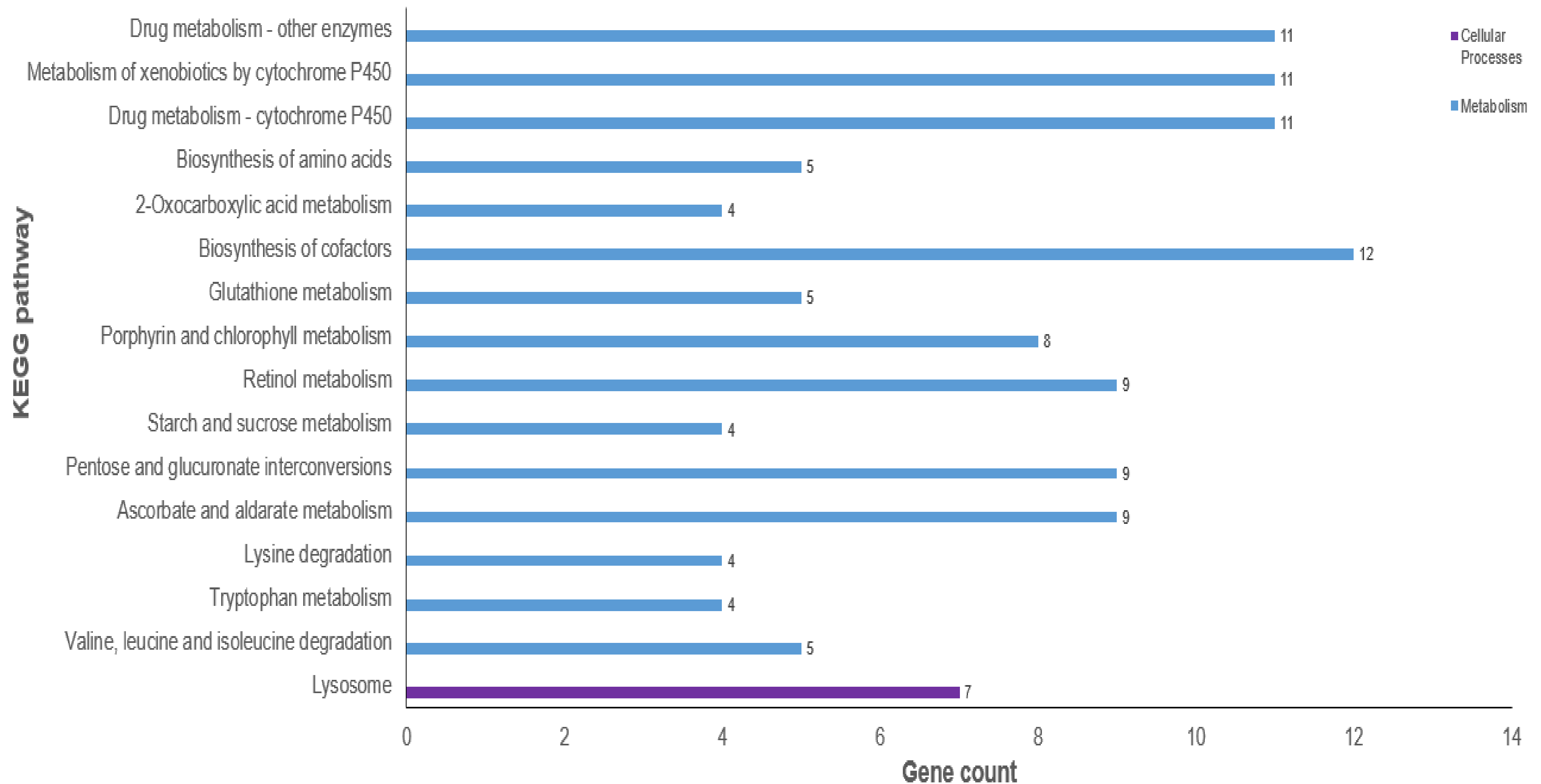


Figure 4.7. KEGG classification between N127 pale and N127 red aphids. X-axis represents the number of unigenes in the correponding functional class. The genes were further classified into two wider categories, blue=carbohydrate metabolism, purple=cellular processes. There are 7 genes in cellular processes and 111 in metabolism.

d) Differentially expressed gene, GO and KEGG between N127 red and N116 green wingless aphids

50 differentially expressed genes were filtered out from the list of 5108 genes. These genes were involved in carotene production, reproduction, metabolism, longevity and cuticle formation. 22 genes that were involved in eye pigmentation, insecticide resistance, stress response, metabolism, starvation response, virus attenuation, histone modification and cuticle synthesis were upregulated in N127 red morphs. In contrast, 28 genes involved in carotene production, wing morphogenesis, ATP synthesis, longevity, cuticle synthesis, survival, reproduction, insect hormone and wing development were downregulated in the N127 red morphs (Table 4.5).

Next, 1737 genes with (P-adjust<0.1 and log2fold>1) from the DEG gene list were subjected to GO analysis. Significant enrichment was obtained for 662 genes that fall under 26 functional groups (Figure 4.8). The 662 genes were classified into three categories: 364 genes in biological process (BP), 285 molecular function (MF) and 13 cellular components (CC). Of the genes in the biological process category, 24% were involved in proteolysis, 22% in transmembrane transport, 15% in DNA integration process and the remaining genes in response to oxidative stress, lipid and carbohydrate metabolic process and nucleotide catabolic process. In the molecular function category, 15% of genes were involved in heme binding, 14% in transmembrane transporter activity, 10% in serine endopeptidase activity, and the remaining genes in G-couple receptor activity, hydrolase activity, ion-channel activity and other amino acid activity. Lastly, all genes in the cellular components category were involved in egg chorion activity.

In addition to the GO analysis, the 1737 genes with (P-adjust<0.1 and log2fold>1) from DEG gene list were subjected to the Kyoto Encyclopedia of Genes and Genomes analysis. Significant enrichment was obtained for 277 genes that fall under 17 different KEGG function descriptions (Figure 4.9). The top four KEGG functional description group with the

highest gene count is biosynthesis of cofactors, lysosome, drug metabolism - other enzymes and metabolism of xenobiotics by cytochrome P450. The KEGG function description can be further categorised into three wider categories with 86% of genes in metabolism pathways 5% in environmental information processing and 9% in cellular processes.

Table 4.5. Differentially expressed genes between N127 red and N116 green wingless aphids. 49 genes were filtered out from 5108 genes based on their function such as (metabolites, flight development, ecdysteroid pathway signalling)

Gene ID	Protein Information from NCBI	log2fold	p-value adjusted
LOC100167145	heat shock protein 68-like	-3.32	8.52E-06
cprr1-2	RR1 cuticle protein 2 precursor	-3.10	8.23E-24
LOC100160300	juvenile hormone epoxide hydrolase 1	-2.19	3.40E-24
LOC100161104	bifunctional lycopene cyclase/phytoene synthase	-2.03	1.32E-27
LOC100159332	bifunctional lycopene cyclase/phytoene synthase	-1.41	0.0077
LOC100161475	fructose-1,6-bisphosphatase 1	-1.27	7.62E-05
LOC100161380	phytoene desaturase	-1.14	1.95E-07
LOC100572714	V-type proton ATPase subunit B	-0.98	3.37E-09
LOC100169464	insulin-like peptide receptor isoform X1	-0.90	5.78E-14
LOC100574964	bifunctional lycopene cyclase/phytoene synthase	-0.87	0.0269
LOC100161594	short-chain specific acyl-CoA dehydrogenase, mitochondrial	-0.85	0.0048
LOC100161043	trehalase isoform X2	-0.82	0.0002
LOC100159050	phytoene desaturase	-0.76	0.0157
Idgf	imaginal disk growth factor precursor	-0.75	0.0005
LOC100161053	ecdysone-induced protein 78C isoform X3	-0.73	4.35E-05
LOC100164834	Krueppel homolog 1 isoform X3	-0.73	0.0007
LOC100162796	homeotic protein spalt-major isoform X2	-0.68	0.0001
LOC100169645	heat shock protein 75 kDa, mitochondrial	-0.61	6.75E-12
LOC100159694	DNA N6-methyl adenine demethylase	-0.46	0.0088
LOC100168097	forkhead box protein O	-0.45	0.0137
LOC100165557	trithorax group protein osa isoform X4	-0.45	0.0012
LOC100166128	fructose-bisphosphate aldolase	-0.44	0.0015
LOC100167980	homeobox protein homothorax isoform X2	-0.37	0.0316
LOC100168563	heat shock protein 60A isoform X2	-0.36	0.0009
Awd1	abnormal wing discs 1 isoform X1	-0.31	0.0156
LOC100160060	protein gustavus isoform X4	-0.30	0.0009
Ecr	ecdysone receptor isoform A	-0.30	0.0134
LOC100160060	protein gustavus isoform X4	-0.30	0.0009

LOC100158682	set1/Ash2 histone methyltransferase complex subunit ASH2	0.38	0.0038
LOC100569254	juvenile hormone acid O-methyltransferase isoform X2	0.44	0.0022
LOC100168026	heat shock protein 70 B2	0.47	0.0251
LOC100163097	histone deacetylase complex subunit SAP18	0.52	6.10E-10
LOC100166213	succinate dehydrogenase	0.56	0.0002
LOC100168479	esterase E4	0.83	1.08E-15
LOC100574469	heat shock 70 kDa protein cognate 4	0.88	5.59E-14
ORF3	chemosensory protein-like precursor	1.20	0.0004
LOC100159282	cuticle protein 7	1.29	0.0003
LOC100162620	leucine-rich repeat extensin-like protein 2	1.63	0.0013
LOC100168775	histone deacetylase 8 isoform X1	1.70	1.69E-16
cp60	cuticular protein 60 precursor	1.70	4.03E-06
LOC100568695	G-protein coupled receptor Mth2	2.05	5.60E-11
LOC100162836	cytochrome P450 4C1-like isoform X1	2.46	3.91E-06
LOC100168000	probable G-protein coupled receptor Mth-like 2	3.39	1.63E-18
LOC100158748	histone acetyltransferase KAT7-like	3.48	9.23E-05
LOC100165740	retinol dehydrogenase 13	4.08	4.08E-19
LOC100167954	glyceraldehyde-3-phosphate dehydrogenase-like	4.46	0.0221
LOC100159685	takeout-like precursor	6.48	4.87E-25
LOC100570971	protoheme IX farnesyltransferase, mitochondrial-like, partial	6.60	7.86E-32
Tor	carotene dehydrogenase	7.34	6.16E-24
LOC100168987	UDP-glucuronosyltransferase 2B13 isoform X1	8.48	1.13E-20

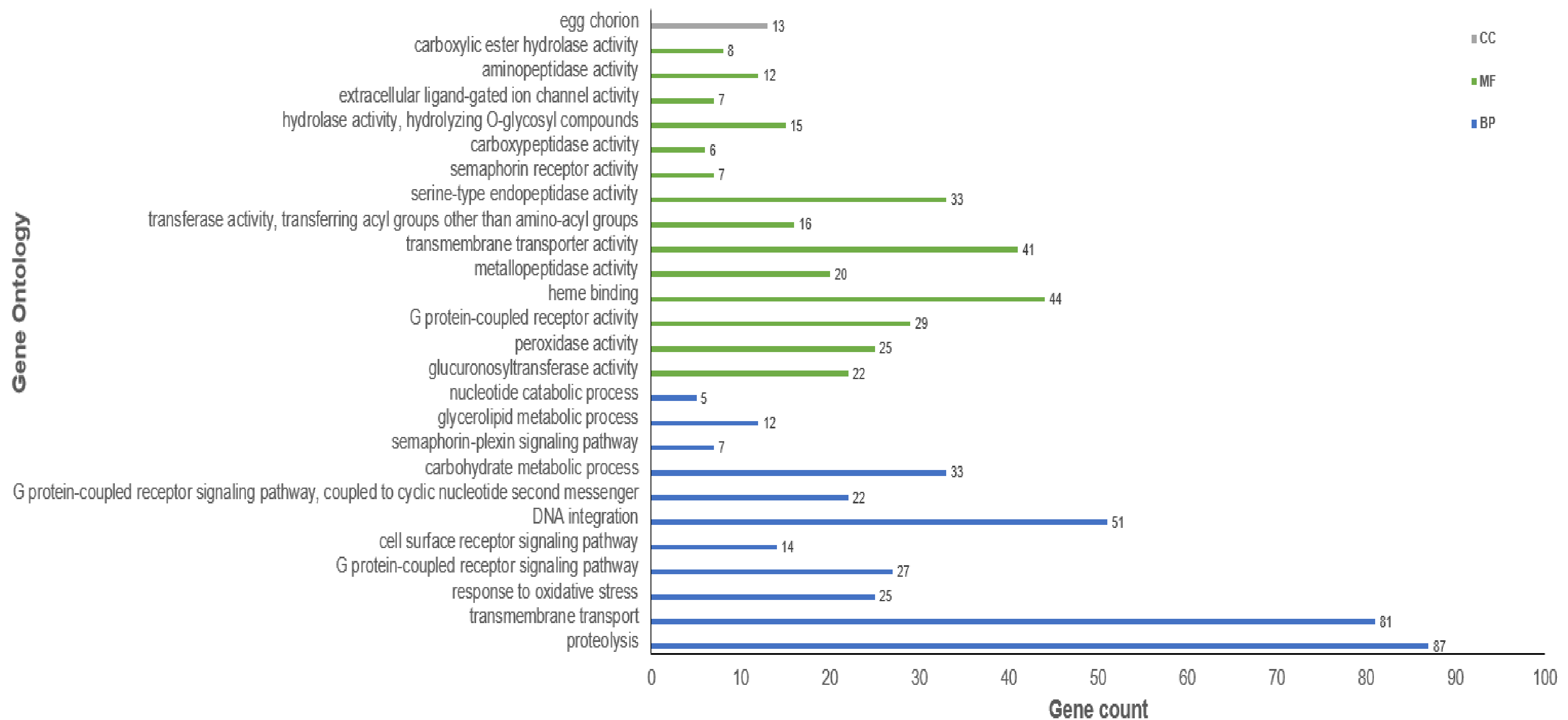


Figure 4.8. GO enrichment between N127 red and N116 green wingless aphids. The gene ontology was classified into three main categories: BP=Biological process (blue), MF=Molecular Function (green), CC=Cellular Components (grey). There are 364 genes in the biological process, 285 genes in molecular function and 13 in cellular component.

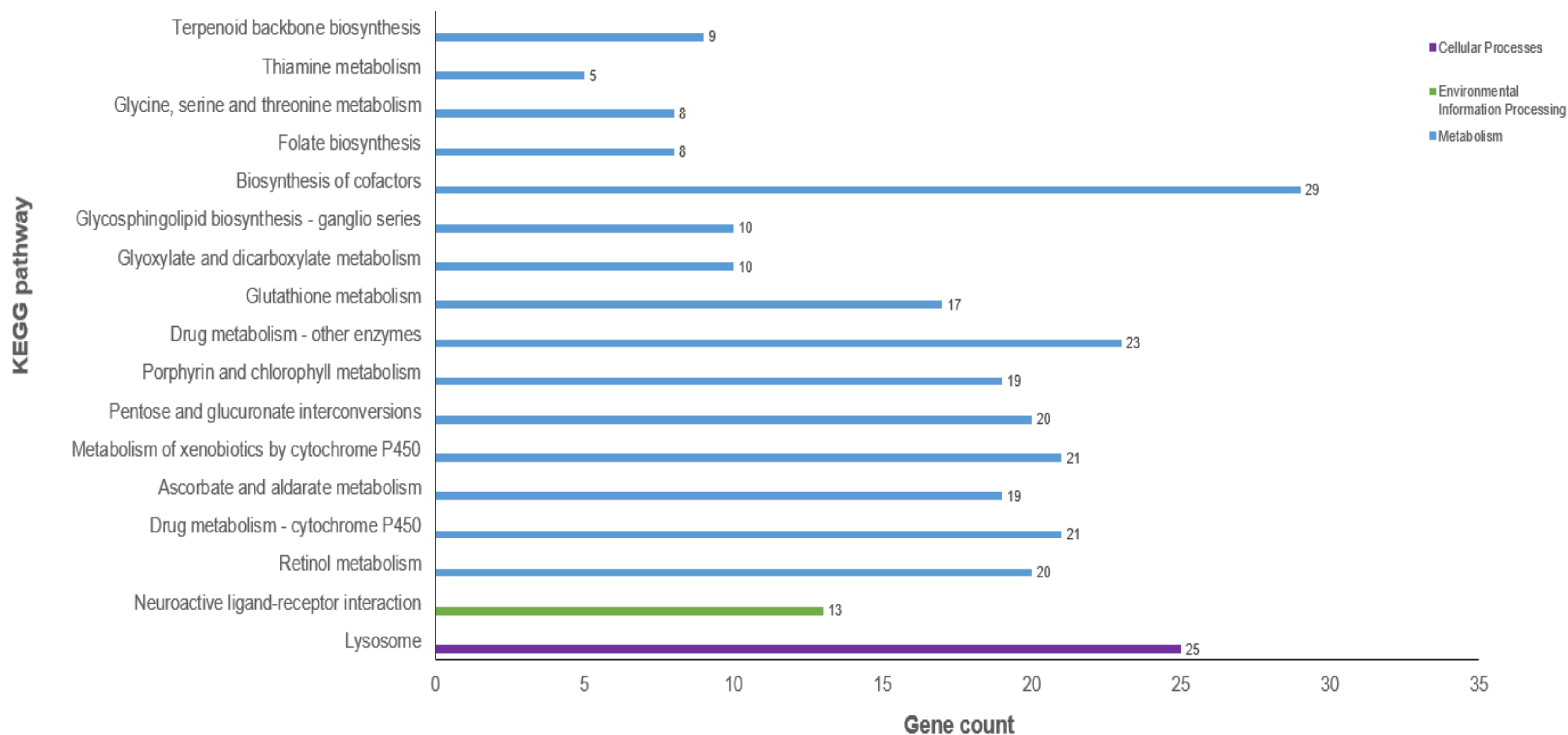


Figure 4.9. KEGG classification between N127 red and N116 green wingless aphids. X-axis represents the number of unigenes in the corresponding functional class The genes were classified into three different wider categories, blue=metabolism, green=environmental information processing, purple=cellular processes. There are 25 genes in cellular processes, 13 genes in environmental information processing and 239 genes in metabolism.

e) Differentially expressed genes, GO and KEGG between N127 pale and N116 winged aphids

33 differentially expressed genes were filtered out from the list of 3153 genes. These genes were involved in carotene production, reproduction, metabolism, longevity and cuticle formation. 16 genes involved in extended lifespan, eye pigmentation, insecticide resistance, insect hormone synthesis, response to stress, cuticle formation, metabolism and response to starvation, were upregulated in N127 pale aphids. In contrast,17 genes involved in wing development, wing muscle and wing bud development, wing expansion, insect hormone, trehalose transport, and reproduction genes were downregulated in the N127 pale morphs (Table 4.6).

Next, 1832 genes with (P-adjust<0.1 and log2fold>1) from the DEG gene list were subjected to the GO analysis. Significant enrichment was obtained for 496 genes that fall under 23 functional groups. The 496 genes were classified into three categories: 199 genes in biological process (BP), 257 molecular function (MF) and 28 cellular components (CC) (Figure 4.10). Of the genes in the biological process category, 33% were involved in proteolysis, 23% in DNA integration activity, 15% in carbohydrate metabolic process and the remaining genes in response to oxidative stress, apoptotic process, and cell surface receptor signalling. Next, 16% of genes in the molecular function category were involved in heme binding, 11% in transmembrane transporter activity, 10% in serine endopeptidase activity, and the remaining genes in oxidoreductase activity, monooxygenase activity, peptidase activity and others. 61% of genes in cellular component category were involved in nucleosome activity and the remaining genes in egg chorion activity.

In addition to GO analysis, the 1832 genes with (P-adjust<0.1 and log2fold>1) from the DEG gene list were subjected to the Kyoto Encyclopedia of Genes and Genomes analysis. Significant enrichment was obtained for 220 genes that fall under 16 different KEGG function descriptions (Figure 4.11). The top four KEGG functional description group with

the highest gene count is biosynthesis of cofactors, pentose and glucuronate interconversions, lysosome, drug metabolism - other enzymes. The KEGG function description can be further categorised under three wider categories with 90% of genes in metabolism pathways 1.5% in environmental information processing and 8.5% in cellular processes.

Table 4.6. Differentially expressed genes between N127 pale and N116 winged aphids. 33 genes were filtered out of 3153 genes based on their functions such as (metabolites, flight development, ecdysteroid pathway signalling)

Gene ID	Protein Information from NCBI	log2fold	p-value adjusted
LOC100161369	Flightin	-5.33	4.27E-09
LOC100168346	troponin C	-2.95	7.72E-07
cprr1-2	RR1 cuticle protein 2 precursor	-2.71	1.21E-17
LOC100166624	gustatory receptor for sugar taste 64f-like isoform X3	-2.52	0.0013
LOC100161104	bifunctional lycopene cyclase/phytoene synthase	-1.82	2.06E-22
LOC100574964	bifunctional lycopene cyclase/phytoene synthase	-1.62	4.75E-05
LOC100160300	juvenile hormone epoxide hydrolase 1	-1.55	1.87E-12
LOC100169115	facilitated trehalose transporter Tret1	-1.55	1.30E-12
Rlpa	RlpA family protein-like precursor	-1.06	0.0089
LOC107882419	trehalase-like	-0.92	0.0413
LOC100163068	S-adenosylmethionine sensor upstream of mTORC1	-0.73	0.0003
LOC100169645	heat shock protein 75 kDa, mitochondrial	-0.72	1.03E-15
LOC100159920	histone deacetylase Rpd3-like	-0.68	0.0041
LOC100166282	partner of bursicon	-0.64	0.0040
LOC100164578	octopamine receptor Oamb	-0.49	0.0003
LOC100169464	insulin-like peptide receptor isoform X1	-0.43	0.0012
LOC100169493	mitogen-activated protein kinase p38b isoform X1	-0.30	0.0142
LOC100162071	protein Wnt-1	0.36	0.0328
LOC100168026	heat shock protein 70 B2	0.58	0.0067
LOC100168479	esterase E4	0.74	5.26E-12
LOC100574469	heat shock 70 kDa protein cognate 4	0.88	8.44E-14
cp60	cuticular protein 60 precursor	0.99	0.0183
LOC100168775	histone deacetylase 8 isoform X1	1.14	2.00E-07
Jhamt	juvenile hormone acid methyltransferase isoform X1	1.78	0.0034
LOC100166187	farnesyl pyrophosphate synthase-like	1.91	1.08E-05
LOC100165740	retinol dehydrogenase 13	3.17	1.03E-13
LOC100568695	G-protein coupled receptor Mth2	3.31	3.89E-13
LOC100162620	leucine-rich repeat extensin-like protein 2	4.50	1.81E-21

LOC100168000	probable G-protein coupled receptor Mth-like 2	4.52	1.16E-16
LOC100159685	takeout-like precursor	7.62	3.67E-28
LOC100570971	protoheme IX farnesyltransferase, mitochondrial-like, partial	7.98	1.16E-20
tor	carotene dehydrogenase	9.63	2.53E-17
LOC100168987	UDP-glucuronosyltransferase 2B13 isoform X1	9.79	7.90E-21

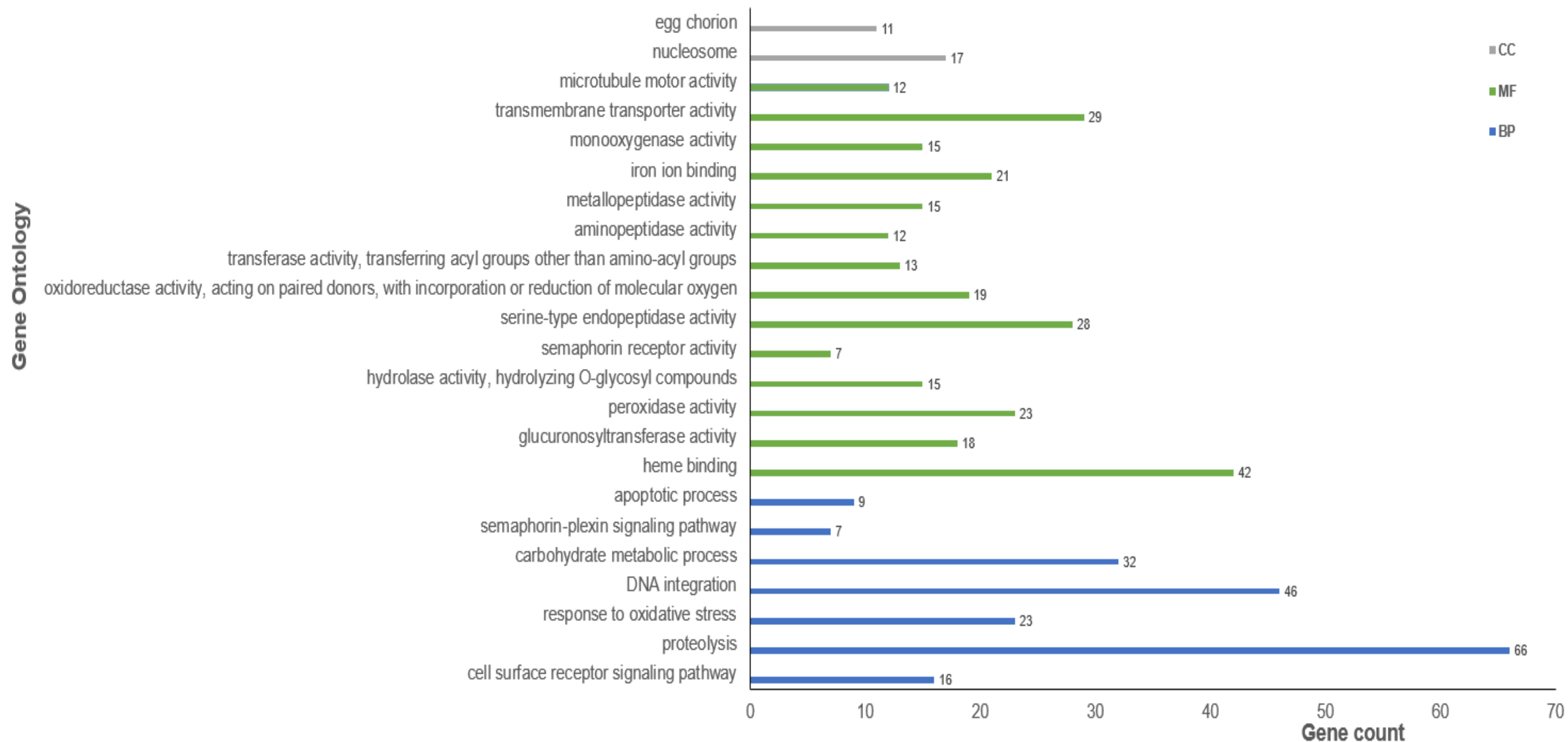


Figure 4.10. GO enrichments between N127 pale and N116 winged aphids. The gene ontology was classified into three main categories: BP=Biological process (blue), MF=Molecular Function (green), CC=Cellular Components (grey). There are 199 genes in biological process , 257 molecular function and 28 cellular components.

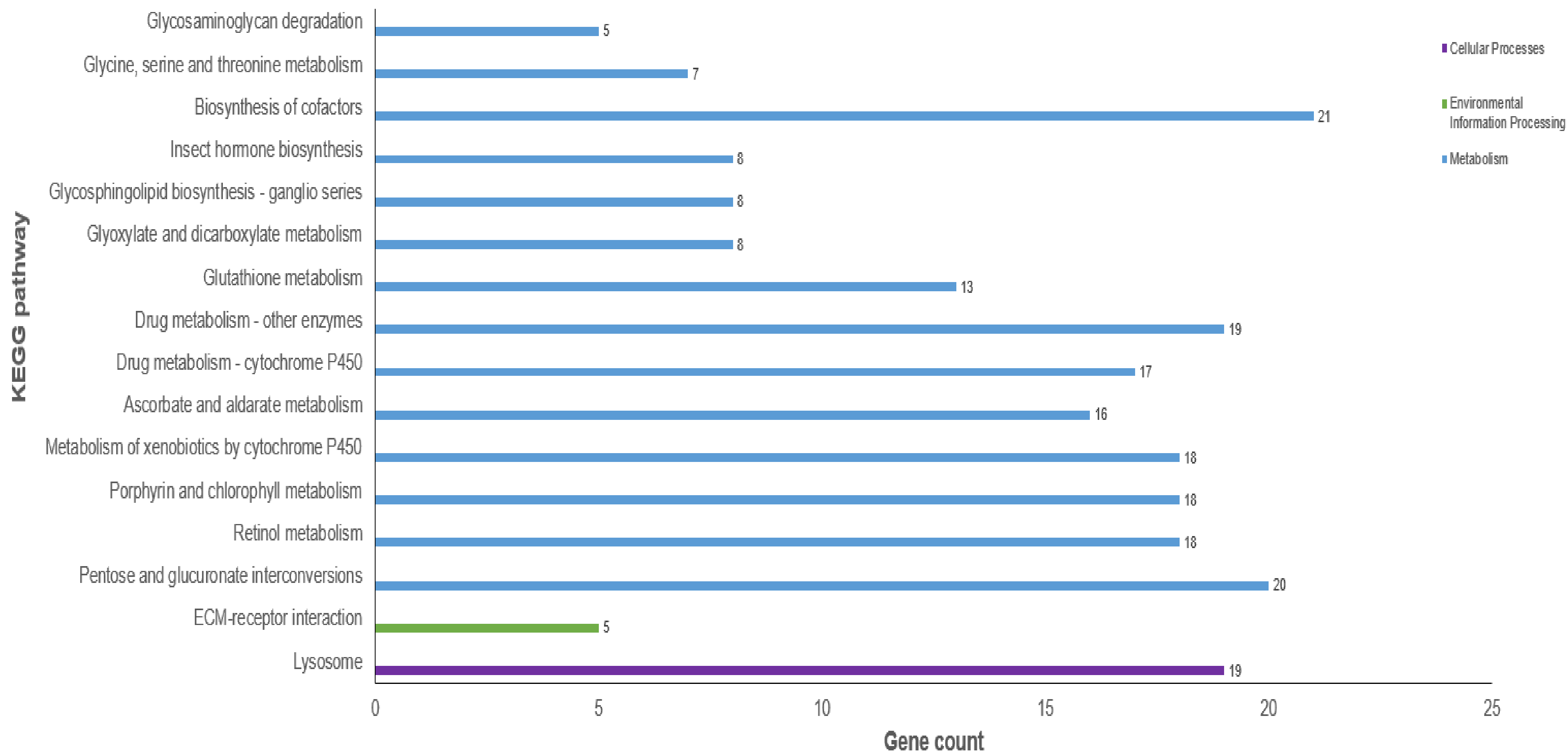


Figure 4.11. KEGG classification between N127 pale and N116 winged aphids. X-axis represents the number of unigenes in the corresponding functional class The genes were classified into three different wider categories, blue=metabolism, green=environmental information processing, purple=cellular processes. There are 19 genes in cellular processes, 5 genes in environmental information processing and 196 genes in metabolism.

f) Differentially expressed genes, GO and KEGG between N127 (red+pale) and N116 (winged+wingless) aphids

38 differentially expressed genes were filtered out from the list of 5262 genes based on their potential role in aphid development. 18 genes that were involved in response to stress, insect hormone synthesis, insecticide resistance, longevity, response to starvation, and metabolism of tricarboxylic acid cycle were upregulated in the N127 aphid (red+pale). In contrast, 20 genes that were involved in carotene production, wing morphogenesis, ATP synthesis, longevity, cuticle synthesis, survival, reproduction, insect hormone synthesis and wing development were downregulated in the N127 aphid (red+pale) (Table 4.7).

Next, 2257 genes with (P-adjust<0.1 and log2fold>1) from the DEG gene list were subjected to the GO analysis. Significant enrichment was obtained for 626 genes that fall under 26 functional groups (Figure 4.12). The 626 genes were classified into three categories: 334 genes in biological process (BP), 266 molecular function (MF) and 26 cellular components (CC). Of the genes in the biological process category, 24% were involved in proteolysis, 18% in transmembrane transport activity, 17% in DNA integration activity and the remaining genes in lipid and carbohydrate metabolic process, response to oxidative stress, apoptotic process and also cell surface receptor signalling pathway. In the molecular function category, 18% of genes were involved in heme binding, 13% in transmembrane transporter activity, 11% in peroxidase activity, and the remaining genes in peptidase activity, semaphoring-plexin receptor activity, monooxygenase activity, and hydrolase activity. In the cellular components category, 54% of genes were involved in egg chorion activity and 46% in nucleosome activity.

In addition to the GO analysis, the 2257 genes with (P-adjust<0.1 and log2fold>1) from the DEG gene list were subjected to the Kyoto Encyclopedia of Genes and Genomes analysis. Significant enrichment was obtained for 248 genes that fall under 16 different KEGG function descriptions (Figure 4.13). The top four KEGG functional description group with

the highest gene count is biosynthesis of cofactors, lysosome, drug metabolism - other enzymes and metabolism of xenobiotics by cytochrome P450. The KEGG function description can be further categorised into three different wider categories with 87.5% in metabolism, 3.5% in environmental information processing and 9% in cellular processes.

Table 4.7. Differentially expressed genes between N127 (red+pale) and N116 (winged+wingless) aphids. 38 were filtered out from 5262 genes based on their function such as (metabolites, flight development, ecdysteroid pathway signalling)

Gene ID	Protein Information from NCBI	log2fold	p-value adjusted
LOC100161369	flightin	-3.57	0.0012
LOC100164416	nose resistant to fluoxetine protein 6	-3.03	0.0027
LOC100166624	gustatory receptor for sugar taste 64f-like isoform X3	-2.93	7.46E-09
LOC100161380	phytoene desaturase	-2.70	1.21E-04
LOC100159920	histone deacetylase Rpd3-like	-2.16	0.0053
LOC100169645	heat shock protein 75 kDa, mitochondrial	-1.92	5.45E-16
LOC100169464	insulin-like peptide receptor isoform X1	-1.85	1.24E-08
LOC100158771	glycogen	-1.24	0.0387
LOC100163068	S-adenosylmethionine sensor upstream of mTORC1	-1.15	0.0007
LOC100168346	troponin C	-1.15	0.0016
LOC100574964	bifunctional lycopene cyclase/phytoene synthase	-0.73	0.0011
Rlpa	RlpA family protein-like precursor	-0.66	7.54E-06
LOC100168413	heat shock protein 68-like	-0.66	0.0002
Idgf	imaginal disk growth factor precursor	-0.62	0.0047
LOC100159332	bifunctional lycopene cyclase/phytoene synthase	-0.51	0.0470
LOC100161043	trehalase isoform X2	-0.50	0.0012
LOC100159050	phytoene desaturase	-0.46	0.0157
LOC100161104	bifunctional lycopene cyclase/phytoene synthase	0.34	4.55E-42

LOC100163097	histone deacetylase complex subunit SAP18	0.36	1.96E-06
LOC100166213	succinate dehydrogenase	0.38	0.012178
LOC100168026	heat shock protein 70 B2	0.44	0.037763
LOC100160300	juvenile hormone epoxide hydrolase 1	0.51	2.87E-26
LOC100166984	protein phosphatases pp1 regulatory subunit-like isoform X1	0.63	6.98E-46
LOC100168479	esterase E4	0.79	8.07E-28
LOC100574469	heat shock 70 kDa protein cognate 4	0.88	3.71E-14
cprr1-2	RR1 cuticle protein 2 precursor	1.15	1.32E-36
LOC100168775	histone deacetylase 8 isoform X1	1.43	3.60E-20
Jhamt	juvenile hormone acid methyltransferase isoform X1	1.46	7.92E-05
LOC100169473	pupal cuticle protein C1B-like	1.47	2.40E-05
LOC100569361	cuticle protein 7	2.21	2.30E-05
LOC100568695	G-protein coupled receptor Mth2	2.44	9.10E-22
LOC100571987	trehalase-like	3.57	0.000543
LOC100570971	protoheme IX farnesyltransferase, mitochondrial-like, partial	6.60	7.71E-51
LOC100159685	takeout-like precursor	6.94	1.03E-53
Tor	carotene dehydrogenase	7.34	2.10E-33

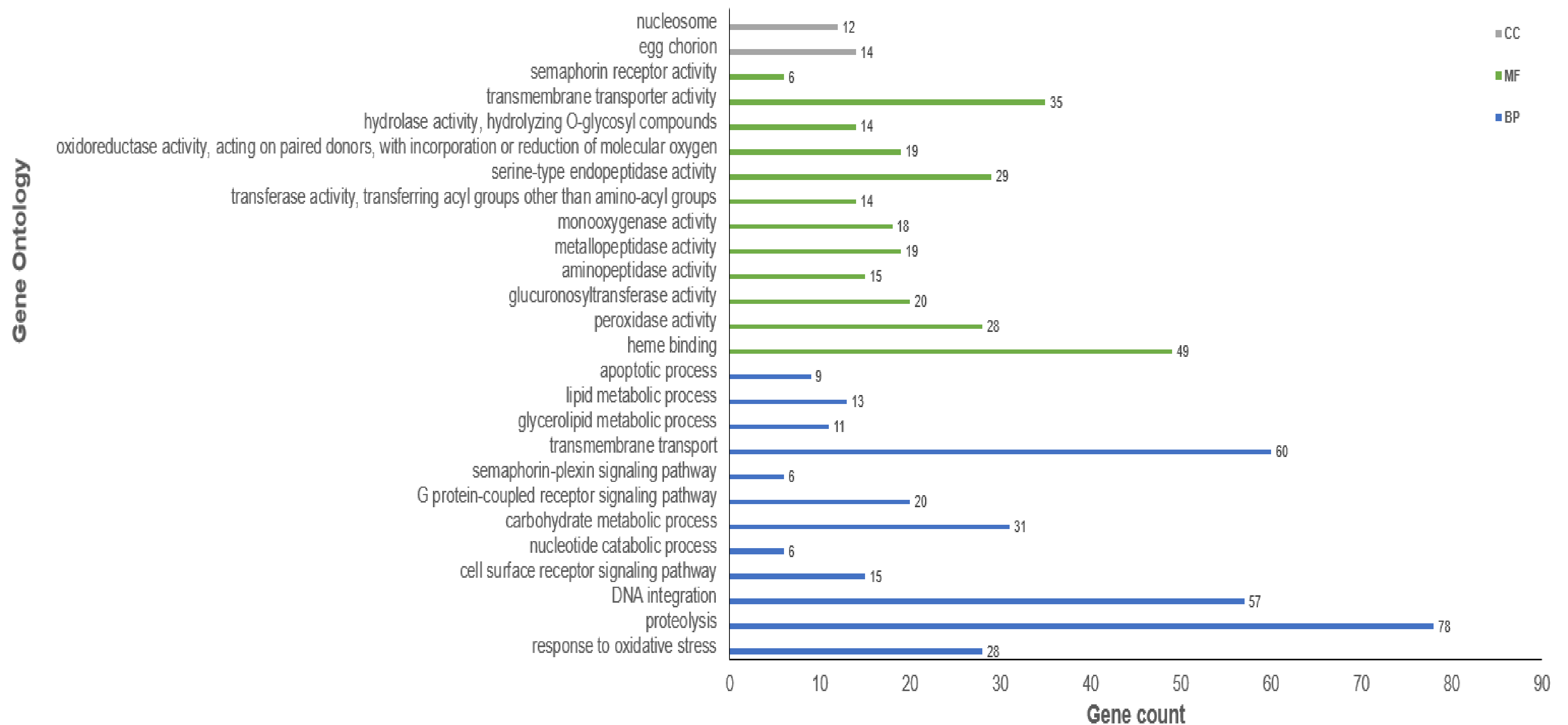


Figure 4.12. GO enrichments between N127 (pale+red) and N116 (wingless+winged) aphids. The gene ontology was classified into three main categories: BP=Biological process (blue), MF=Molecular Function (green), CC=Cellular Components (grey). There are 334 genes in biological process, 266 molecular function and 26 cellular components.

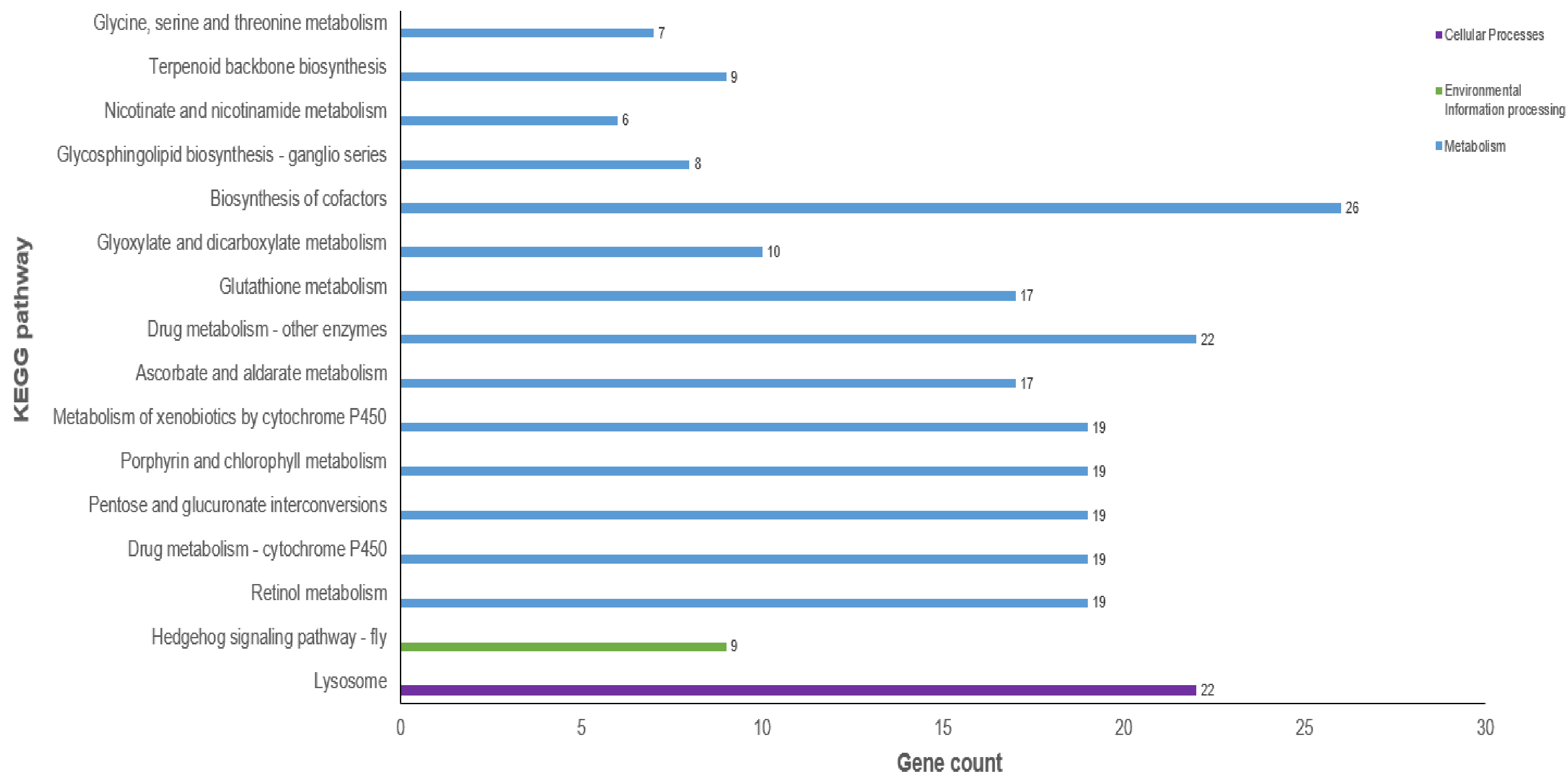


Figure 4.13. KEGG classification between N127 (red+pale) and N116 (winged+wingless) aphids. X-axis represents the number of unigenes in the corresponding functional class The genes were classified into three different wider categories, blue=metabolism, green=environmental information processing, purple=cellular processes. There are 22 genes in cellular processes, 9 genes in environmental information processing and 217 genes in metabolism.

g) Differentially expressed genes, GO and KEGG between alternative morphs (N127 pale+N116 winged) and wild type (N127red+N116 wingless)

23 differentially expressed genes were filtered out from the list of 5287 genes. 12 genes involved in wing muscle development, volatile reception, ecdysone hormone, longevity, DNA methylation and metabolism were upregulated in the aphid morphs (N127 pale + N116 winged). In contrast, 11 genes that were involved in response to stress, carotenoid biosynthesis, reproduction and tricarboxylic acid cycle were downregulated in the aphid morphs (N127 pale + N116 winged). (Table 4.8).

Next, 163 genes with (P-adjust<0.1 and log2fold>1) from DEG gene list were subjected to the GO analysis. Significant enrichment was obtained for 116 genes that fall under 15 functional groups (Figure 4.14). The 116 genes were classified into three categories: 45 genes in biological process (BP), 71 molecular function (MF) and 0 cellular components (CC). Of the gene in the biological process category, 51% were involved in transmembrane transport activity, 31% in proteolysis, and the remaining genes in riboflavin metabolic process, vitamin transmembrane transport and oligopeptide transport. In molecular function category, 30% of genes were involved in structural constituent of cuticle, 23% in transmembrane transporter activity, 8% in both heme binding and cysteine-type peptidase activity, and the remaining genes in ATPase coupled transmembrane transporter activity, acid phosphatase activity, iron binding, oxidoreductase activity and vitamin transmembrane transporter activity.

In addition to the GO analysis, the 163 genes with (P-adjust<0.1 and log2fold>1) from DEGS gene list were subjected to the Kyoto Encyclopedia of Genes and Genomes analysis. Significant enrichment was obtained for 23 genes that fall under four different KEGG function descriptions. The top four KEGG functional description group with the highest gene count is lysosome, autophagy – animal, ABC transporters and biosynthesis of amino acids. The KEGG function description can be further categorised into three different wider

categories with 17% in metabolism, 18% in environmental information processing and 65% in cellular processes. (Figure 4.15).

Table 4.8. Differentially expressed genes between alternative morphs (N127 pale+N116 winged) and wildtype (N127 red+N116 wingless). 23 genes were filtered out from 5262 genes based on their function such as (metabolites, flight development, ecdysteroid pathway signalling)

Gene ID	Protein Information from NCBI	log2fold	p-value adjusted
LOC100570344	esterase E4	-1.99	0.0001
LOC100574272	actin-1, partial	-1.11	0.0019
LOC100574951	cuticle protein 7	-1.10	0.0091
LOC100160736	heat shock protein 83	-0.68	0.0054
LOC100168026	heat shock protein 70 B2	-0.65	0.0040
LOC100159050	phytoene desaturase	-0.65	0.0515
LOC100165352	histone-lysine N-methyltransferase eggless	-0.44	0.0011
LOC100162792	isocitrate dehydrogenase	-0.41	0.0008
LOC100163748	heat shock 70 kDa protein cognate 4	-0.39	0.0004
LOC100162429	dihydrolipoyl dehydrogenase, mitochondrial	-0.36	9.34E-06
LOC100168563	heat shock protein 60A isoform X2	-0.36	0.0042
LOC100159682	Catalase	0.38	0.0305
LOC103309827	DNA methyltransferase 1-associated protein 1	0.39	0.0015
LOC100162614	mothers against decapentaplegic homolog 4	0.45	0.0039
LOC100165833	ecdysone 20-monooxygenase isoform X2	0.48	0.0007
LOC100159349	G-protein coupled receptor Mth2	0.54	0.0124
LOC100568675	nose resistant to fluoxetine protein 6 isoform X3	0.79	0.0039
ORF1	chemosensory protein-like precursor	1.23	0.0007
LOC100160700	phosphoenolpyruvate carboxykinase	1.51	0.0002
LOC100169576	facilitated trehalose transporter Tret1 isoform X1	2.15	6.07E-05
LOC100168346	troponin C	2.36	0.0023
LOC100161369	Flightin	3.01	0.0018

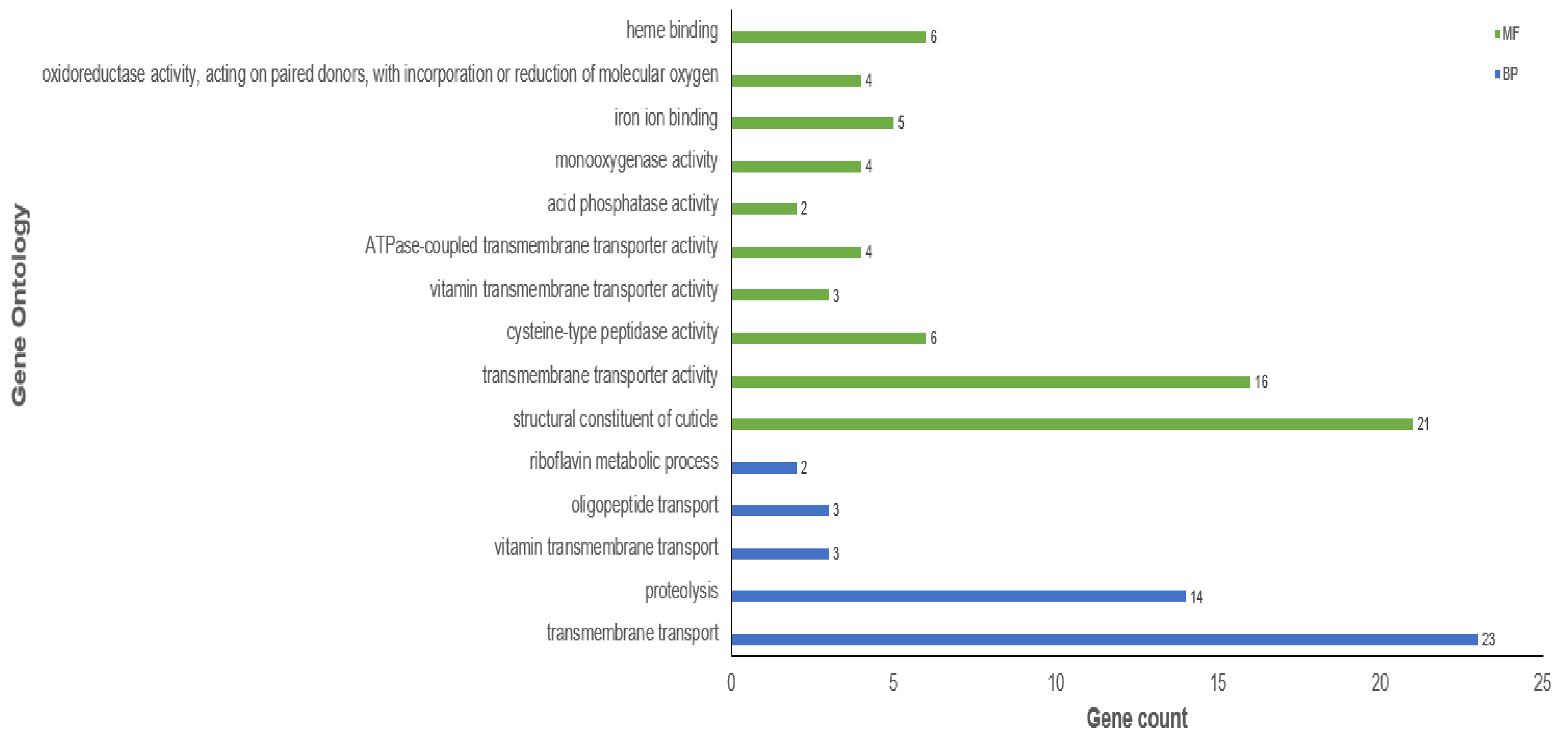


Figure 4.14. GO enrichment between alternative morphs (N127 pale+N116 winged) and wildtype (N127 red+N116 wingless). BP=Biological process (blue), MF=Molecular Function (green), CC=Cellular Components (grey). There are 45 genes in biological process and 71 genes in molecular function

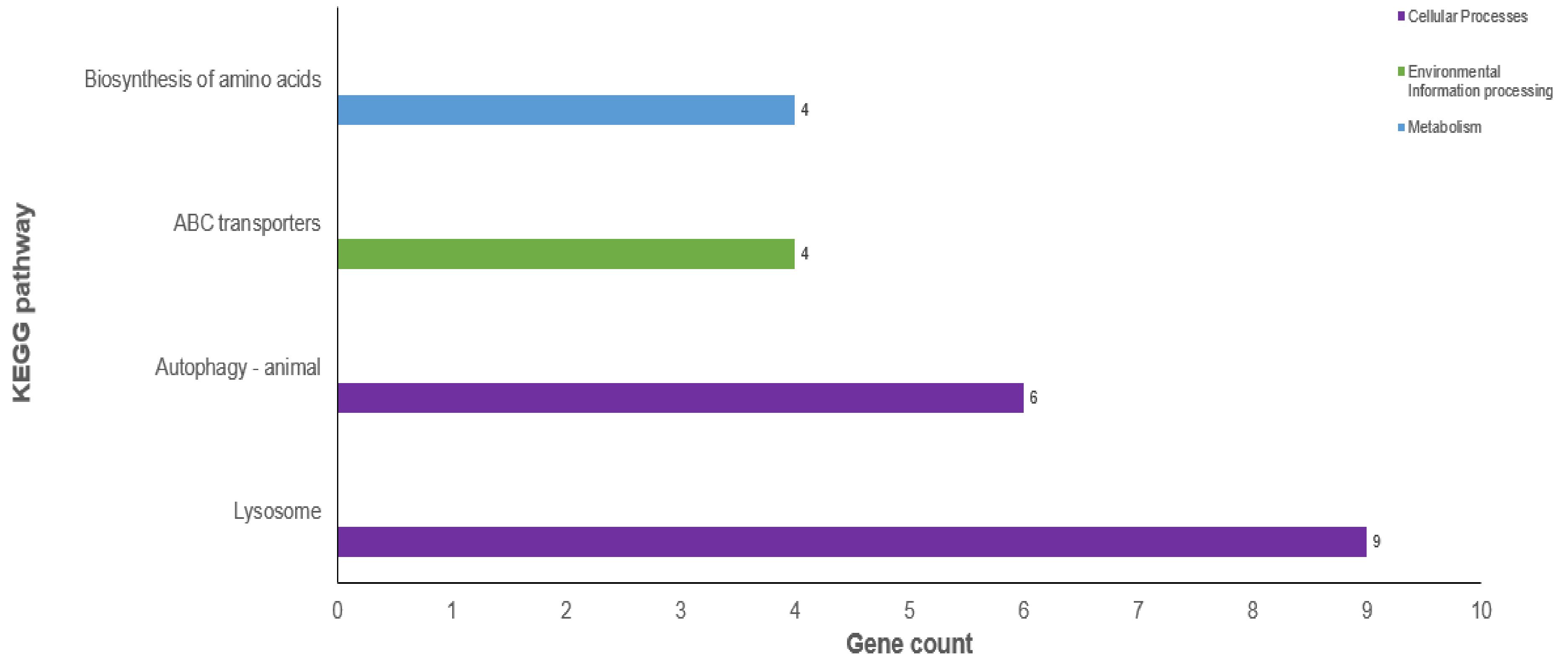


Figure 4.15. KEGG classification between alternative morphs (N127 pale+N116 winged) and wildtype (N127 red+N116 wingless). X-axis represents the number of unigenes in the corresponding functional class The genes were classified into three different wider categories, blue=metabolism, green=environmental information processing, purple=cellular processes. There are 15 genes in cellular process, 4 genes in environmental information processing and 4 genes in metabolism.

h) Differentially expressed genes, GO and KEGG between interaction of the aphid genotype and aphid morphs

1986 genes were significantly differentially expressed (P-adjust<0.1, -0.3<log2fold>0.3) between the interaction of aphid genotype and aphid morphs. 22 differentially expressed genes were filtered out from the list of 1986 genes. These genes were involved in wing development, reproduction, metabolism, longevity and cuticle formation. 14 genes involved in lipid storage, aphid endosymbiont systems, glycolysis, locomotion, trehalose transportation, stress response, and wing morphogenesis were upregulated. In contrast, seven genes that were involved in aphid immune system, wing muscle development, epigenetic maintenance, histone modification, pigmentation and wing expansion behaviour were downregulated (Table 4.9).

Next, 498 genes with (P-adjust<0.1 and log2fold>1) from the DEG gene list were subjected to a GO analysis. Significant enrichment was obtained for 116 genes that fall under 15 functional groups resulting in 13 functional groups (Figure 4.16). The 150 genes were classified into three categories: 68 genes in biological process (BP), 82 molecular function (MF) and 0 cellular components (CC). Of the gene in the biological process category, 44% were involved in transmembrane transport activity, 35% in proteolysis, 10% in ion transmembrane transport and the remaining genes in potassium ion transport and nucleotide catabolic process. In the molecular function category, 17% of genes were involved in catalytic activity, 16% in both iron ion binding and heme binding, and the remaining genes in oxidoreductase activity, glucuronosyltransferase activity, transmembrane signalling receptor activity, monooxygenase activity and also extracellular ligand-gated ion channel activity.

In addition to the GO analysis, the 498 genes with (P-adjust<0.1 and log2fold>1) from the DEGS gene list were subjected to the Kyoto Encyclopedia of Genes and Genomes analysis. Significant enrichment was obtained for 107 genes that fall under 11 different KEGG

function descriptions (Figure 4.17). The top four KEGG functional description group with the highest gene count is lysosome, autophagy – animal, ABC transporters and biosynthesis of amino acids. The KEGG function description can be further categorised into three different wider categories with 77% in metabolism, 9% in environmental information processing and 14% in cellular processes.

Table 4.9. Differentially expressed gene between interaction of aphid genotypes and aphid morphs. 22 genes were filtered out from 1986 genes based on their function such as (metabolites, flight development, ecdysteroid pathway signalling)

Gene ID	Protein Information from NCBI	log2fold	p-value adjusted
LOC100161369	Flightin	-5.29	0.0008
LOC100168346	troponin C	-3.09	0.0044
LOC100166282	partner of bursicon	-0.74	0.0470
LOC100169493	mitogen-activated protein kinase p38b isoform X1	-0.58	0.0018
LOC100162657	histone deacetylase 8 isoform X1	-0.57	0.0297
LOC100167839	polycomb protein EED	-0.54	0.0055
LOC100163097	histone deacetylase complex subunit SAP18	-0.31	0.0435
LOC100169464	insulin-like peptide receptor isoform X1	0.47	0.0365
LOC100159065	heat shock 70 kDa protein cognate 4	0.66	0.0061
LOC100162796	homeotic protein spalt-major isoform X2	0.73	0.0186
Samdc	S-adenosylmethionine decarboxylase isoform X1	0.75	0.0006
LOC100164834	Krueppel homolog 1 isoform X3	0.80	0.0371
LOC100169318	nose resistant to fluoxetine protein 6	0.86	0.0196
LOC100163393	phenoloxidase 2	1.04	0.0374
LOC100571236	locomotion-related protein Hikaru genki isoform X1	1.11	0.0013
LOC100160293	Phosphoglucomutase	1.15	0.0048
LOC100159380	facilitated trehalose transporter Tret1	1.20	0.0086
LOC100160570	protein yellow-like	1.68	0.0147
LOC100167084	lipid storage droplets surface-binding protein 1 isoform X2	1.69	1.22E-15
cprr1-8	RR1 cuticle protein 8 isoform X1	1.81	0.0050
LOC100161878	esterase E4 isoform X1	2.15	0.0009

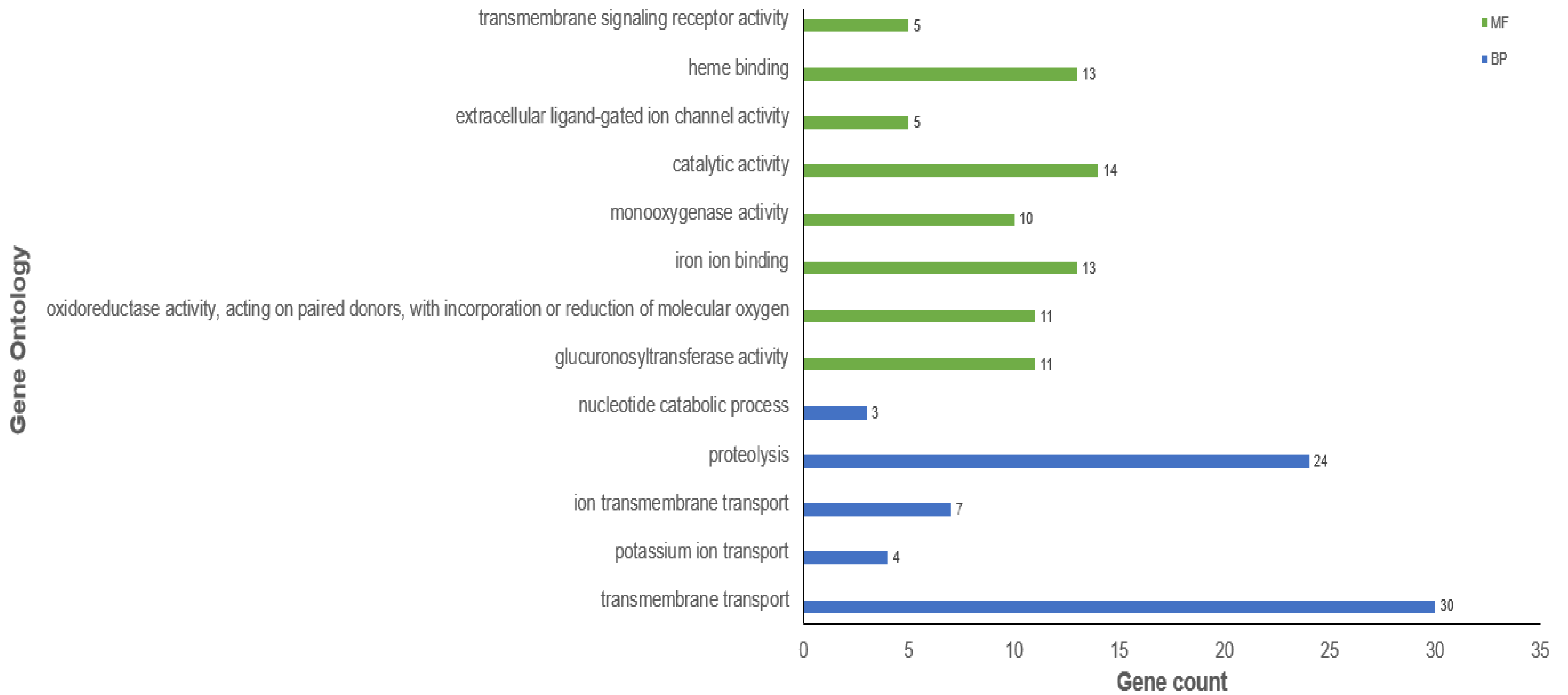


Figure 4.16. GO enrichment in the interaction of aphid genotypes and aphid morph. BP=Biological process (blue), MF=Molecular Function (green), CC=Cellular Components (grey). There are 68 genes in biological process and 82 genes in molecular function.

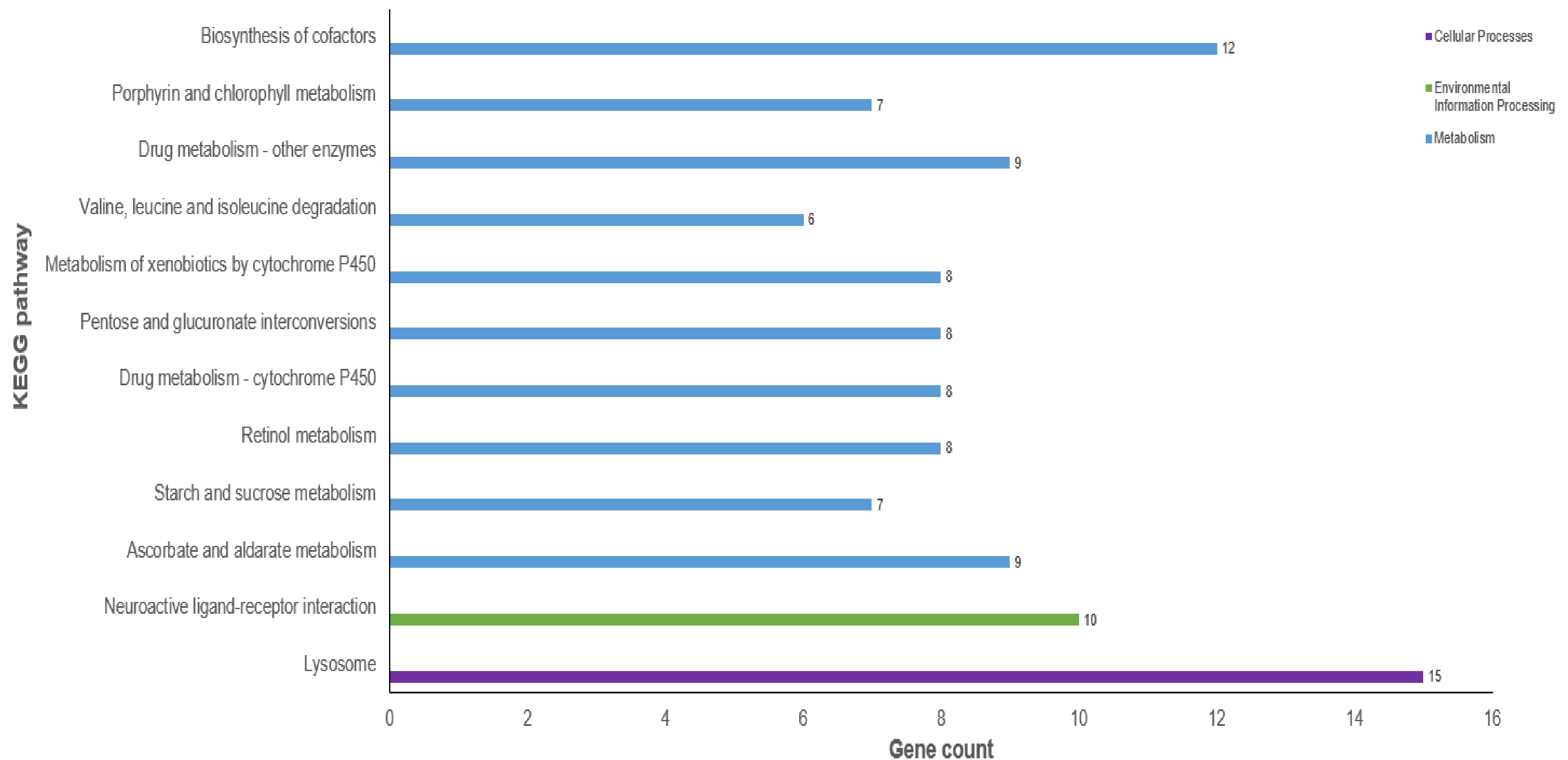


Figure 4.17. KEGG classification in the interaction of aphid genotypes and aphid morph. X-axis represents the number of unigenes in the corresponding functional class The genes were classified into three different wider categories, blue=metabolism, green=environmental information processing, purple=cellular processes. There are 15 genes in cellular processes, 10 genes in environmental information processing and 82 genes in metabolism.

4.3.3 Gene expression

We selected eight genes from the differentially expressed gene list obtained from RNA-seq from the (N116 winged vs N116 wingless) group based on their potential role in aphid development such as wing formation, stress response, ecdysone signalling, reproduction for qPCR validation. Another additional eight genes were selected to investigate their expression level between aphid morphs and genotypes based on their role in aphid development (wing development, stress response, carotene production, reproduction) (Table 4.10) as published in previous studies (Brisson et al. 2010; Zhang et al. 2017; Parker et al. 2019; Zhang et al. 2019).

a) geNorm analysis

Analysis of the list of reference genes commonly used in aphid sample determined a geNorm M-value (stability of each reference gene where a lower value represents higher stability across experimental groups, Figure 4.18) threshold for which all three of the five reference genes fell below (M<0.5), resulting in the combinations to the two-reference genes with lowest M-value being used for normalisation.

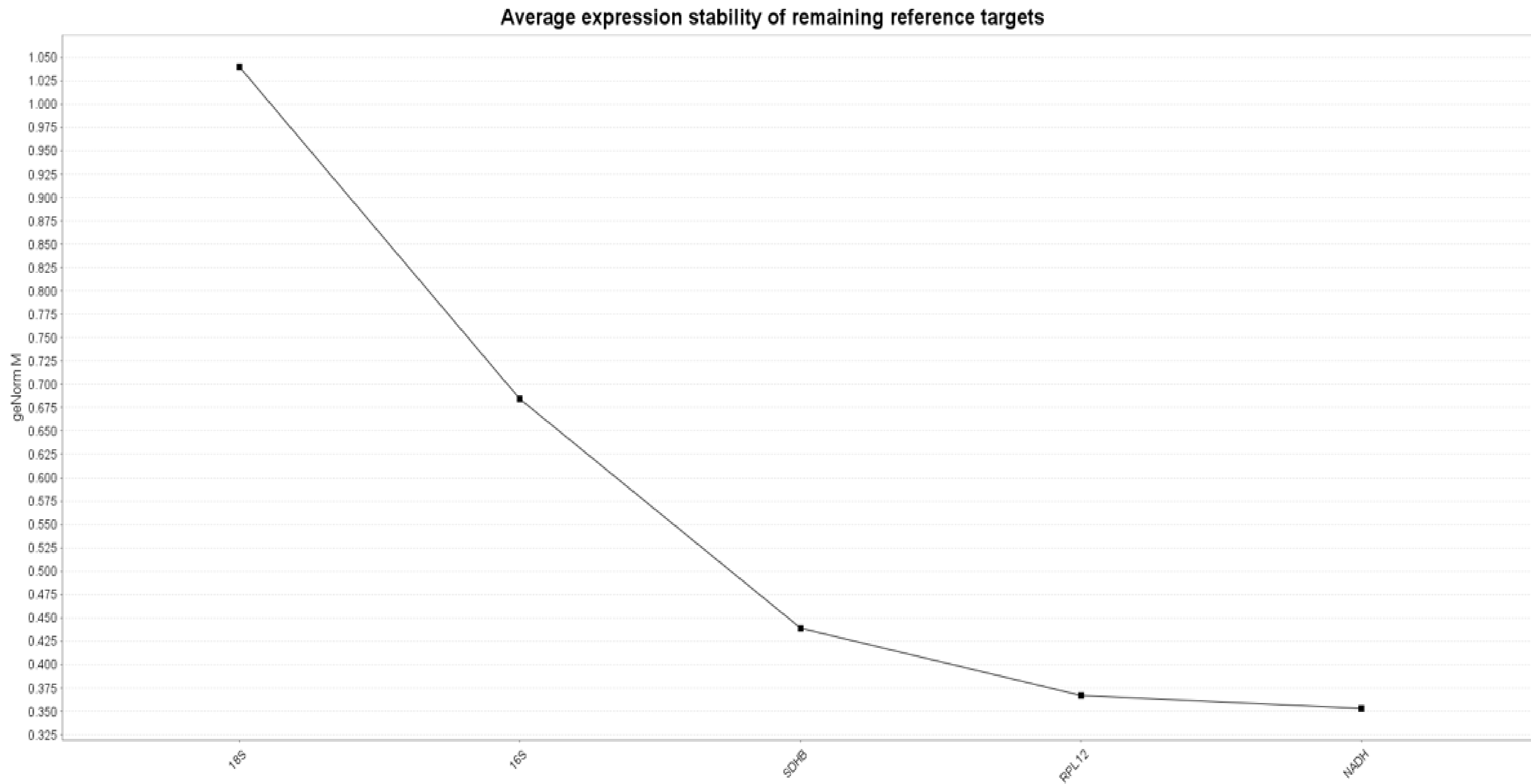


Figure 4.18. Representative geNorm analysis using qBase+ software to identify the optimal number of reference genes. The geNorm M-value relates to stability of each reference gene, where a lower M-value signifies increased stability (Yang et al. 2014). The genes selected here are those with the lowest M-value and highest stability.

b) RNA-seq gene validation

The expression of 8 different genes selected from RNA-seq based on the group comparison of (N116 winged vs N116 wingless). The genes were validated using qPCR and the overall visualising of gene expression is presented in (Figure 4.19). Next, the gene expression of another 8 different is also analysed and were selected based on their function in pea aphids reported in previous studies (such as stress response, wing formation, carotenoid production) (Figure 4.20). Lastly, the significant predictors for each gene were reported in (Table 4.10).

Table 4.10. Summary of statistical model used for gene expression analysis and the significant predictors for each gene. The graph representing first 8 genes (validation from RNA-seq results) is presented in Figure 4.19 and the remaining genes were presented in Figure 4.20 (selected based on previous literature and research).

Gene	Model	Significant predictors	F-value	P-value
Troponin C	GLM	Interaction of morphs and genotype	19.899	<0.001
Trehalase	GLM	Main effect (morphs)	8.186	0.011
Krueppel-homolog 1	GLM	Main effect (morphs)	16.861	<0.001
Mother against decapentaplegic homolog 4	GLM	Main effect (morphs)	16.861	<0.001
Partner of Bursicon	GLM	Interaction of morphs and genotype	10.845	0.002
DNA methyltransferase 1 associated protein 1	GLM	Main effect (morph)	5.493	0.009
Histone deacetylase Rpd3	GLM	Main effect (morph)	5.493	0.009

Ecdysone 20-monooxygenase	GLM	Interaction of morphs and genotype	6.097	0.025
Flightin	GLM	Interaction of morphs and genotype	19.899	<0.001
Apterous 1	GLM	Interaction of morphs and genotype	6.369	0.023
Engrailed	GLM	None of the predictors were significant. **reported value here is from the interaction of morphs and genotype	0.371	0.551
Vestigial	GLM	Interaction of morphs and genotype	4.063	0.037
Short-chain specific acyl-coA dehydrogenase	GLM	Main effect (morphs)	7.031	0.011
Heat Shock Protein 83	GLM	None of the predictors were significant. **reported value here is from the interaction of morphs and genotype	0.789	0.388
Carotene dehydrogenase	GLM	Main effect (morphs) ** morphs here represent only pale vs red aphid in N127 genotype as no expression of these genes for N116 genotype	1.732	0.225
Acyrthosiphon pisum densovirus	GLM	Main effect (morphs) ** morphs here represent only pale vs red aphid in N127 genotype as no expression of these genes for N116 genotype	0.29	0.605

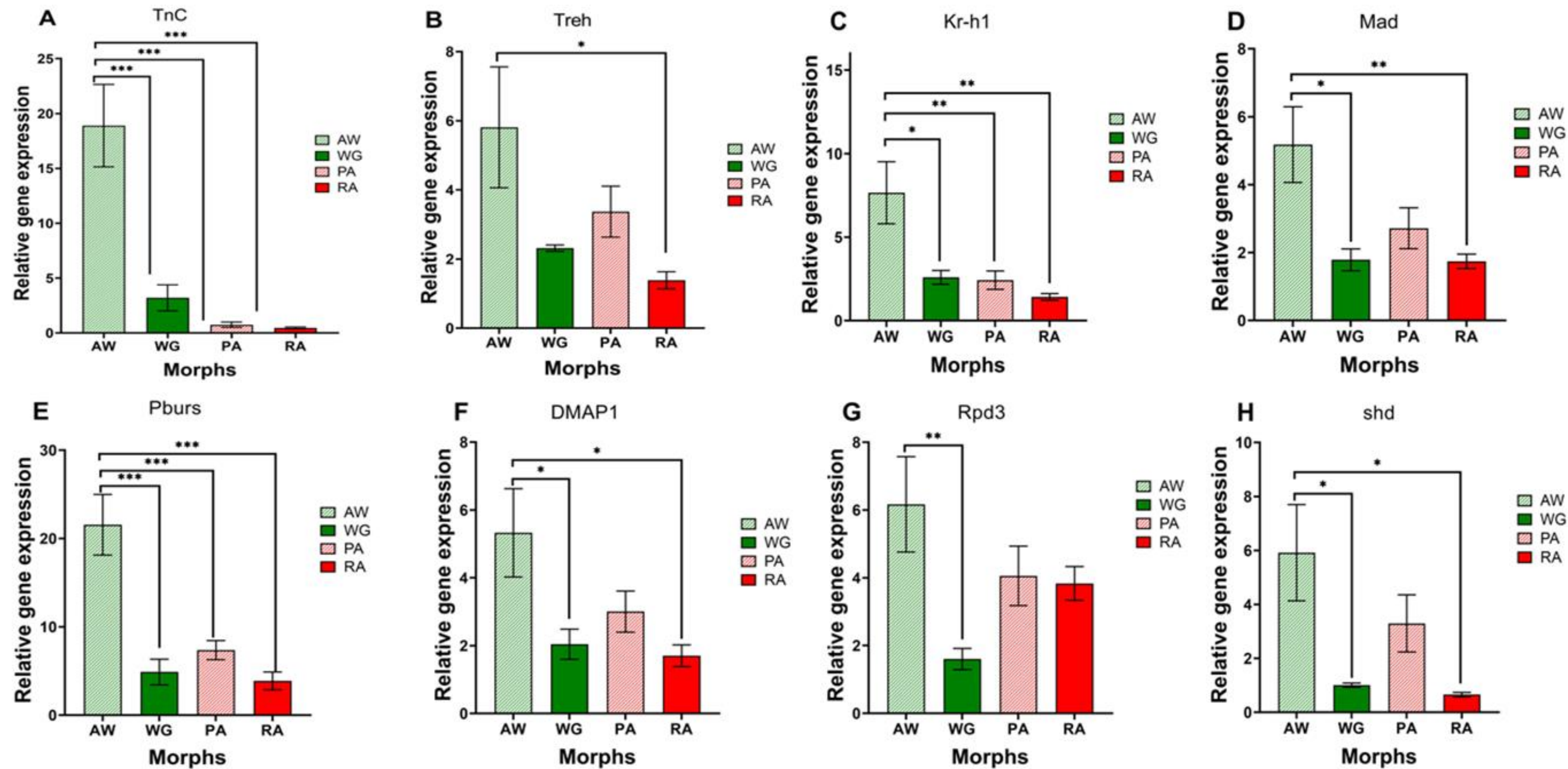


Figure 4.19. Validation through qPCR of RNA-seq differentially expressed genes between aphid morphs. AW= adult wing, WG= adult wingless, PA= pale adult, RA= red adult. Each of the bars represents the average measurement of expression from five biological replicate with error bars indicating the SEM. The Y-axis scale is specific to each gene. *P-values <0.05, **P-values <0.01, ***P-values<0.001 obtained from Tukey's post-hoc test. From left to right: Troponin C (TnC), Trehalase (Treh), Kruppel-homolog1 (Kr-h1), Mothers against decapentaplegic homolog 4 (Mad), Partner of bursicon (Pburs), DNA methyltransferase 1 associated protein 1 (DMAP1), Histone deacetylase Rpd3 (Rpd3), Ecdysone 20-monooxygenase(shd). The p-value, F-value and predictors of each model for the genes is tabulated in Table 4.10.

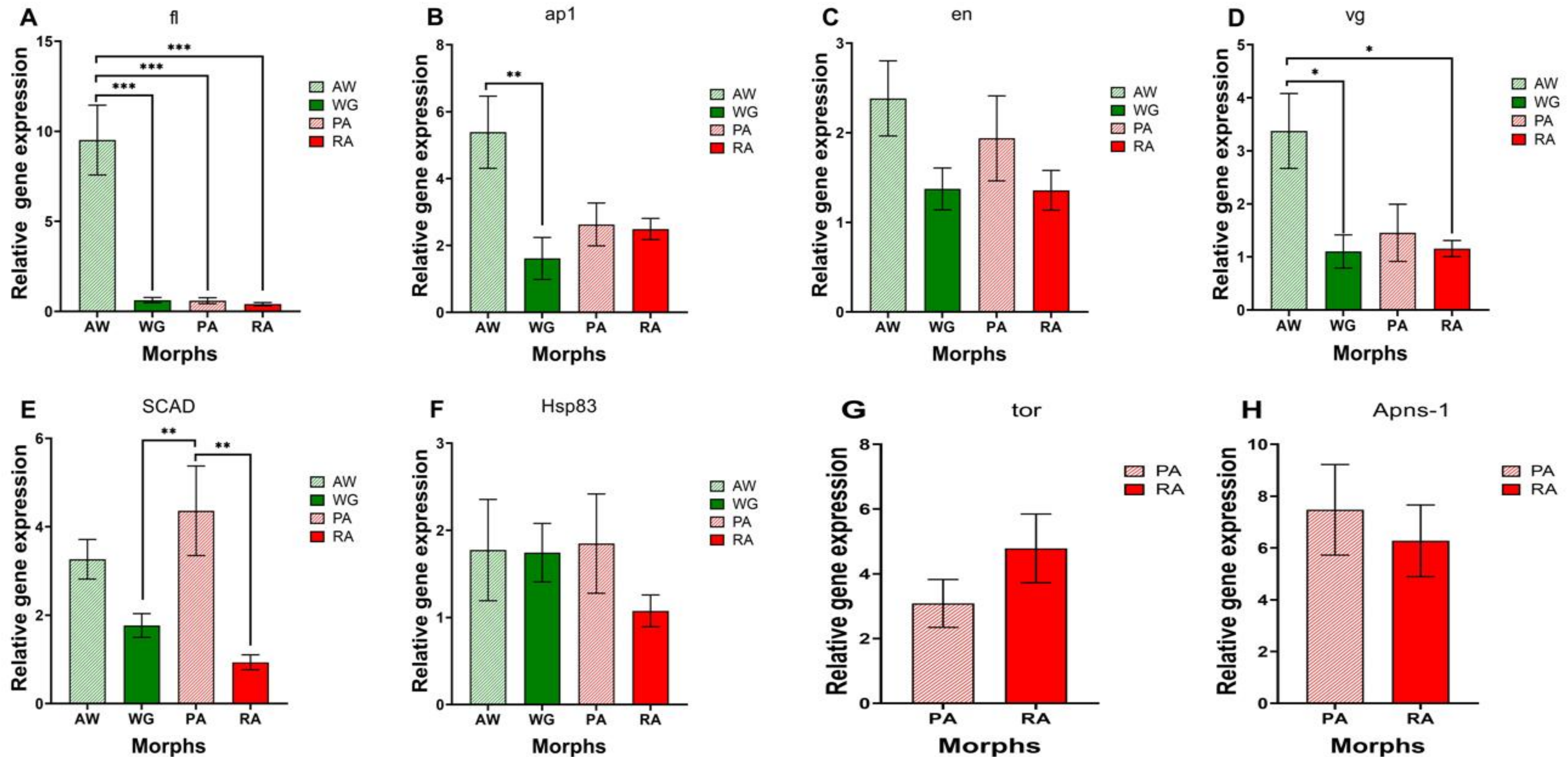


Figure 4.20. Expression level of candidate genes between aphid morphs. AW= adult wing, WG= adult wingless, PA= pale adult, RA= red adult. Each of the bar represents the average measurement of expression from five biological replicates with error bars indicating the SEM. The Y-axis scale is specific to each gene. *P-values <0.05, **P-values <0.01, ***P-values<0.001 obtained from Tukey's post-hoc test. From left to right: Flightin (fl), Apterous (ap1), Engrailed (en), Vestigial (vg), Short chain specific acyl-coA dehydrogenase (SCAD), Heat shock protein 83 (Hsp83), Carotene dehydrogenase (tor), Acyrthosiphon pisum desnovirus (Apns-1). The p-value, F-value and predictors of each model for the genes is tabulated in Table 4.10.

4.3.4 Protein expression

Protein samples from different aphid morphs were quantifiedy using western blotting with the Hsp70 antibodies. Although, all the morphs showed a higher protein expression for Hsp70 against N116 wingless morphs, there are no significant differences between the expression level between the morphs. (GLM, $F(2,9)=1.554$, $P=0.286$) (Figure 4.21).

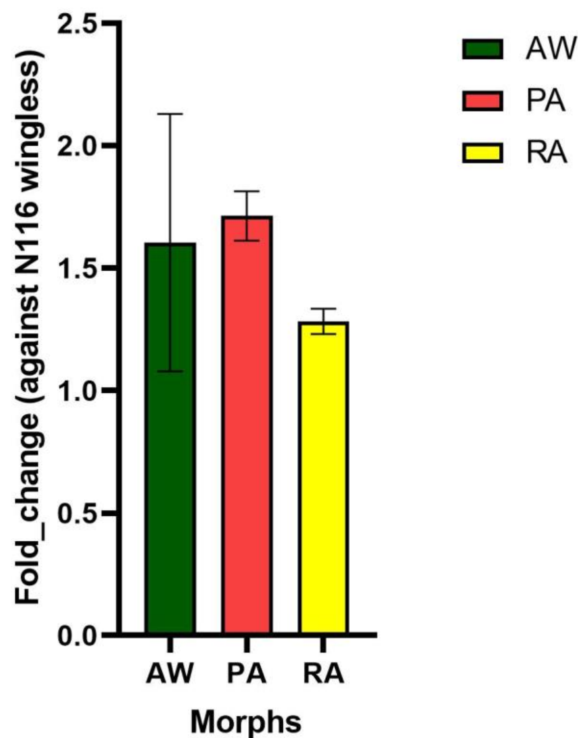


Figure 4.21. Protein level of Hsp70 (against N116 wingless). AW=N116 adult winged (green), PA=N127 pale aphid (PA), RA=N127 red aphid (yellow). Each of the bars represents the average measurement of expression from three replicates with error bar indicating the SEM.

4.4 Discussion

This chapter had several aims. First, we aimed to investigate the transcriptome profile between different pea aphid morphs and genotypes. Our RNA-seq suggest that many genes were differentially expressed between the alternative morphs in N116 genotype (N116winged vs N116 wingless). For examples, many genes that were upregulated in the N116 winged morphs compared to N116 wingless were involved in the wing formation and also neuroendocrine signaling, while downregulated genes were mostly involved in histone modifications and also metabolites. Further, our RNA-seq also reveals thousands of genes that were differentially expressed between the N127 genotype (N127pale vs N127 red) with most genes that were upregulated in the N127 pale morphs were involved trehalose and insulin production with most genes downregulated involved in glycogen fatty acid pathway and cuticular production. In contrast to the N116 alternative group comparison, most genes in the N127 alternative group comparison were involved metabolites rather than wing development. Next, our RNA-seq also reveals many genes that were differentially regulated between the alternative morphs from the two different genotypes (N127 pale vs N116). One of the major differences between the alternative morph in N127 and N116 is that N127 changes it body colour but does not produce wing while the N116 genotype usually produces wing. Our RNA-seq support this as most upregulated gene in the N127 were involved in metabolism, carotene production and down-regulated genes were involved in wing formation. Additionally, our RNA-seq also reveals many differentially expressed genes between the wild type of two different genotype (N127ered vs N116 wingless). Most of the genes that were upregulated in the N127 were involved in carotene production, histone modifications, cuticle formation. In contrast, most genes that were downregulated were involved in neuroendocrine signalling (ecdysteroid pathway) and wing formation. The neuroendocrine signaling are essential for wing triggering and wing development suggesting that these pathways might explain the different in response between the two aphid genotype wild type.

Apart from that, we found that 533 overlapping genes between the N116 winged vs N116 wingless group comparison and N127 pale vs N127 red comparison. Most of the overlapping genes between this two group comparisons has uncharacterized protein whereby the function of the gene is unknown. Next, a lot of the overlap genes were involved in metabolite regulation. For example, the facilitated trehalose transporter Tret1 is involved in the transfer of newly synthesized trehalose to the hemolymph. Trehalose has been shown to regulate the feeding behaviour of pea aphids suggesting this might be responsible for the different feeding behaviour observed in the pea aphid morphs (Wang et al. 2021). Next, the metabolite genes such as phosphoenolpyruvate carboxykinase (PEPCK) were also found to overlap across the group comparisons. PEPCK is usually involved in gluconeogenesis and has been reported to be differentially regulated between the alate and apterous morphs in brown citrus aphids (Shang et al. 2016). Shang also showed that inhibiting the PEPCK genes through RNAi resulted in underdeveloped and impaired flight function in the brown citrus aphid. Then, the G-couple receptor Mth2 were also found to overlap across different group comparison. Mth2 gene has been reported to be involved in increasing the lifespan of pea aphids during starvation (Li et al. 2014) and also involved in other crucial development such as growth, and immunity. Further, some of the overlapping genes were involved in extracellular communication function such as kelch-like protein (KLHL). Next, some other genes that were overlapped between these two groups comparison were involved in ecdysone signalling such as shd, EcR, Kr-h1. These genes plays important role in regulating the ecdysone signalling and juvenile hormone in pea aphid which has shown to be important in regulting wing development. Apart from that, Kr-h1 has always been shown to regulate oocyte reproduction in insects. Together, the results suggest that these genes might be important in regulating pea aphid morph development and reproduction.

My second aim was to investigate some the expression of aphid development gene such as (wing formation, stress response) based on previously published studies. Our results

suggest that most of the wing development genes were differentially expressed between aphid morphs and genotypes with the N116 winged having highest expression for wing development genes compared to other morphs. Lastly, I aimed to investigate the protein expression of the candidate gene Heat shock protein 70 (due to lack of availability of antibody for pea aphid). Although our RNA-seq showed that Hsp70 was differentially expressed between aphid morphs and genotype, our western blot analysis showed no difference in protein expression.

4.4.1 Gene expression

Gene expression was assessed by qPCR on whole adult aphids of each morph (N116 winged, N116 wingless, N127 red, N127 pale) from Chapter 3 with specific primers for genes involved in different aspects of aphid development. Pea aphid morphs showed differential expression on gene expression involved in wing formation, stress, response, carotene synthesis and ecdysteroid pathway.

fl

In insects, the muscle responsible for wing movements falls into two different categories known as the direct flight muscles (DFM) and indirect flight muscles (IFM). The flight capacity in the insect order Hemiptera, Hymenoptera, Diptera are usually powered by the asynchronous IFM. Flightin is a protein that is uniquely expressed in the IFM and was first reported in *Drosophila* (Vigoreaux et al. 1993). Flightin plays an essential role in maintaining the integrity of IFM sarcomere which is important for the assembly of thick filament and sarcomere integrity. Apart from that, fligthin is also important in regulating the myofilament stiffness and stretching of flight muscles (Henkin et al. 2004). In aphids, IFM helps move the wing indirectly through the deformation of thoracic exoskeleton (Crossley et al. 1978) that contains three different types of muscle known as dorsoventral muscles (DVM), dorsal longitudinal muscles (DLM) and oblique dorsal muscle (ODM) (Ishikawa et al. 2008, Ogawa

et al. 2012). The main function of DLM is to act as wing depressors (downstroke) that cause the stretching of DVM which results in a contraction that helps elevate the wings (upstroke) and stretching of DLM (Ogawa and Miura 2013). A recent paper by Chang et al. (2022) showed that knocking out flightin in pea aphids results in severely deformed wings. The most prominent changes in flightin knockdown aphid were observed in DLM where it becomes significantly wider and looser resulting in an abnormal flight apparatus where the wings were tilted to one side. The data presented here further support the role of flightin in aphid wing development whereby the N116 winged aphid has the highest expression in comparison to all other morphs.

ap1

Apterous (ap) is a LIM-homeodomain protein that is first reported to play a role in regulating wing development in *Drosophila*. Ap is usually expressed in the dorsal cells of *Drosophila* wing disc and plays an essential role in the establishment of the dorsal-ventral boundary during wing formation in *Drosophila*. In the dorsal-ventral boundary of *Drosophila*, ap1 helps activate notch signalling, which then induces the expression of the wingless gene (wg) which in turn activate the vg expression at different threshold (Tomoyasu et al. 2009). The importance of ap in wing development is shown by Cohen et al. (1992), whereby *Drosophila* with ap1 knockout showed a severe defect in wing and haltere imaginal disc development. Recently, Brisson et al. (2010) reported wing development gene homologs network in aphids deduced from the wing development genes in *Drosophila melanogaster.* Interestingly, in aphids there are two copies of ap gene (ap1 and ap2) in comparison to *Drosophila* and *Apis mellifera,* which only had one copy. Brisson reported only significant gene expression of ap1 in the 1st and 2nd nymphs from crowded and non-crowded mothers. The data presented here showed that ap1 also showed a significant morph effect in adult aphids with N116 winged morphs having the highest expression in comparison to the other morphs. One of the explanations is that winged-destined aphid and

wingless aphid nymphs are physically impossible to be distinguished between the 1st and 2nd nymph stage, therefore the sample in Brisson et al. (2010) might have contained the opposite morphs which would result in the underestimate of the true-difference in the expression of the genes. Apart from that, by using RNAi in early last larval stage of *Tribolium castaneum*, Tomoyasu et al. (2009) showed that the formation of the hindwing in the beetle was repressed. This suggests that the expression of ap during insect development might vary depending on the species. In aphids, the wings are only fully formed once the 4th instar stage moult to adult stage, therefore the expression of ap1 is still highly expressed at the adult stage in the winged morph in comparison to the other wingless morphs (Song et al. 2018). Our results are in contrast to previous results that showed lower expression of ap1 in adult macropterous strain (long-winged) of planthopper in comparison to branchypterous strain (short-winged) (Liu et al. 2015). However, the difference in expression could be due to adult branchypterous strain planthopper still have wings while in wingless aphid does not have any wings as the wing bud has been completely degenerated by the 2nd nymph instar stage.

en

In the aphid wing gene network deduced from *Drosophila*, the gene that orchestrates the anterior-posterior (AP) compartment in wing development is known as engrailed (en). En is an important transcription factor that is involved in wing patterning and neurodevelopment in *Drosophila* (Morata and Lawrence 1975). One of the main functions of en is to help establish and maintain the P cell identity. Apart from that, en has also previously been described to regulate the expression of other AP genes such as decapentaplegic (dpp) and hedgehog (hh) (Solano et al. 2003). The data presented here suggest that there is no morph effect for engrailed expression in the adult stage. Our data is in addition to the results from Brisson et al. (2010) which showed no morph effect for engrailed expression during the embryo stage to the 4th instar nymph stage. Interestingly, our results disagree with a

previous study by Zhang et al. (2019). Zhang found that en is differentially expressed between the bird cherry-oat aphid with higher expression of en in the wingless morphs in comparison to the winged morph. Dufour et al. (2020) showed that en only affects wing morphology in *Drosophila* at a critical time point whereby loss or gain of function of en after this time point does not have any effect on wing morphology. Apart from wing development, engrailed is also involved in neurodevelopment. In *Drosophila*, overexpression of engrailed results in collapsed motor neurons and thereby affecting the central nervous system (Siegler and Jia 1999). However, the role of engrailed in aphid neurons remains widely unknown. Together with the results above, it is possible engrailed expression is lowered in the winged aphid after the critical timepoint in wing development to avoid any negative effect on the neuron development in aphids.

vg

One of the most well-known and studied genes in *Drosophila* wing development is known as vestigial. Vg is often referred to as a 'master gene' in aphid wing development whereby it can trigger and initiate wing development regardless of the tissues where it is expressed (Baena- López and García-Bellido 2003). Vg has been reported to interact with other wing development genes such that the expression of vg at wing margin were controlled by wg and Notch while in the proximal part of wing blade, the expression of vg is regulated by itself and another gene dpp (Williams et al. 1994; Kim et al. 1996). Next, vg is also able to form a complex with the gene scalloped (sd) to regulate downstream genes involved in wing development. Even though the role of vg in other insects remains understudied, the role of principal wing development homologs remains highly conserved across different insects. Simmonds et al. (1997) reported that *Drosophila* with vg knockout showed a severe defect in wing development whereby the wing and haltere imaginal disc fail to proliferate normally. Next, a similar observation in wing deformity is also observed in the red flour beetle *Tribolium castaneum* when vg is inhibited (Courtney et al. 2013). Further, a recent study by

Zhang et al. (2021) showed that green peach aphids with vg suppress through RNAi showed severe defects in the wing apparatus Our results showed significant morphs effect of vg expression with N116 adult winged having highest expression level in comparison to other morphs. Our result is the first to report a difference in vg expression between winged morph and wingless morph and agrees with a previous study that reported higher expression of vg in winged bird-cherry oat aphid (Fan et al. 2020). Apart from that, Fan also reported the highest expression of vestigial in bird-cherry oat aphids during the 3rd instar stage. In pea aphids, Brisson et al. (2010) investigated 11 wing development genes across different aphid instar stages and morphs and reported no significant difference. However, the expression level of vg was not investigated in the paper, therefore it would be interesting to look at the expression level of vg between the early stage of winged-destined nymphs and wingless nymphs to determine if a similar pattern is observed. Together, the results suggest that vg might play a vital role in pea aphid wing development.

SCAD

SCAD is a "mitochondrial-like" gene that is involved in energy production specifically in fatty acid/ lipid metabolism. In insects, lipids are mainly stored in the fat body as triacylglycerol and they are hypothesised to be the main fuel used during flight. The content of triacylglycerol has been reported to be higher in multiple winged insects' species such as long-winged planthopper (Itoyama et al. 1999), crickets (Zera et al. 1994) and also aphids (Xu et al. 2011). Interestingly, our results showed that N127 pale aphid has the highest expression for SCAD in comparison to all the other morphs. Although not significant, SCAD also showed a higher expression in N116 winged morphs compared to the N116 wingless morphs. Our results indicate that aphids use lipids as the main fuel for dispersal rather than sugar or glycogen (Yang et al. 2014). The higher expression of SCAD in N127 in pale aphids agrees with previous studies that reported fatty acid oxidation level increase when insects are exposed to starvation stress (McCue et al. 2015). In aphids, the lifespan of

winged aphids is usually longer in comparison to the wingless aphid (Li et al. 2016). However, the difference in the lifespan of pale vs red aphids remains unknown. A previous study by Lee et al. (2012) reported an increased level of genes involved in fatty-acid oxidation in *Drosophila* when exposed to starvation. The overexpression of fatty-acid genes in *Drosophila* helps increase the stress tolerance to oxidation and starvation stress. Therefore, it might be possibile that the higher expression of SCAD in N127 pale aphids observed in our experiment could be related to increasing lifespan and starvation tolerance. However, further research on the lifespan and starvation tolerance differences in N127 pale with other morphs is needed to confirm this speculation.

Hsp83

Heat shock proteins (Hsps) are molecular chaperones that play an important role in preventing protein misfolding and denaturation resulting from environmental stressors such as heat, oxidative stress and others (Miao et al. 2020). The ecological role of Hsp suggests that their expression is triggered by environmental stressors and could play a role in influencing fitness parameters of organisms such as lifespan, fecundity and development (Freitak et al. 2012). The homolog of Hsp83 known as Hsp90 has shown to be involved in different developmental processes in insects such as spermatogenesis, oogenesis and embryogenesis (Song et al. 2007). A recent paper by Will et al. (2017) reported the pleiotropic roles of Hsp83 in pea aphid fecundity, longevity and embryogenesis whereby aphids with a reduction in Hsp83 showed lower survival and fecundity. Our results showed that there is no significant difference in Hsp83 expression levels between all the morphs. This suggests that the lower fecundity in dispersal morphs might be regulated by other genes or pathways rather than Hsp83. Apart from that our results also agree with a previous study that showed that crowding does not induce any expression of large heat shock genes (Hsp40, Hsp90) (Chapuis et al. 2011). Since large Hsps are constitutively expressed it is plausible that they might still be involved in resistance to stress when exposed to crowded

conditions without altering their expression. Next, a recent study has shown that Hsp83 plays an important role in caste-specific ovary development in honeybees with the queen ovary showing significantly higher expression of Hsp83 in comparison to the worker ovary (Lago et al. 2016). In our experiment, the whole aphid is used which could result in underestimation of the expression of the Hsp83 since Hsp83 is ubiquitously expressed across tissues. Therefore, further study of the expression of Hsp83 in the ovary of dispersal and non-dispersal aphid morphs would be needed to determine their role in morph-specific ovary development in the aphid. Lastly, in our experiment two different aphid genotypes N116 green and N127 red aphid were used. One of the main differences is that N127 has a lower fecundity compared to N116. Although we did not find a significant difference in Hsp83 between the N116 wingless and N127 red, our data showed that N127 has a lower expression. In *Drosophila* a mutant strain with lower Hsp83 reduction showed less fecundity compared to the normal flies (Chen et al., 2012) Therefore, future work on manipulation of Hsp83 expression would be needed to determine if Hsp83 could be a potential gene that regulates the differences of fecundity observed between these two aphid genotypes.

tor

Carotenoids are pigments that play an important physiological function in many organisms. Carotenoids are usually biosynthesised in photosynthetic organisms such as plants, and algae with most animals unable to synthesize carotenoids. Surprisingly, carotenoid biosynthesis gene is found in a few insect species such as aphids resulting in a range of body colour (Takemura et al. 2021). In aphids, the carotenoids torulene is responsible for the red pigmentation and is only found in the red morphs, while the remaining carotenoid such as β- carotene is observed across aphid of different body colour morphs (Moran et al. 2010). We investigate the carotene biosynthesis gene tor and find no significant morph effect between the pale and red morph and no expression in N116 (winged and wingless) since they do not express this gene. A study by Wang et al. (2019) has suggested that the

changes from red to pale morph when exposed to starvation is a result of the breakdown of the red pigment to increase energy reserve. The fact that we found no significant difference in the tor gene might indicate that the red pigmentation is still being synthesised by aphids but is subsequently broken down to increase their energy reserve in starvation conditions. Further research on other genes involved in carotenoid process is needed to determine if changes from red to pale morph is due to higher expression in other carotenoid gene resulting in paler colour or due to a higher breakdown rate of carotenoid pigments in starvation conditions.

Apns-1

Densovirus falls under the family of Parvoviridae and is a type of virus that possesses a single-stranded DNA genome often found in arthropods (van Munster et al. 2003). The role of densovirus in aphid phenotypic plasticity was first reported in the rosy apple aphid (Ryabov et al. 2009). Ryabov found that rosy apple aphids infected with densovirus were able to produce winged morphs compared to virus-free clones. In our experiment, we investigated the expression of this gene in two different genotypes (N116 and N127) that showed differences in response to crowding. Our data presented showed that there is no significant expression of APNS between the N127 pale and N127red and the N116 genotype showed no expression for this gene. Our results disagree with a previous study by Parker et al. (2019) that reported the ability to trigger wing formation in different pea aphid genotypes is dependent on the densovirus gene (Apns-1) and (Apns-2) expression with highly inducible genotype having a higher expression of Apns. One possible explanation for the observation in our data could be due to the fact that whole organisms were used, which could underestimate the expression of this gene as Parker et al. 2019 reported that Apns-1 is more highly enriched in aphids' heads in comparison to the whole body.

TnC

Myofilament is an important protein in the flight muscle conserved in both invertebrates and vertebrates. Myofilament usually contains three main proteins known as the F-actin activated myosin motors, thin filament regulatory proteins tropomyosin and troponin (Tn) (Cao and Jin 2020). Tn complex is made up of three different protein subunits known as the inhibitory subunit troponin I (TnI), thin filament anchoring subunit troponin T (TnT) and the Ca^{2+} binding subunit troponin C (TnC). Our data showed that TnC showed is highly expressed in the N116 winged morphs in comparison to the other morphs. Our results agree with the previous study which showed that TnC plays an essential role in insects flight whereby *Drosophila* with the TnC knockout showed a defect in flying and result in flightless phenotypes (Chechenova et al.2017). Apart from that, a study in planthoppers also showed that long-winged planthoppers showed higher expression for TnC in comparison to short-winged planthopper which further support that TnC is essential for flight and its function is also conserved in aphids (Gao et al. 2019). TnC also plays important role in regulating insect fitness apart from its function in flight muscle contraction. A recent study by Lan et al. (2018) showed that rice dwarf virus reduced the survival and fecundity of its host leafhopper *Nephotettix cincticeps* by downregulating the expression of TnC. Therefore, the higher lifespan observed in the winged morphs might be a result of the upregulation of the TnC expression. Further research on the manipulation of TnC in different morphs is needed to confirm its role in survival in aphid morphs.

Treh

Trehalose is a non-reducing disaccharide that plays a major role as energy resource in insects. The hydrolysis of trehalose to produce glucose molecules is regulated by an enzyme known as Trehalase (Treh). Studies have shown that inhibition of trehalase activity influences insect development such as flight capacity, and chitin synthesis (Silva et al. 2004; Shukla et al. 2015). We found Tre at a significantly higher level in winged adults

compared to N127 red aphids. Even though Tre is not significant between N116 winged and wingless, the winged morphs still showed higher expression. The role of Treh in winged development has been studied in several insect species. For example, Li et al. (2019) reported wing deformity in *Harmonia axyridis* when Treh is silenced. Next, Zhang et al. (2017) also reported similar wing deformity in *Nilapavarta lugens* when Treh is inhibited. Apart from wing deformity, Zhang found that the expression of the wing development gene ap is significantly reduced after the injection of dsTreh. This suggests that Treh might play a role in regulating wing development genes and therefore affecting wing development in insects. Further studies on silencing Treh in pea aphids are needed to shed light on the role of Tre in aphid wing development.

Kr-h1

In insects, two hormonal systems known as juvenile hormone (JH) and 20-hydroxyecdysone (20E) play an essential role in insects' metamorphosis and development. Kr-h1 is a juvenile hormone response transcription factor that was first discovered in *Drosophila* (Schuh et al. 1986). The expression of Kr-h1 varies between insects for e.g. in B.mori the level increased during larval stages and disappeared during the final larval instar star. Next during the prepupal stage Kr-h1 reappears and completely disappears again in pupae. At the end of the late pupal stage, the expression increased again and was maintained at an elevated level during adult stage (Kayukawa et al. 2014). In hemimetabolous insects, the level usually increased in early nymphal stage followed by complete disappearance for several days and reappears after the adult has moulted (Konopova et al. 2011; Liu et al. 2018). Our data showed that Kr-h1 is highly expressed in the N116 winged morph in comparison to other morphs during adult stage. To our knowledge, this is the first-time differences in Kr-h1 have been reported in aphids.

In brown planthoppers, the disruption of Kr-h1 expression during the critical wing formation stage (3rd and 4th instar) results in adults with stunt wing development and malformation

in genitalia (Jin et al. 2014). In aphids, the critical stage of winged development happens during the 2nd instar whereby wing buds from wingless nymph degenerated while the winged-destined nymph wing buds continue to grow. However, the wing formation in winged morphs is not completely formed until they reached the adult which could possibly explain why we observed the highest expression of Kr-h1 in N116 winged morph. Further support for the role of Kr-h1 in wing development is also observed in the German cockroach *B. germanica* whereby roaches with reduced expression of Kr-h1 showed deformity in wing structure such that the wing pad does not expand perfectly. (Lozano and Belles 2011). Next, Strassburger et al. (2021) also reported defects in wing growth and formation in *Drosophila* with reduced Kr-h1 levels. Interestingly, Kr-h1 also plays a role in the ecdysone hormone pathway whereby it represses the insect ecdysone through regulating primary genes involved in ecdysone response such as ecdysone receptor (EcR), BR-C and ecdysone-inducible protein E93 (Liu et al. 2018). Apart from that, Kr-h1 also acts on steroidgenic enzymes such as Spok to reduce ecdysone synthesis (Zhang et al. 2018). It is important to note that, although Kr-h1 represses the gene involved in ecdysone biosynthesis, the expression of Kr-h1 is also inducible by 20E in a tissue-specific manner. Next, some studies have shown that Kr-h1 can regulates oocyte production in insects. (Yue et al. 2018; Tang et al. 2020; Hu et al. 2020). For example, Hu et al. (2020) show that knocking out Kr-h1 in insects leads to a reduction in oocyte production However, contradicting results have been reported whereby knocking out Kr-h1 does not affect insect ovarian development (Smykal et al. 2014; Gujar and Palli 2016). Taken together, the results from these studies suggest that the role of Kr-h1 might be species-dependent and reveals the importance and a complex crosstalk between the JH-20E pathways in insect development. In aphids, the function of Kr-h1 remains widely unknown and future studies on manipulation of the expression level of Kr-h1 across different developmental stages and tissues could shed light on wing development and reproduction in different aphid morphs.

Mad

Next, dpp is another gene that is involved in morphogenesis and has been well characterised in *Drosophila* (Teleman and Cohen 2000; Zecca and Struhl 2021). In *Drosophila*, dpp is expressed in a precise gradient to regulate multiple aspects of *Drosophila* such as cell morphogenesis, cell survival and cell differentiation during wing development (Teleman et al. 2000; Akiyama et al. 2008; Zhang et al. 2013). Dpp usually binds to the gene Thickviens (Tkv) and forms a complex known as the Dpp-Tkv complex (Nahmad et al. 2008). The complex will further trigger phosphorylation in Mad which in turn acts as the transcription activator of dpp (Nahmad et al. 2008). Here we investigate another gene Mad that is hypothesised to mediate dpp signalling and therefore influences the wing development and insect ecdysis process (Tanimoto et al. 2000; Santos et al. 2016). We found that Mad is highly expressed in the adult winged morph compared to the other morphs. This is in accordance with a study by Sekelsky et al., 1994 which shows that *Drosophila* lacking Mad gene shows a defect in wing development corresponding to the dpp mutant phenotype. Next, Santos et al. (2016) also showed that german cockroaches that had a depleted level of Mad showed defects when molting to adult stages whereby the cockroach showed an unextended wing. Other studies on RNAi of Mad in the cricket *Gryllus bimaculatus* also reported defects in wing extension (Ishimaru et al. 2016). Above we discussed the role of Kr-h1 in wing development; interestingly a recent study by Ishimaru et al. (2019) showed that Mad might play a role in regulating the expression of Kr-h1 whereby Mad depleted cricket showed a lower level of Kr-h1. Kr-h1 is involved in wing and ovary development through regulating the 20E induced Broad complex BR-C (Kayukawa et al. 2014). Therefore, it is possible that there is a crosstalk between the Mad and Kr-h1 expression in regulating wing polyphenism in aphids. Our results above also showed that the level of Kr-h1 is also highly expressed in the N116 winged morph in comparison to the other morphs. In aphids, the role of Mad remains unknown and future studies on

manipulating its expression and the gene it interacts with might shed light on the wing development in aphids.

Pburs

In insects, molting is an important process that allows them to shed and replace their exoskeleton as they grow (Luo et al. 2005). The neurohormone required in this process is known as bursicon and is encoded by the gene burs and the newly identified homologous gene known as partner of bursicon (Pburs) (Luo et al. 2005). Burs are important in the synthesis of (burs-α) protein, while Pburs are responsible for encoding the (burs-β) protein. Burs-α and burs-βprotein will dimerize to form the bursicon neurohormone (Melnattur et al. 2020). In our study, we found that Pburs was significantly highly expressed in the N116 winged morphs in comparison to the other morphs. Our results are in accordance with other studies that found that *Drosophila* mutants without the Pburs gene showed a defect in wing expansion (Lahr et al. 2012). Next, Peabody et al. (2008) showed that the alteration of the neurohormone burs in *Drosophila* results in failure of wing expansion. Another study by Arakane et al. (2008) showed that the reduction of Pburs and burs in red flour beetles results in incomplete expansion of the wing. Apart from wing development, Pburs also plays an important role in cuticle sclerotization. Sclerotization is an important process that results in the hardening and stiffness of tissues. In winged aphids, the head and thorax are heavily sclerotized in comparison to other morphs. Taken together, Pburs might play an essential role in wing development and also cuticle sclerotization in aphids and future research on manipulation of expression Pburs is needed to confirm this.

DMAP1

DNA methyltransferase 1 plays an important role in epigenetic regulation. The activation of DNMT 1 is regulated by the DNA methyltransferase 1 associated binding protein 1 (DMAP1) such that knocking out the gene will result in hypomethylation (Lee et al. 2010).

Interestingly, our results showed that N116 winged aphid had the highest expression of DMAP1 in comparison to all the other morphs. The function of DMAP1 has been widely studied in mammals whereby DMAP1 knockout mice showed severe embryonic lethality (Mohan et al. 2011) while the role of DMAP1 in insects remains largely unknown. Recently, a study has found that DMAP1 might play a role in insect reproduction (Gegner et al. 2020). Gegner showed that silencing of the DMAP1 gene in the harlequin ladybird *Harmonia axyridis* leads to a reduction in fertility and fecundity. Next, Li et al. (2021) showed that DMAP1 might also play a role in thermal tolerance in the predatory mite *Neoseiulus barkeri* whereby silencing the expression of DMAP1 in high temperature acclimated strains (HTAS) of predatory mites results in decreased thermal tolerance and increase mortality rate. Apart from that, a study by Xu et al. (2020) showed that silencing the expression of Dnmt1 affects the wing formation in the silkworm *Bombyx mori*. Our results are the first to report such differences in DMAP1 expression between aphid morphs. Taken together these studies and our findings, it is possible that DMAP1 might play an important role in regulating wing development and also fecundity in aphids.

Rpd3

Apart from DNA methylation, histone acetylation is another epigenetic mechanism that plays an important role in eukaryotes. In eukaryotes, histone acetylation is usually involved in gene regulation, cell signalling and metabolism regulation (Ali et al. 2018). The process of histone acetylation is tightly regulated by the activities of histone acetyltransferases (HATs) and histone deacetylases (HDACs) (Kirfel et al. 2020). Rpd3 is one of the major histone deacetylases that primarily target H3 and H4 (lysine 5 and 12). Ehrentraut et al. (2010) showed that reducing levels in Rpd3 will lead to an increase in global acetylation. In our study, we found that the histone deacetylases Rpd3 (Rpd3) are highly expressed in the N116 winged morphs in comparison to the other morphs. A previous study by Nakajima et al. (2016) reported that Rpd3 helps increase starvation resistance in *Drosophila*. In aphids,

the winged morphs are usually more resistant to starvation stress in comparison to the wingless morphs, suggesting that Rpd3 could also be regulating the starvation tolerance in aphids.

Rpd3 is also involved in regulating lifespan in *Drosophila* (Kopp and Park 2019). Kopp showed that flies with a reduced level of Rpd3 have a longer lifespan. Our study agrees with the results as winged morphs usually have a longer lifespan in comparison to wingless morphs. Apart from regulating lifespan, another study by Kirfel et al (2020), has shown that inhibiting Rpd3 results in a reduction in the number of viviparous offspring and the number of premature nymphs in pea aphids. However, our results disagree with this study as we found that N116 winged morphs have the highest expression in comparison to the other morphs. In addition to the role mentioned above, Rpd3 has also been shown to be involved in wing formation in *Drosophila*. In *Drosophila*, the dpp gene plays an important role in wing development. Zhang et al. (2013) showed that in *Drosophila* the gene Atrophin (Atro) can recruit the Rpd3 to form an Atro-rpd3 complex which then binds to the dpp locus and reduce dpp transcription through histone acetylation. Lastly, Rpd3 also seems to be involved in regulating caste-specific behavior. Simola et al. (2016) showed that major workers ants with reduced Rpd3 levels exhibit minor worker-specific behavior. Together this study shows a remarkably diverse function of Rpd3 in insects and future studies in manipulating Rpd3 expression are needed to elucidate the function of Rpd3 in aphids.

shd

20-hydroxyecsdyone (20E) is a type of insect molting hormone and plays an essential role in insect molting process and development. In most insects, the prothoracic gland usually secretes the ecdysone (Lafont et al. 2012; Qu et al. 2015). The ecdysone will be further hydroxylated to 20E by the enzyme ecdysone 20-monooxygenase (shd). In our data, we found that the expression of shd is highest in the N116 winged morph in comparison to the other morphs. Recent studies have reported that knocking down shd in the small brown

planthopper *Laodelphax striatellus* leads to delayed development and also nymphal lethality (Jia et al. 2013). Similar results have also been observed in the silkworm *Bombxy mori* when shd is knocked down. (Fujinaga et al. 2020). Further, a recent study by Liu et al. (2020) reported higher expression of shd in the high fecundity genotype of brown planthoppers *Nilaparvata lugens* in comparison to the low fecundity genotype suggesting that ecdysone might also play a role in regulating insects' fecundity. Apart from that ecdysone also plays an important role in butterfly eyespot plasticity; a study by Tian and Monteiro (2022) reported higher expression of shd in butterflies producing larger eyespots suggesting that shd helps induce the production of 20E which in return regulates eyespot plasticity. In aphids, ecdysone plays an important role in transgenerational wing polyphenism (Vellichirammal et al. 2017). Vellichirammal reported that knocking down the gene involved in ecdysone signalling resulted in an increase in winged progeny production while increasing 20E resulted in more wingless progeny. In the pea aphid, the winged morph usually produces higher wingless offspring in comparison to the wingless morph; it is possible that shd plays a role in this through increasing the expression of 20E. Apart from that, a study has shown that English grain aphid *Sitobion avenae* feeding on resistant plants showed downregulation of shd (Lan et al. 2012). This suggests that ecdysone signalling might have a role in regulating aphid feeding bevaiour. Taken together, ecdysone seems to play an essential role in insect development and further studies on manipulating ecdysone through manipulating shd expression are needed to elucidate their role in other aspects of aphid development such as fecundity and embryogenesis.

4.4.2 Transcriptome profile expression

N116 winged vs N116 wingless

The transcriptomic difference between N116 winged and N116 wingless was examined using RNA-seq to explore the underlying molecular mechanisms in regulating wing polyphenism, reproduction and behavior of winged morphs. We found 2308 genes that were differentially expressed between winged and wingless morphs. Next, we filtered out 44 genes from 2308 genes based on their function in different aphid development. 19 of the 44 were upregulated in the N116 winged morph. For example, we found some genes that are involved in wing formation such as fl, TnC, Mad, Pburs were upregulated in the winged morph in comparison to the wingless morph. The function of fl, TnC, Mad, Pburs on wing development has been discussed above. Apart from the wing development genes, we found that some chemosensory genes such as ORF2, and ORF4 are highly expressed in the winged morphs. The chemosensory system in insects plays an important role in plant selection in herbivorous insects. Our data agree with other studies that found higher expression of the chemosensory genes in cotton aphid *Aphis gossypii* (Peng et al. 2020). During dispersal, winged morphs usually locate the new host plant through volatile cues. Therefore, the winged morph may require a more fully developed chemosensory system to detect these volatile cues. Apart from that, winged morphs are usually more resistant to starvation and also have longer lifespan compared to the wingless morph. The N116 winged morphs showed higher expression such as apolipoprotein D (ApoD), and G-protein coupled receptor Mth2 which has been shown to extend lifespan and survival during starvation in other insects (Walker et al. 2006; Li et al. 2014). Apart from the wing development and hormone gene discussed in section (6.4.2), some other wing-related and hormone genes such as abnormal wing disc 2 (Awd2), p38 mitogen-activated protein kinase (p38) and eclosion hormone-like (EH) were also upregulated in winged morph. These genes have been reported to play an important role in moulting process and wing development (Adachi-Yamade et al. 1999; Romani et al. 2017).

In contrast, 25 of the 44 genes were downregulated in the winged morphs (Supplementary). For example, phenoloxidase 1 (PO) was downregulated in the winged morphs. PO is an

important enzyme and plays an essential role in insect immunity whereby reduced expression results in lower immunity toward fungal and bacterial infection (Ma et al. 2020). In aphids, winged morphs usually have lower immunity toward pathogens REF compared to wingless aphids therefore the phenoloxidase gene might be important in regulating immunity in insects. Further, many genes involved in epigenetic modification such as DNA N6-methyl adenine demethylase (DMAD), histone-lysine N-methyltransferase E(z) E(z), histone acetyltransferase p300 (EP300) were also downregulated in the winged morphs. These genes are involved in regulating chromatin accessibility, histone modification and methylation process (Maksimoka 2014; Stepanik and Harte 2012; Yao et al. 2018). These studies suggest that epigenetic regulation might play an important role in regulating wing polyphenism in pea aphids.

Next, most DEGs between N116 winged and N116 wingless aphids were distributed to GO terms such as transmembrane transport, proteolysis, structural constituent of cuticle and transmembrane transport activity and similar units were observed in KEGG such as transport and catabolism, membrane transport, signalling molecules and interactions and amino acid metabolism. The results suggest that these functional terms and units might be important in regulating aphid wing polyphenism and the genes that fall under these terms may be the key that underlies the difference between N116 winged and N116 wingless morph.

N127 pale vs N127 red

The transcriptomic difference between N127 pale and N127 red was examined using RNA-seq to explore the underlying molecular mechanisms that regulate the red and pale morphs in aphids. We found 2076 genes that were differentially expressed between pale and red morph. Next, we filtered out 40 genes from 2076 genes based on their potential function in different aphid development (e.g. stress response, neuroendocrine signalling, wing formation, metabolite regulation and carotene production). 23 of the 40 genes were

upregulated in the pale morphs. For example, a lot of the genes that were involved in lipid and insulin regulation were upregulated in the pale morphs. For example, lipid storage droplets surface-binding protein 1 (LSD-1), short-chain specific acyl-CoA dehydrogenase, mitochondrial (SCAD), insulin receptor substrate 1 (IRS), and insulin-like peptide receptor (IGF). These genes play a major role in regulating metabolism such as carbohydrates, lipids and were shown to be upregulated in insects during starvation stress (Xu et al. 2015, Wood et al., 2018). Our result also agrees with a previous study by Tabadkani et al. (2013) that showed a big difference in metabolite levels between the red and pale aphids. Next, similar to the N116 winged morphs, N127 also has increased resistance to starvation in comparison to the red morphs. N127 pale morphs also showed an increase in genes such as apolipoprotein D (ApoD), and forkhead protein O (FoxO). These genes have shown to be upregulated in insects during starvation and are essential for survival during these conditions (Kramer et al. 2008; Johnson and Stolzing 2019).

In contrast, 17 genes were downregulated in the pale morph (Supplementary). For example, succinate dehydrogenase (FH), isocitrate dehydrogenase (OXCT1), pyruvate dehydrogenase E1 component subunit beta, mitochondrial (PDHB), probable phosphoglycerate kinase (PGK), were all downregulated in the pale morphs in comparison to the red morphs. All of these were involved in the glycolysis-tricarboxylic acid (glycolysis-TCA) pathway which is important for energy metabolism. Our findings support the observation from previous studies whereby starvation in the cotton leafworm *Spodoptera litura* shows similar expression profiles in these genes (Hu et al. 2016). Next, genes involved in carotene production such as phytoene desaturase and phytoene synthase were also differentially expressed which could be responsible for the difference in body colour observed between the two morphs. Lastly, genes that are involved in cuticle formation such as RR1 cuticle protein 4 (RR1), and cuticular protein like-precursor were downregulated. A similar observation in these cuticle genes expression was reported in soybean aphids when exposed to starvation (Enders et al. 2014). It is possible that the reduction of expression of

CP was metabolically costly and the reduction in this protein production is related to energy conservation during environmental stress.

Next, most DEG between winged and wingless aphids was distributed to GO terms such as transmembrane transport, proteolysis, glucuronosyltransferase activity and transmembrane transport activity and similar units were observed in KEGG such as transport and catabolism, carbohydrate metabolism, metabolism of other amino acids and xenoiotics biodegradation and metabolism. The results suggest that these functional terms and units might be important in regulating differences in aphid locomotion and body colours and the genes that fall under these terms may be the key that underlies the difference between N127 pale and N127 red morphs.

N127 red vs N116 wingless

Two polymorphic pea aphid genotypes (N116 and N127) were used in the experiments, and they respond very differently to environmental stress. The transcriptomic difference between the N127 red and N116 wingless was examined using RNA-seq to explore the underlying differences in these two genotypes in their wildtype forms. We found 5108 genes that were significantly differentially expressed between N127 red wingless aphid and N116 green wingless aphid. 50 differentially expressed genes were filtered out from the list of 5108 genes. These genes were involved in carotene production, reproduction, metabolism, longevity and cuticle formation. 28 of the 50 genes were upregulated in N127 red morphs. For example, protoheme IX farnesyltransferase, mitochondrial-like, partial (Cox10), carotene dehydrogenase (tor) was upregulated in red morphs. Our result agrees with the study by Zhang et al. (2018) which reported a similar expression profile of these genes between red and green pea aphids of different genotypes. tor is essential in producing the carotenoid pigment torulene which gives the red colour pigment only found in the red aphid which suggests that this gene is responsible for the body colour differences observed between the two aphid genotypes. On the other hand, the absence of protoheme IX

farnesyltransferase, mitochondrial-like has been reported to decrease virus attenuation and enhance pigmentation in *S.auerus* (Xu et al. 2016). However, their role in aphids remains unknown. Next, many genes involved in histone modification such as histone deacetylase 8 (HDAC 8), histone deacetylase complex subunit SAP18 and histone acetyltransferase KAT7-like (KAT-7) were upregulated in the red morphs. A study by Kirfel et al. (2020) showed that silencing the HDAC 8 and SAP18 has no effect on aphid life-history traits (survival, total number of offspring and total number of premature offspring) but silencing Kat-7 results in significantly more offspring. Our results also showed that Kat-7 is more highly expressed in the N127 red morphs. Since the N127 red morphs usually have lower reproduction than the N116 green morph, it is possible this gene might be responsible for the reproduction differences observed in these two genotypes. Apart from that, some genes involved in insecticide resistance such as esterase E4 (E4) and cytochrome P450 4C1-like (p450) were upregulated in red morphs. Previous studies have shown that aphids with increased expression in this gene have higher resistance to insecticide (Silva et al. 2012; Hirata et al. 2017) However, differences in insecticide resistance between these two genotypes remain unknown and further study is needed to elucidate the role of these genes in pea aphid insecticide resistance.

Apart from that, 28 of the 50 genes involved in carotene production, wing morphogenesis, ATP synthesis, longevity, cuticle synthesis, survival, reproduction and insect hormone and wing development were downregulated in the N127 red morphs. For example, a lot of the genes involved in ecdysone signalling such as ecdysone-induced protein 78C (Eip78C), ecdysone receptor (EcR), Krueppel homolog 1 (Kr-h1) were downregulated in red morphs. Ecdysone plays an essential role in insect development (Cruz et al. 2006). In *Drosophila*, knocking down Ecdysone related genes has been shown to result in a decrease in oogenesis (Ables et al. 2015). Since red morphs have lower reproduction in comparison to green morphs it is possible that ecdysone might play a role in regulating the reproduction between these two genotypes. Apart from that, ecdysone also plays an important role in

regulating wing polyphenism in aphids. Aphids with higher ecdysone levels usually produce more wingless offspring (Vellichirammal et al. 2017). Even though the red morphs have a lower expression for EcR and Eip78c they also have a lower level of expression of Kr-h1 which is usually involved in ecdysone inhibition (Zhang et al. 2018). Since the red morph has lower expression Kr-h1 it means that there is less inhibition of ecdysone which could result in higher ecdysone levels despite lower expression of ecdysone related genes. Therefore, it is possible that the inability of red morphs to produce winged offspring when exposed to environmental stress is related to the ecdysone pathway and further study is needed to confirm this. Apart from that, another gene Gustavus (Gus) was also downregulated in the red morph. Gao et al. (2021) showed that silencing this gene in *Drosophila* leads to a reduction in reproduction in the green peach aphid. Therefore, it is possible that the reduction of expression of this gene could be responsible for the lower reproduction observed in N127 red morphs.

Next, most DEGs between N127 red wingless morph and N116 green wingless aphid were distributed to GO terms such as transmembrane transport, proteolysis, DNA integration and transmembrane transport activity and similar units were observed in KEGG such as transport and catabolism, nucleotide metabolism, metabolism of cofactors and vitamins and xenobiotics biodegradation and metabolism. The results suggest that these functional terms and units might be important in regulating differences in life history traits between the two genotypes.

N127 pale vs N116 winged morphs

Two alternative morphs were produced from each genotype (pale morphs in N127 and winged morphs in N116) respectively when exposed to environmental stress. The transcriptomic difference between the N127 pale and N116 winged morph was examined using RNA-seq to explore the underlying differences between the alternative morphs in two genotypes. We found 3153 genes that were significantly differentially expressed between

N127 pale aphid and N116 winged aphid. 33 differentially expressed genes were filtered out from the list of 3153 genes. These genes were involved in carotene production, reproduction, metabolism, longevity and cuticle formation. 16 of the 33 genes were upregulated in the N127 pale morphs compared to N116 winged morphs. For example, the gene G-protein coupled receptor Mth2 (Mth2) was upregulated in the pale morphs. This gene has been reported to increase lifespan and survival during starvation (Li et al. 2014). Both pale and winged morphs have increased lifespan compared to their counterpart (N127 red and N116 wingless). However, the lifespan and survival differences between these two alternative morphs are unknown and further research is needed to elucidate the role of this gene in the lifespan and survival of the alternative morphs. Interestingly, the gene responsible for the red pigmentation in aphid tor was highly upregulated in the pale morph. The pale morph is usually formed by breaking down the red pigmentation for energy when exposed to starvation stress (Tadbakani et al. 2013). Therefore, it is possible that the red pigmentation is still being produced in aphids but subsequently broken down to produce more energy in the pale morph. Interestingly, the gene Wnt-1 was also highly upregulated in the pale morphs. Wnt-1 plays an essential role in insect development, especially wing development (Murat et al. 2010; Ding et al. 2019). This is surprising such that the pale morph does not possess any wing appendages. However, Wnt-1 also has other roles such as the development of central nervous system, trachea and eye, therefore the higher expression in the pale morph might be related to this development rather than wing development.

In contrast, 17 of the 33 genes were downregulated in the N127 pale morphs. For example, fligthin (fl), partner of bursicon (Pburs), and mitogen-activated protein kinase p38b (p38 MAPK) were downregulated in N127 pale morph. These genes are involved in wing development in insects (Adachi-Yamada et al., 1999; Peabody et al. 2008). This is not surprising as the N127 pale morph does not possess any wing appendages in comparison to the N116 winged morphs. Apart from that, the gene gustatory receptor for sugar taste

64-like (Gr64f) was also downregulated in the pale morph. Gr64f plays an important role in feeding motivation during starvation and for detection for detecting sugars such as fructose, sucrose and others (Mishra et al. 2013). The N127 pale aphid only has increased motion and cannot fly in comparison to the N116 winged morph. Therefore, the lower expression might be due to the fact they require a less heightened sensitivity in detecting sugar as they cannot disperse as far compared to the winged morph. Lastly, the gene histone deacetylase Rpd3-like was also downregulated in the N127 pale morphs. Kirfel et al. (2020) showed that knocking out the Rpd3 results in a reduction in reproduction in pea aphids. In our study, we found that the N127 pale morphs have lower reproduction in comparison to the N116 winged morphs (Chapter 3) and the lower expression of this gene might be responsible for the difference in the reproduction observed between these two morphs. Lastly, genes such as trehalase-like (Treh) and facilitated trehalose transporter Tret1 (Tret1-1) are downregulated in N127 pale morph. Both genes are involved in glucose production and transport (Shi et al. 2017). Since N127 pale morph only increases its locomotion in comparison to the N116 winged morphs which usually fly away, it is possible that flying requires more energy consumption compared to increase locomotion.

Next, most DEGs between the N127 pale morph and the N116 winged morphs were distributed to GO terms such as transmembrane transport, proteolysis, DNA integration, heme binding and carbohydrate metabolic process and similar units were observed in KEGG such as amino acid metabolism, carbohydrate metabolism, metabolism of cofactors and vitamins and xenobiotics biodegradation and metabolism. The results suggest that these functional terms and units might be important in regulating differences in life history traits and development between the two alternative morphs across two genotypes.

N127 (N127 red+N127 pale) genotype vs N116 genotype (N116 winged+N116 wingless)

The transcriptomic difference between the N127 genotype (both red and pale morphs) and N116 (both winged and wingless morph) were examined using RNA-seq to explore the underlying difference in transcriptome profile for the genotype as a whole. We found 5262 genes that were significantly differentially expressed between N127 and N116 aphids. 39 differentially expressed genes were filtered out from the list of 5262 genes based on their potential role in aphid development. Some of the genes that were highly upregulated in the N127 morphs are involved in carotenoid pigmentation (tor). As discussed above, this gene is mainly involved in producing the torulene which is responsible for the red pigmentation observed in the red morphs therefore is not surprising that the N127 genotype has a higher expression for this gene. Next, the N127 genotype also showed higher expression for the gene histone-lysine N-methyltransferase, H3 lysine-79 specific (H3K79). This gene is involved in H3 lysine 79 methylation and is usually associated with gene activation (Dottermusch-Heidel et al. 2014). However, the role of H3K79 methylation function in aphids remains unclear and requires further study. Apart from that, N127 also showed upregulation for the heat shock protein 70 (Hsp70). Interestingly, N127 always showed higher expression even in individual group comparison (N127 pale vs N116 winged) and (N127 red vs N116 wingless). Heat shock protein plays important role in reaction to environmental stress such as heat stress, starvation and others (Farahani et al. 2020). Previous studies on *Rhodnius prolixus* and *Harmonia axyridis* have shown that Hsp70 plays an important role in insect development and resistance to starvation (Shen et al. 2015, Paim et al. 2016). Next, Freitak et al. (2012) reported a difference in Hsp expression in *Tribolium castaneum* when exposed to starvation and heat stress. The role of Hsp70 in aphids remains widely unclear and further research is needed to elucidate the role of Hsp70 in regulating aphid starvation resistance.

In contrast, 20 of the 38 genes were downregulated in the N127 genotype in comparison to the N116 genotype. Some of the downregulated genes in N127 were involved in wing development such as imaginal disk growth factor precursor (Idgf), fligthin (fl) and the

function of these genes has been discussed above. Since only the N116 genotype can produce winged offspring, it is not surprising that these genes were downregulated in the N127 genotype. Next, we found that apart from heat shock protein 70, other heat shock proteins such as heat shock protein 68 (Hsp68), and heat shock protein 75 (Hsp75) were downregulated in the N127 morphs. Most studies on Hsp75 and Hsp68 have been reported in the context of thermal stress. For example, Chen et al. (2015) reported an increased expression of Hsp75 in *Neoseiulus cucumeris* when exposed to heat stress. Chen et al. 2019 also reported a similar trend in Hsp68 expression when exposed to heat stress. Therefore, Hsp75 and Hsp68 may play a major role in thermal stress reaction rather than starvation stress. Apart from that, Rank et al. (2007) reported differences in the expression of heat shock protein expression across the willow beetle *Chrysomela aeneicollis* genotypes. Therefore, different Hsp may be expressed across aphid genotypes when exposed to starvation stress.

Next, most DEGs between N127 and N116 were distributed to GO terms such as transmembrane transport, proteolysis, DNA integration, heme binding and zinc ion binding and similar units were observed in KEGG such as transport and catabolism, carbohydrate metabolism, metabolism of cofactors and vitamins and xenobiotics biodegradation and metabolism. The results suggest that these functional terms and units might be important in regulating differences in life history traits and development between the two genotypes.

Morphs (N127 pale+N116 winged) vs wildtype (N127 red+N116 wingless).

The transcriptomic difference between the alternative morphs (both N116 winged and N127 pale morphs) and wild-type (both N116 wingless and N127 red morph) was examined using RNA-seq to explore the underlying difference in transcriptome profile between dispersal vs non-dispersal morphs. We found 1467 genes that were significantly differentially expressed between alternative morphs (N127 pale aphid + N116 winged aphid) vs wild types (N127 red + N116 wingless). 23 differentially expressed genes were filtered out from the list of

5287 genes. Some of the genes that were upregulated in the alternative morphs were involved in resistance to starvation. For example, the gene phosphoenolpyruvate carboxykinase (Pepck). Pepck is an enzyme that is usually involved in the gluconeogenesis process resulting in phosphoenol pyruvate (PEP) (Yang et al. 2009). PEP can be used to generate pyruvate that can be used for cellular respiration and therefore could play a potential role in response to stress such as starvation. A recent study by Spacht et al. (2018) showed that Pepck was upregulated in the flesh fly *Sarcophaga bullata* when exposed to starvation conditions. Our results agree with this study as the alternative morphs both showed higher expression for this gene in comparison to the wild type. Apart from that, other genes such as chemosensory protein-like precursor (Csp) are also upregulated in the alternative morphs. Csp plays an important role in volatile detection and searching for a new habitat (Song et al. 2018). Since dispersal morphs need to disperse to other plants and look for new habitats, therefore it is possible that Csp might play an essential role in aphids in detecting suitable host plants to colonise.

In contrast, genes that were downregulated in the alternative morphs were involved in stress response. A lot of heat shock protein genes such as Hsp60, Hsp70, and Hsp83 were all downregulated in the alternative morphs. Heat shock protein plays an important role in stress response but can also regulate other aphid development such as reproduction, resistance to starvation and other functions (discussed above). Apart from that, cuticle genes such as cuticle protein 7 (Cp7) were downregulated in the alternative morphs. Since the production of cuticles can be energetically costly, it is possible that the alternative morphs reduced the production of energy to invest energy for dispersal (both flight and increase locomotion).

Next, most DEGs between alternative morphs and wild type were distributed to GO terms such as transmembrane transport, proteolysis, structural constituent of cuticle, transmembrane transporter activity and heme binding. Similar units were observed in KEGG such as ABC transporter, biosynthesis of amino acids, autophagy-animal and

lysosome. The results suggest that these functional terms and units might be important in regulating differences in life history traits and development between the alternative morphs and wild type.

4.5 Summary

In this chapter the expression of candidate genes involved in wing development, stress response and metabolite regulation were investigated between different pea aphid morphs. Our results are the first to reveal that the candidate genes involved in wing development for e.g., vg, ap1 were differentially expressed between the pea aphid morphs with the N116 winged aphids having the highest expression in these genes. This suggests that these genes could be important in helping the understanding the underlying mechanism regulating wing development in aphids. From an agroecology perspective, understanding wing development in aphids is important as it could help us develop potential insecticide or pesticide in controlling aphid migration to prevent the huge damage in agriculture.

Further, the transcriptome profiles between the pea aphid morphs were also investigated in this chapter. We found many DEGs between the genotype group comparison which could play an important role in wing polyphenisms and aphid development. In the within-genotype comparison (e.g., N116 winged vs N116 wingless) and (N127 pale vs N127 red), we found that the transcriptome profiles are very different such that N116 shows many genes that are involved in wing development while the N127 group comparison shows mostly genes that are involved in metabolite regulation and carotene productions. This is the first time such a difference in transcriptome profile is reported between aphid genotypes that either produce winged offspring or changes their body colour. Our transcriptome profile not only reveals potential candidate genes in regulating aphid response to environmental stress but also potential underlying mechanisms in regulating their development. Wing development has been contributing to the evolutionary success of aphids. Our results not only help provide insight into genes in regulating wing development in aphids but also alternative

genes that help regulate aphid responses to stress for genotypes that do not produces wing which could contribute to their evolutionary success in the wild.

This chapter also reveals many DEG between the genotype group comparison (e.g., N127 red vs N116 wingless) that are involved in carotene production, metabolite regulation, stress response and reproduction. This is important as it could help us understand the prevalence of both genotypes in the wild. Future research on manipulating these genes (e.g., RNA interference, knockout gene) and the effect on aphid development (e.g., developmental time, reproduction, survival, wing development) should be investigated to provide insight into the role of these genes in regulating pea aphid life-history traits and also polyphenisms.

Chapter 5: DNA methylation of pea aphid developmental genes and methylome profile in different pea aphid morphs

5.1 Introduction

Most organisms will experience changes in the environment during their lifetime. Organisms usually perceive changes in environmental conditions and adjust their phenotype. The phenotypic changes can range from minor metabolite changes to modifying their morphology and physiology, which increase their ability to survive in the environment. This phenomenon is also known as phenotypic plasticity (Fordyce 2006). Phenotypic plasticity is commonly defined as the ability of an organism's genotype to produce multiple discrete phenotypes under different conditions (West-Eberhard 2003). One of the most extreme cases of phenotypic plasticity is known as polyphenism. In polyphenism, the output is usually discrete which as opposed to continuous (usually observed in plasticity) resulting in distinct phenotype (Yang et al. 2019). Polyphenism is widely observed across the insect species, ranging from phase polyphenism in the migratory locust *Locusta migatoria* to wing polyphenism in aphids *Acyrthosiphon pisum* (Ma et al. 2011; Ogawa and Miura 2014). In insects, environmentally induced polyphenism plays an important role in the evolutionary success as it allows the insects to quickly adapt to environmental changes. Polyphenism in insects can be triggered by a wide range of abiotic factors such as temperature and food resources. However, the trigger can also be biotic such as increase population density, the presence of predators, and others (Roessingh et al. 1998). Although there are many different types of environmental factors that can induce polyphenism, they all share a common regulatory scheme. First, the environmental signal is perceived by various groups of cells (e.g. nerve cell, muscle cells) followed by signal transduction through different endocrine systems such as ecdysteroid pathways, insulin pathways, juvenile hormone pathways and others. Next, the target tissues and cells will undergo epigenetic changes such as DNA methylation, histone modification, and non-coding RNAs which in turn regulate the expression profile of genes required for the alternative phenotype (polyphenism) (Figure 5.1)

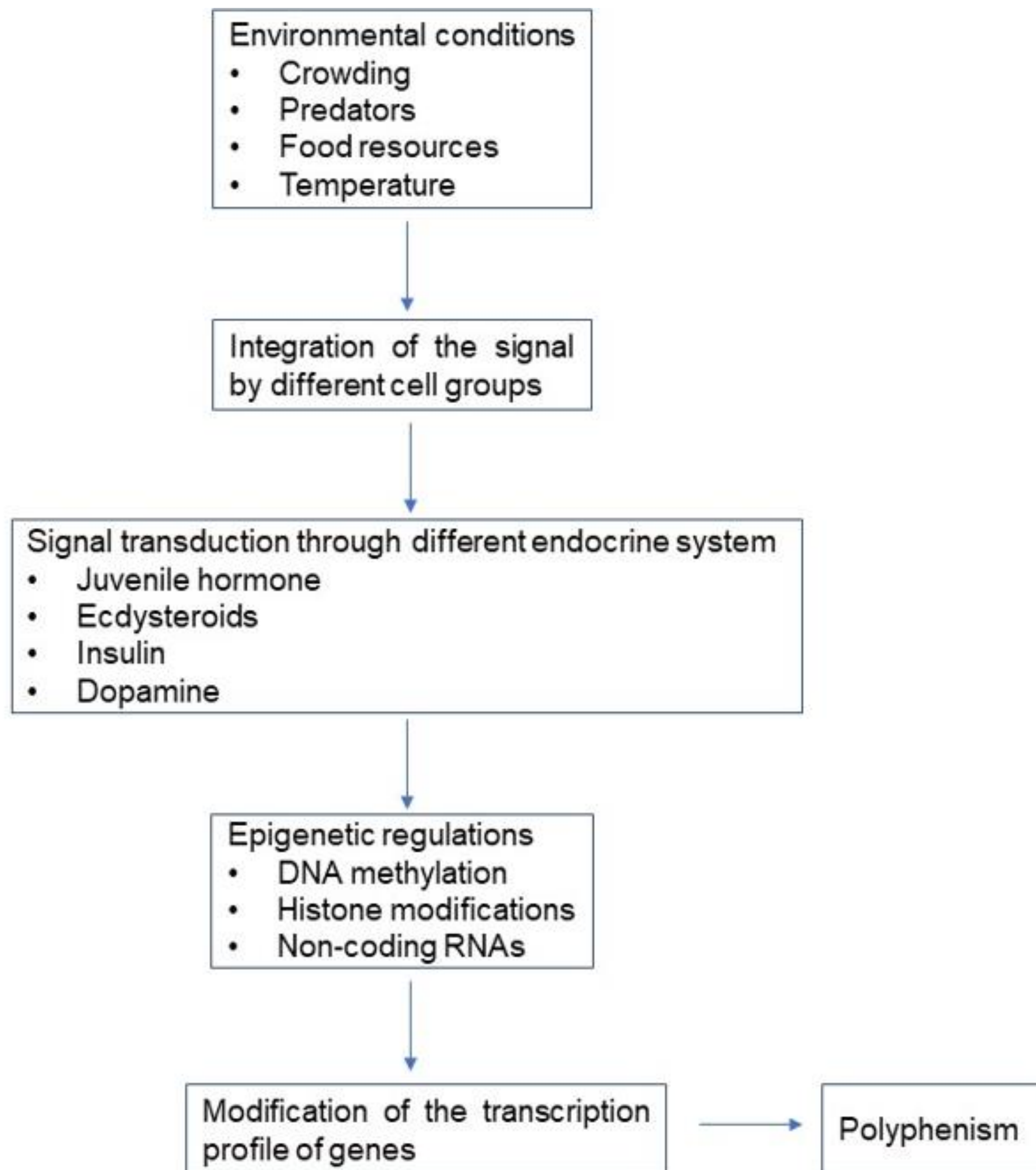


Figure 5.1 The regulation framework for insect polyphenism. Firstly, the environmental signal is detected and integrated by a specific group of cells/tissues in the organisms after exposure to different environmental conditions. Next, the signal will be transduced to target tissues or cell groups through different neuroendocrine systems. After that, different epigenetic mechanisms such as DNA methylation, and histone modifications, are involved in modulating the gene expression which in turn results in the production of alternative phenotypes.

One of the main epigenetic mechanisms is DNA methylation. DNA methylation has been the most studied epigenetic mechanism in the context of polyphenisms and has been shown to regulate the genome of many species (Bewick et al. 2017). DNA methylation usually involves the addition of the methyl group to the C-5 position of the cytosine ring.

The process of DNA methylation is facilitated by a family of enzymes known as DNA methyltransferase (Dnmt) such as Dnmt1, Dnmt 2, and Dnmt 3. In most organisms, DNA methylation can be observed in any of the cytosine throughout the genome but is more commonly found in the CpG dinucleotide area and is established through the help of Dnmt 3. Furthermore, DNA methylation status is usually maintained by Dnmt 1 (Goll and Bestor 2005; Cheng and Blumenthal 2008; Kim et al. 2009; Morandin et al. 2021). The Dnmt family also varies across insect species. For example, the honeybee *Apis mellifera* possesses two copies of Dnmt1, the parasitoid wasp Nasonia vitripennis has three copies of Dnmt1 and *Drosophila* has lost both Dnmt1 and Dnmt3 (Bird 2002; Wang et al. 2006). In insects, DNA methylation is often found in the gene body (exon + intron) compared to mammals where most DNA methylation occurs throughout the genome (Suzuki and Bird 2008, Lyko et al. 2010; Bonasio et al. 2012). DNA methylation has also been reported to regulate insect development, ranging from embryogenesis, memory processing to age-related changes in worker behavior (Kay et al. 2018; Herb et al. 2012). The first reported case of DNA methylation in aphids was found in the E4 esterase gene in the insecticide-resistant green peach aphid, *Myzus persicae* (Field 2000).

Aphids are an excellent model for studying epigenetic mechanisms underlying polyphenisms for a number of reasons. First, aphids show multiple polyphenism such as sexual polyphenism and wing polyphenism. Second, aphids usually reproduce parthenogenetically, therefore all offspring are essentially clones of the mothers, which makes them an excellent model to study the role of epigenetic changes in polyphenism. Third, the aphid genome has also been fully sequenced, which is beneficial when it comes to next-generation sequencing. Apart from that, aphids also possess a large number of genes (~35000 genes) with many genes having undergone duplication events (Srinivasan and Brisson 2012). Some of the genes that were duplicated were involved in multiple epigenetic mechanisms such as DNA methylation, histone modification, chromatin

remodelling and others (Srinivasan and Brisson 2012). Interestingly, the aphid genome also possesses one of the lowest G/C contents (29.6%) in comparison to other fully sequenced insect genomes. Finally, the pea aphid also possesses all the necessary Dnmt enzymes required for the DNA methylation.

This chapter investigates methylation patterns of candidate genes involved in aphid development including wing development, stress response, metabolite regulation and hormone regulation building upon the data from previous Chapters and previous studies (Brisson et al., 2010).

The overall aim of this chapter was to investigate whether the candidate genes involved in aphid development are differentially methylated in the dispersal and non-dispersal morphs, and between genotypes. Recent work has also focused on understanding the difference in aphid genotype and the differences in the degree of phenotypic plasticity between aphid genotypes (Kanvil et al. 2014, Parker et al. 2019, Sentis et al., 2019). Because of the critical role of the ability to disperse in aphid evolutionary success, this chapter also aimed to assess the methylation status between the aphid morphs and genotype to provide insight into the role of epigenetics in regulating aphid polyphenism. Furthermore, we integrated the transcriptome results from (Chapter 4) to provide an insight into the relationship between DNA methylation and gene expression. Specifically, the aims of this Chapter were to:

- Measure the DNA methylation level of genes involved in different pea aphid developmental processes such as reproduction, wing formation, carotene production and stress response by pyrosequencing
- Investigate the methylome profile of different pea aphid morphs in two genotypes (N116 winged vs N116 wingless and N127 pale vs N127 red) by MBD-seq
- Integrate the results of transcriptome profile (chapter 4) and methylome profile to provide an insight into the correlation between gene expression and DNA methylation

5.2 Methods

5.2.1 Materials

All samples used in the experiment were obtained from the mesocosm experiment described in section 2.3.3. The number of samples used in the pyrosequencing analysis is outlined in the (Table 5.1), stratified by genotype and morphs.

Table 5.1. Summary of aphid numbers used for molecular analysis in Chapter 5.

Assay	Genotype	Morphs	Replicates
Pyrosequencing analysis	N116	wingless	x5 (pooled samples)
	N116	winged	x5 (pooled samples)
	N127	red	x5 (pooled samples)
	N127	pale	x5 (pooled samples)

5.2.2. DNA methylation level by pyrosequencing

Total DNA was extracted from whole adult aphids obtained as described in Chapter 2 (section 2.4.2), quantified (section 2.4.3b), quality checked for integrity (section 2.4.3c) and undergone bisulfite conversion (section 2.4.3d). DNA methylation of candidate genes is quantified using pyrosequencing (section 2.4.3e).

5.2.3 MBD-seq library preparation

DNA was extracted from seven adult aphids using the Qiagen DNeasy Blood & Tissue Kit (Qiagen, UK) according to the manufacturer's instructions. DNA concentration and quality checks were done using nanodrop, gel electrophoresis and Qubit (Thermo Fisher Scientific, USA). Genomic DNA was sheared using Next Gen Bioruptor to a size range of 400bp and was checked using TapeStation System 4200 (Agilent Technologies, USA). DNA enrichment was performed using the MethylMiner™ Methylated DNA Enrichment Kit (Thermo Fisher Scientific, USA) with an initial input of 1µg of sheared DNA per reaction. Briefly, DNA was bound to the MBD-biotin protein containing Dynabeads M-280

Streptavidin followed by a series of washing. Lastly, the methylated beads were eluted from the capture beads using a High Elution Buffer. Next, the library was prepared using the Next Gen DNA library kit (Active Motif, USA). Briefly, the eluted DNA undergoes two rounds of repair steps and washing, followed by a ligation step with the addition of a specific index for each sample using the Next Gen Indexing Kit (Active Motif, USA), followed by another round of washing. Lastly, the sample containing a specific index undergoes a last round of ligation with washing and amplified using PCR at 98°C for 30s, followed by six cycles of DNA denaturation at 98°C for 10s, followed by annealing at 60°C for 30s and lastly extension at 68°C for 60s. The amplified library is then washed, and the library size is analysed using TapeStation System 4200 (Agilent Technologies, USA). The library is sequenced on the Illumina Next Seq 550 Series (Illumina, USA) using the NextSeq 500/550 v2.5 Kits (75 cycles) (Illumina, USA). Finally, the output data was demultiplexed (allowing one mismatch) and BCL-to-Fastq conversion was performed using Illumina's bcl2fastq software, version 2.20.0.422.

5.2.4 Sequence assembly and annotation

Unmapped single ended reads of 76bp from an Illumina NextSeq550 sequencer were checked using a quality control pipeline consisting of FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and FastQ Screen (https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/). The reads were trimmed to remove any adapter or poor-quality sequence using Trimmomatic v0.39. The reads were mapped against the reference *Acyrthosiphon pisum* genome v3.0. Next, MACS2 v2.2.5 (Zhang et al. 2008) was used to identify MBD regions and programs from the UCSC repository (http://hgdownload.soe.ucsc.edu/admin/exe/) were used to convert bedGraph files from MACS2 to bigWig, including bedClip and bedGraphToBigWig. Differential binding analysis was performed using DiffBind v3.4.11 (Stark and Brown 2011) in R v4.1.2 (R Core Team 2021). The 'read' input were the final filtered BAM files used in

the MACS2 peak calling. The annotation of differentially bound regions was performed using ChIPseeker v1.30.3 (Yu et al. 2015). Differentially bound MBD regions from were filtered by log2 fold change (positive >0.3 and negative <-0.3, or either separately) and the 'conc' value (mean normalised reads) of 6.64 (~100 reads). The relative location of each differentially bound region to associated genes were determined using the annotatePeak function. Gene Ontology enrichment analysis was performed using the enricher function of clusterProfiler v4.2.2 (Wu et al. 2021) in R v 4.1.2 and GO annotation for the aphid genes was obtained from the Bioinformatics Platform for Agroecosystem Arthropods (BIPAA).

5.2.5 Integration of RNA-seq and MBD-seq data

The data from RNA-seq and MBD-seq were integrated to provide insight into the relationship between DNA methylation and gene expression. First, for differentially methylated genes (DMGs) from MBD-seq, I searched for the corresponding gene expression level in RNA-seq with the criteria of (P<0.05 and <-0.3<fold change>0.3). Next, the relationship between DNA methylation and gene expression was listed as the same trend (increased methylation and increased expression in gene body, increased methylation and decreased gene expression in promoter region) or the opposite trend (increased methylation and decreased expression in gene body, increased methylation and increased gene expression in promoter region) for both hypermethylated and hypomethylated genes. Genes that have multiple regions methylated (promoter, 3’UTR and gene body) or only (3’UTR and 5’UTR) methylated were listed as overlapped and not available (NA), respectively, for the analysis.

5.3 Results

5.3.1 DNA methylation levels

We measured DNA methylation of all the genes discussed in (Chapter 4) to provide insight into the role of DNA methylation in gene expression. Our results reveals that most of the genes have low levels of methylation (1-3%) with some intermediate level (30-40%) and only a few highly methylated genes (70-100%). Among the lowly methylated genes, the DNA methylation levels were consistently low across the CpG sites examined. In contrast, there is a bigger variation level in individual CpG methylation level for genes that were intermediate and highly expressed. Further a few genes (APNS1, Hsp70 and INR) were not tested due to failure to design primers. The DNA methylation level of 8 different genes selected based on the (N116 winged vs N116 wingless expression) RNA-seq experiment were analysed (Figure 5.2). Next, the DNA methylation of another 8 different is also analysed and were selected based on their function in pea aphids reported in previous studies (such as stress response, wing formation, carotenoid production) (Figure 5.3). Lastly, the significant predictors for each genes were reported in (Table 5.2).

Table 5.2. Summary of statistical model used for gene expression analysis and the significant predictors for each genes.

Gene	Model	Significant predictors	F-value	P-value
Troponin C	GLM	None of the predictors were significant **reported value here is from the interaction of morphs and genotype	1.8	0.198
Trehalase	GLM	None of the predictors were significant **reported value here is from the interaction of morphs and genotype	2.148	0.162
Krueppel-homolog 1	GLM	None of the predictors were significant	0.000	1.000

		**reported value here is from the interaction of morphs and genotype		
Mother against decapentaplegic homolog 4	GLM	Main effect (morphs)	11.545	<0.001
Partner of Bursicon	GLM	Interaction of morphs and genotype	5.031	0.039
DNA methyltransferase 1 associated protein 1	GLM	Interaction of morphs and genotype	8.793	0.009
Histone deacetylase Rpd3	GLM	Main effect (morph)	8.593	0.010
Ecdysone 20-monooxygenase	GLM	None of the predictors were significant **reported value here is from the interaction of morphs and genotype	0.000	1.000
Flightin	GLM	Main effect (genotype)	7.443	0.015
Apterous 1	GLM	None of the predictors were significant **reported value here is from the interaction of morphs and genotype	0.400	0.536
Engrailed	GLM	None of the predictors were significant **reported value here is from the interaction of morphs and genotype	1.361	0.260
Vestigial	GLM	None of the predictors were significant **reported value here is from the interaction of morphs and genotype	1.670	0.215
Short-chain specific acyl-coA dehydrogenase	GLM	None of the predictors were significant **reported value here is from the interaction of morphs and genotype	0.000	1.000

Heat Shock Protein 83	GLM	Interaction of morphs and genotype	5.458	0.033
Carotene dehydrogenase	GLM	Main effect (morphs) ** morphs here represent only pale vs red aphid in N127 genotype as no expression of these genes for N116 genotype	0.007	0.934

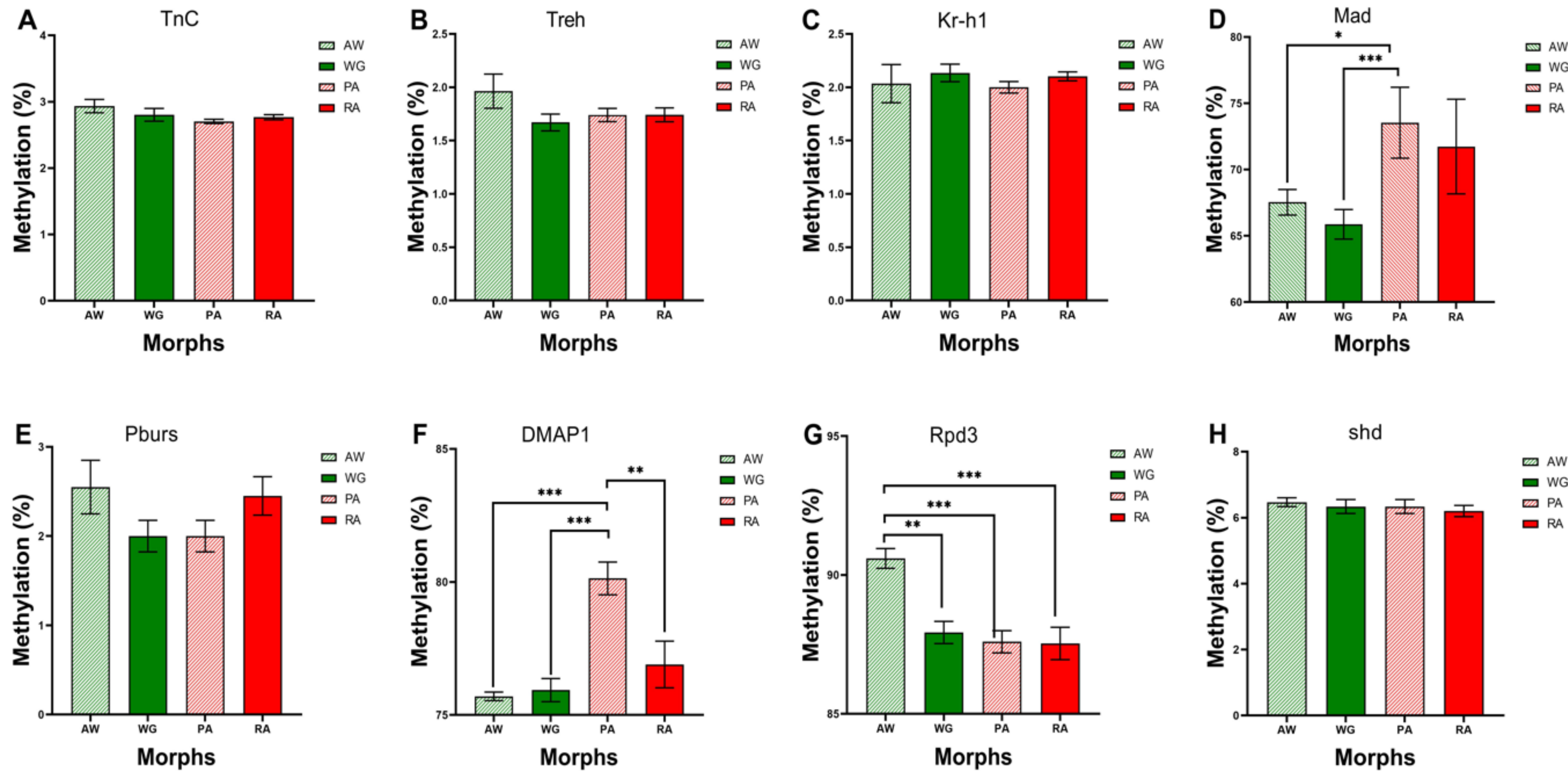


Figure 5.2. DNA methylation level of candidate genes between aphid morphs (from RNA-seq). AW=adult wing, WG=adult wingless, PA=pale adult, RA=red adult. Each of the bars represents the average measurement of expression from five biological replicate with error bars indicating the SEM. The Y-axis scale is specific to each gene. *P-values <0.05, **P-values <0.01, ***P-values<0.001 obtained from Tukey's post-hoc test. From left to right: TnC (Troponin C), Treh (Trehalase), Kr-h1 (Krueppel-homolog 1), Mad (mothers against decapentaplegic homolog 4), Pburs (partner of bursicon), DMAP1 (DNA methyltransferase 1-associated protein 1), Rpd3 (Histone deacetylase Rpd3), shd (Ecdysone 20-monooxygenase).

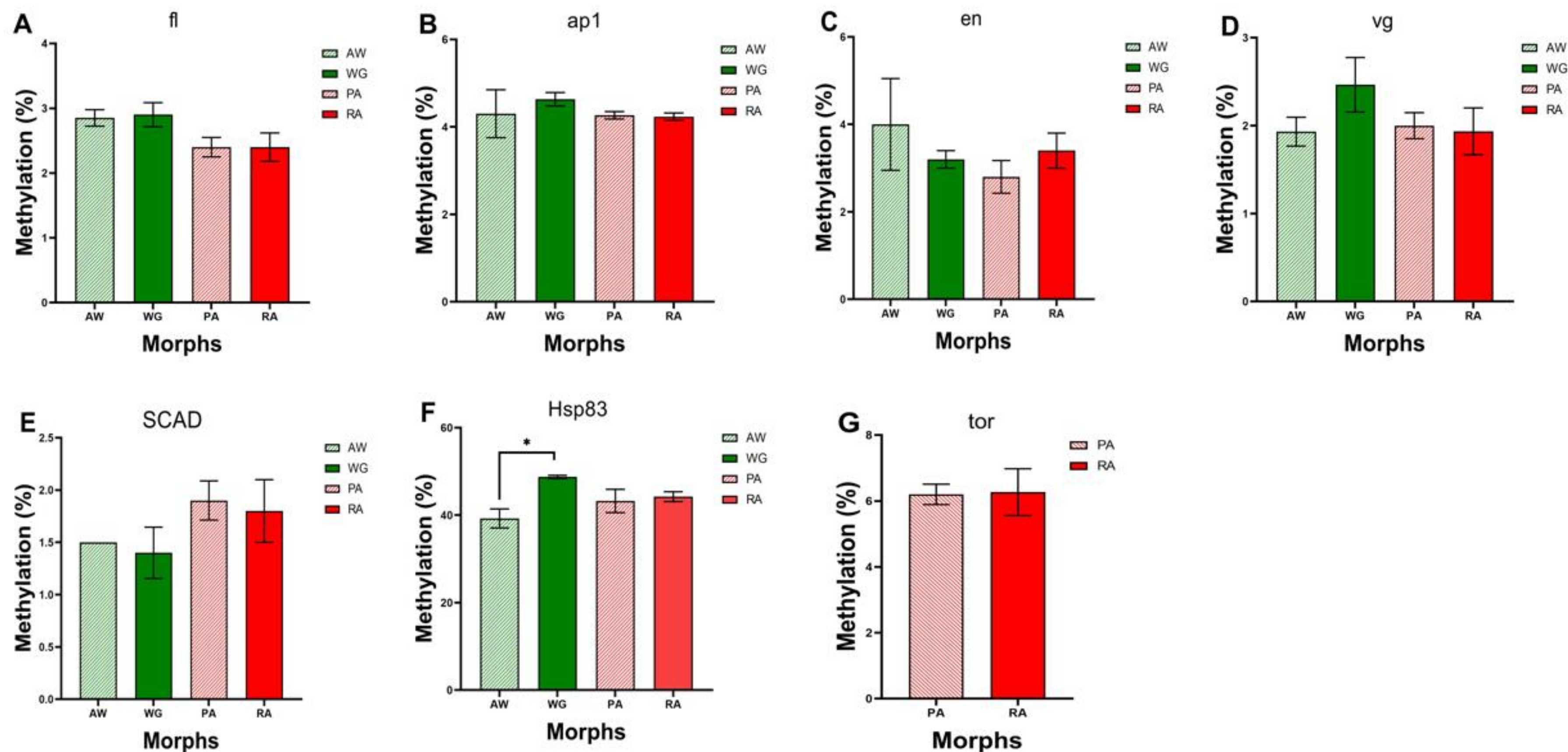


Figure 5.3. DNA methylation levels of candidate genes between aphid morphs. AW=adult wing, WG=adult wingless, PA=pale adult, RA=red adult. Each of the bars represents the average measurement of expression from five biological replicates with error bars indicating the SEM. The Y-axis scale is specific to each gene. *P-values <0.05, **P-values <0.01, ***P-values<0.001 obtained from Tukey's post-hoc test. From left to right: fl (Flightin), ap1 (apterous1), en (engrailed), vg (vestigial), SCAD (Short-chain acyl-CoA dehydrogenase), Hsp83 (Heat shock protein 83), tor (carotene dehydrogenase).

5.3.2 MBD-seq analysis

Four different morphs from the two genotypes were selected for whole methylome profiling: genotype N116 (wingless and winged) and genotype N127 (red and pale). The two genotypes have very different body morph colours: N116 is usually green, while N127 is usually red. Apart from that, they respond differently to crowded conditions, with N116 producing winged offspring and N127 changing their body colour from red to pale.

5.3.3 Methylome profiles of pea aphid genotypes and their morphs

The methylation profile of four different pea aphid morphs was compared. 196 million reads were obtained for N116 (winged), 195 million reads for N116 (wingless), 394 million reads for N127 (pale), and 365 million reads for N127 (red) (Table 5.3). After filtering (includes trimming and mapping to reference genome), 124 million, 127 million, 248 million, and 226 million high-quality reads were obtained for each morph, respectively. Apart from that, principal analysis plot was carried out to visualise the overall patterns of DNA methylation across morphs (Figure 5.4).

Table 5.3. Number of raw reads and unique mapped reads obtained from MBD- seq.

Strain	Number of inputreads	Uniquely mapped read
N116 winged	196,085,634	123,873,182
N116 wingless	194,873,237	126,961,032
N127 pale	394,240,334	247,478,037
N127 red	364,744,675	226,039,832

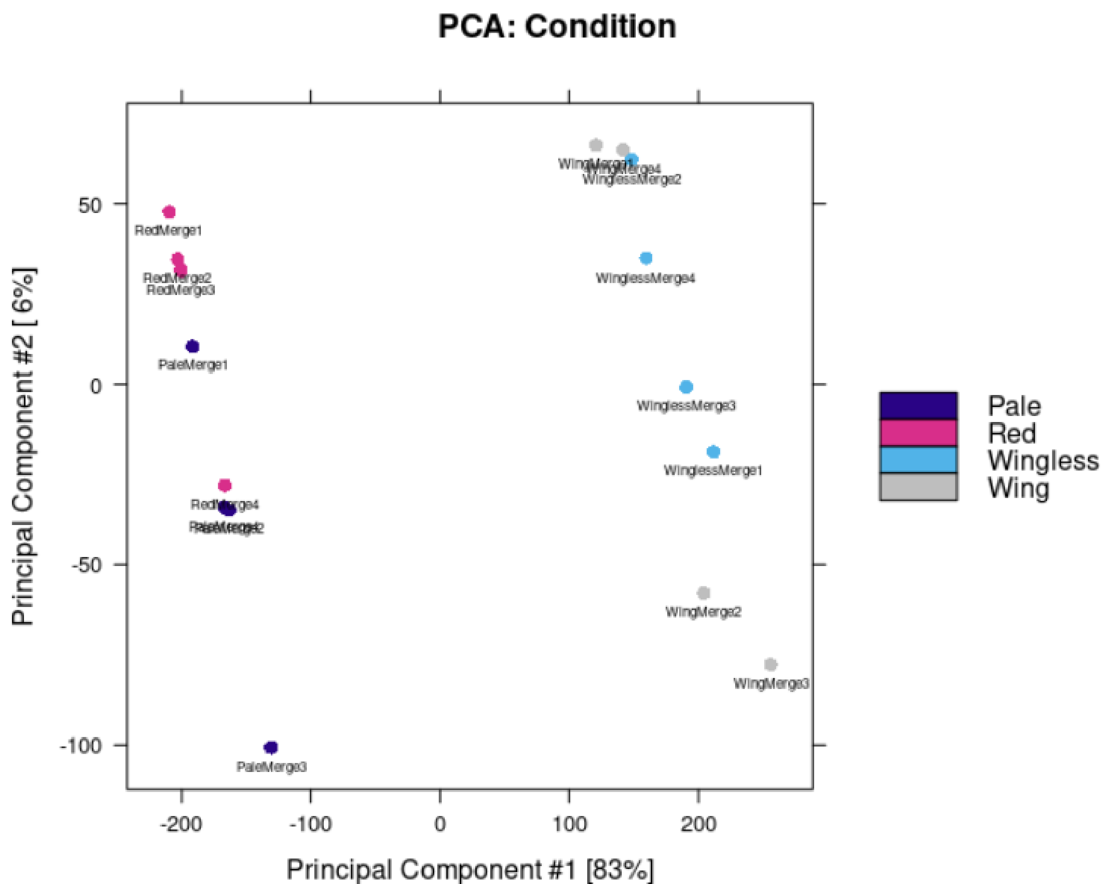


Figure 5.4. Principal component analysis of gene expression of four pea aphid morphs. Each dot represents one biological replicate from the MBD-seq experiment.

For DNA methylation level measurements, a range of 62%–65% of the total reads were uniquely mapped to the genes in the reference database. The significant DMGs for all groups were then further annotated for GO by using the ClusterProfiler (p-adjust <0.1). For further functional categorisation, KEGG pathway analysis was performed using the aphid KEGG database. DEGs were then classified into different KO terms.

a) Differentially methylated genes between different all group comparison

Based on the results of MBD-seq we found 77 genes that were differentially expressed between the N116 winged aphid vs N116wingless aphid, 410 genes between N127 pale aphid with N127 red aphid, 1967 genes between N127 red aphid vs N116 wingless aphid

and 1675 genes between N127 pale vs N116 winged aphid (Figure 5.5). Then, we only found 1 overlapping genes between the N116 winged aphid vs N116wingless aphid group with N127 pale vs N127red aphid group. Further, we found 197 genes overlap between N127pale vs N127red aphid group vs N127red vs N116wingless aphid group. Most overlapping genes (1390) were observed between the N127red vs N116wingless aphid group vs N127pale vs N116 winged aphid group. Then, 41 overlapping genes were found between N116winged vs N116 wingless group vs N127pale vs N116winged aphid group. No overlapping genes were found across all the other comparisons. All differentially methylated genes were obtained based on the parameter of (P-value <0.1 and -0.3<log2fold>0.3).

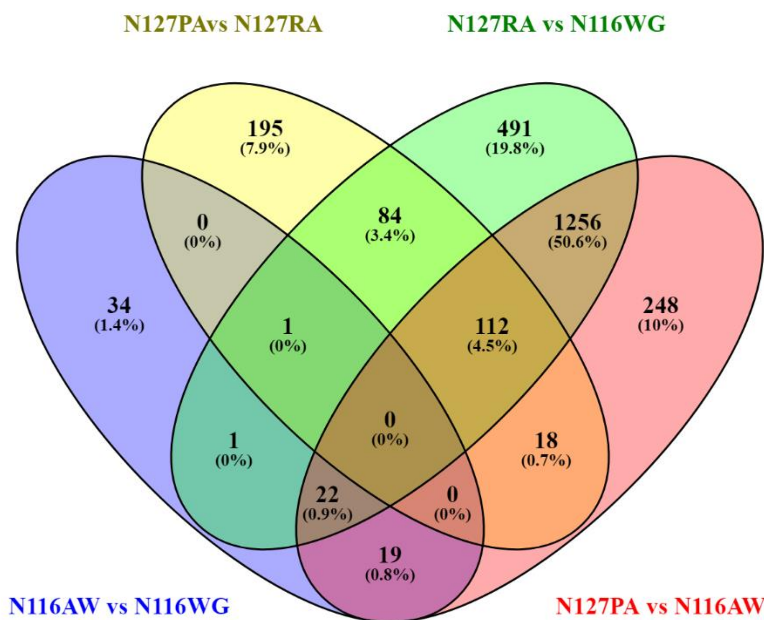


Figure 5.5. Venn diagram comparing DMG identified by MBD-seq between different aphid morphs. Blue=N116 winged +N116wingless morphs, Green=N127red +N116winglesss, Yellow= N127 pale + N127 red morphs, Red=N127 pale+ N116 winged morphs**.** Each sub-category represents the overlapping genes that were found between the two or more different group comparisons.

b) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N116 winged and N116 wingless aphids

Our DMR analysis showed that gene body (exon + intron) (39.5%) and distal intergenic region (39.5%) is the most methylated region followed by UTR (13.5%) and promoter (7%) (Figure 5.6A, B). Most of the methylated regions are located on the X chromosome (79%), followed by the A2 chromosome (9%), A1 chromosome (4%), A3 chromosome (4%) and unplaced scaffold regions (4%) (Figure 5.6C, D). We found four genes to be hypermethylated and 147 hypomethylated in the N116 winged morphs in comparison to the N116 wingless morphs (Figure 5.6E). The top 10 most hypermethylated genes in N116 winged were involved in cellular development. In contrast, the top 10 most hypomethylated genes were involved in ubiquitin processes (Table 5.4).

Next, the genes were subjected to GO analysis. None of the genes showed significant enrichment. In addition to the GO analysis, the genes were also subjected to KEGG analysis, and no significant enrichment was obtained for KEGG analysis either.

Table 5.4. Top 10 most hypermethylated and hypomethylated in N116 winged and N116 wingless aphids

Gene	Chromosome	Fold-change
Uncharacterized	X	0.48
disks large homolog 1	X	0.62
myocardin-related transcription factor A	X	0.55
nucleotide exchange factor SIL1	X	0.38
Uncharacterized	X	-0.47
sentrin-specific protease 1-like	X	-0.47
Uncharacterized	X	-0.45
kelch-like protein 2	X	-0.45
E3-independent) E2 ubiquitin-conjugating enzyme UBE2O-like	X	-0.45
peroxisomal membrane protein PEX13-like	X	-0.44
Uncharacterized	X	-0.44
eukaryotic translation initiation factor 4 gamma 3-like	X	-0.43
Uncharacterized	X	-0.43
26S proteasome regulatory subunit 4	X	-0.42

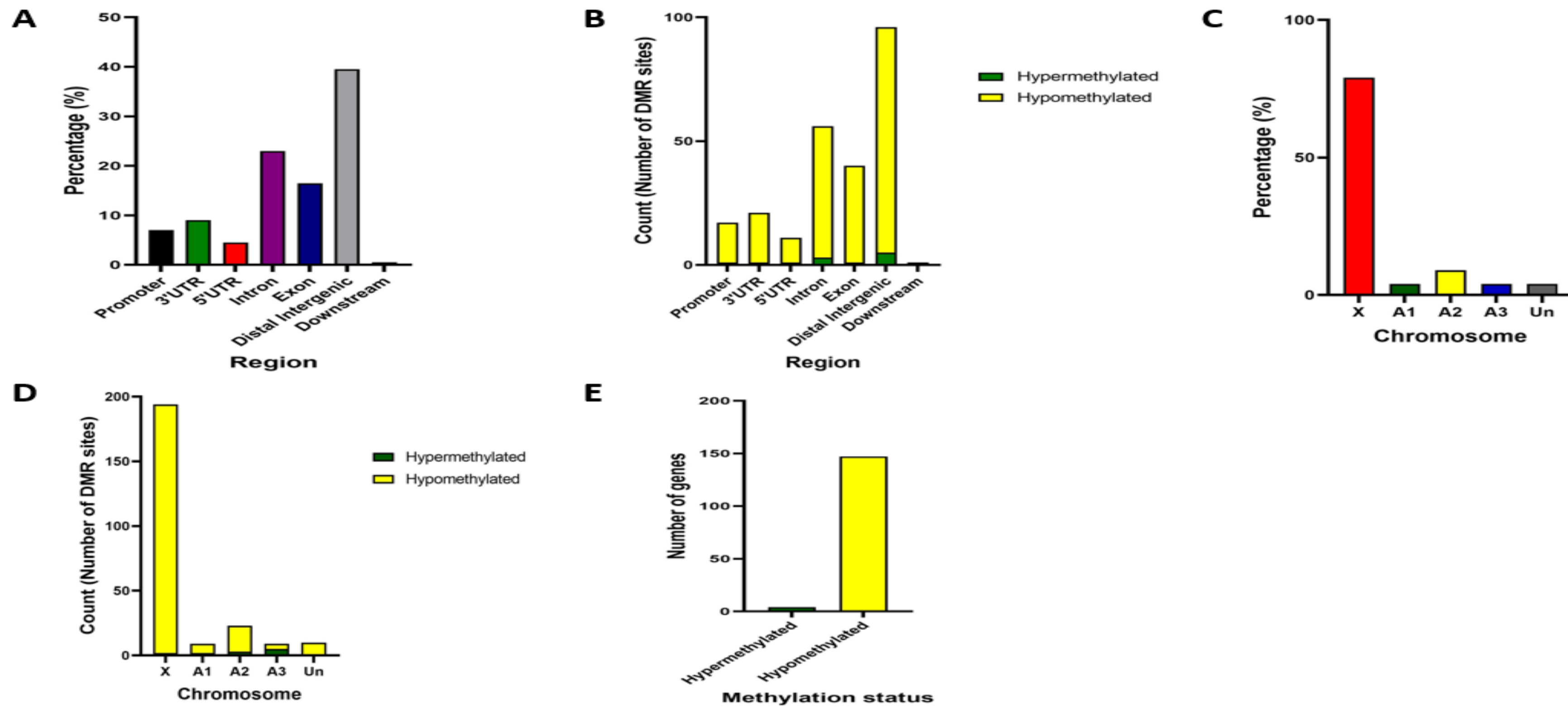


Figure 5.6. MBD-seq results for N116 winged and N116 wingless aphids. A) The DMR region between N116 winged and wingless. Black=promoter, green=3'UTR, red=5'UTR, purple=intron, blue=exon, grey=distal intergenic (DNA region between the coding genes), yellow=downstream Y-axis (percentage)=The number of site for each region for both hypomethylation and hyper methylation together (Percentage calculated as the number of sites for region divided by the total number of sites). **B)** The methylation status of the DMR. Green=hypermethylated, yellow=hypomethylated. **C)** The distribution of the methylation sites by chromosome. Red=X chromosome, green=A1 chromosome, yellow=A2 chromosome, blue=A3 chromosome, grey=unplaced scaffold. **D)** The methylation status of DMR by chromosome. Green=hypermethylated, yellow=hypomethylated**. E)** The total number of genes that are differentially methylated. Green=hypermethylated, yellow=hypomethylated

c) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N127 pale and N127 red aphids

Our DMR analysis showed that the gene body (exon + intron) (40%) is the most methylated region followed by distal intergenic (31%), UTR (19%), promoter (9%) and downstream (1%) (Figure 5.7A, B). Most of the methylated regions are located on the A1 chromosome (31%) followed by A2 chromosome (26%), X chromosome (24.5%), unplaced scaffold region (11%) and A3 chromosome (7.5%) (Figure 5.7C, D). We found 192 genes to be hypermethylated and 153 hypomethylated in the N127 pale morphs in comparison to the N127 red morphs (Figure 5.7E). The top 10 most hypermethylated genes were involved in histone process, cellular development, and signalling. In contrast, the top 10 most hypomethylated genes were involved in cellular development, cell adhesion and signalling. (Table 5.5).

Next, the genes were subjected to GO analysis. None of the genes showed significant enrichment. In addition to the GO analysis, the genes were also subjected to KEGG analysis, and no significant enrichment was obtained for KEGG analysis either.

Table 5.5. Top 10 most hypermethylated and hypomethylated in N127 pale and N127 red aphids.

Gene	Chromosome	Fold-change
soluble guanylate cyclase 88E	A3	1.95
UPF0454 protein C12orf49 homolog	A1	1.44
translocase of outer mitochondrial membrane 20	A1	0.91
V-type proton ATPase 116 kDa subunit a	A1	0.91
Uncharacterized	A2	0.79
mitogen-activated protein kinase 13-A	A1	0.79
collagen alpha-1(XI) chain	A1	0.69
Uncharacterized	X	0.63
rad23 protein	A1	0.62
Histone H2AV	A3	0.62
zinc finger protein OZF-like	A2	-1.25
uncharacterized	X	-1.19
mitochondrial ribosomal protein S18A	X	-1.00
dystonin	A2	-0.96
f-box only protein 9-like	A2	-0.89
adapter molecule Crk	X	-0.67
tyrosine-protein phosphatase Lar	A2	-0.66
Uncharacterized	X	-0.61
Uncharacterized	Unplaced scaffold	-0.60
coiled-coil domain containing 44	A2	-0.56

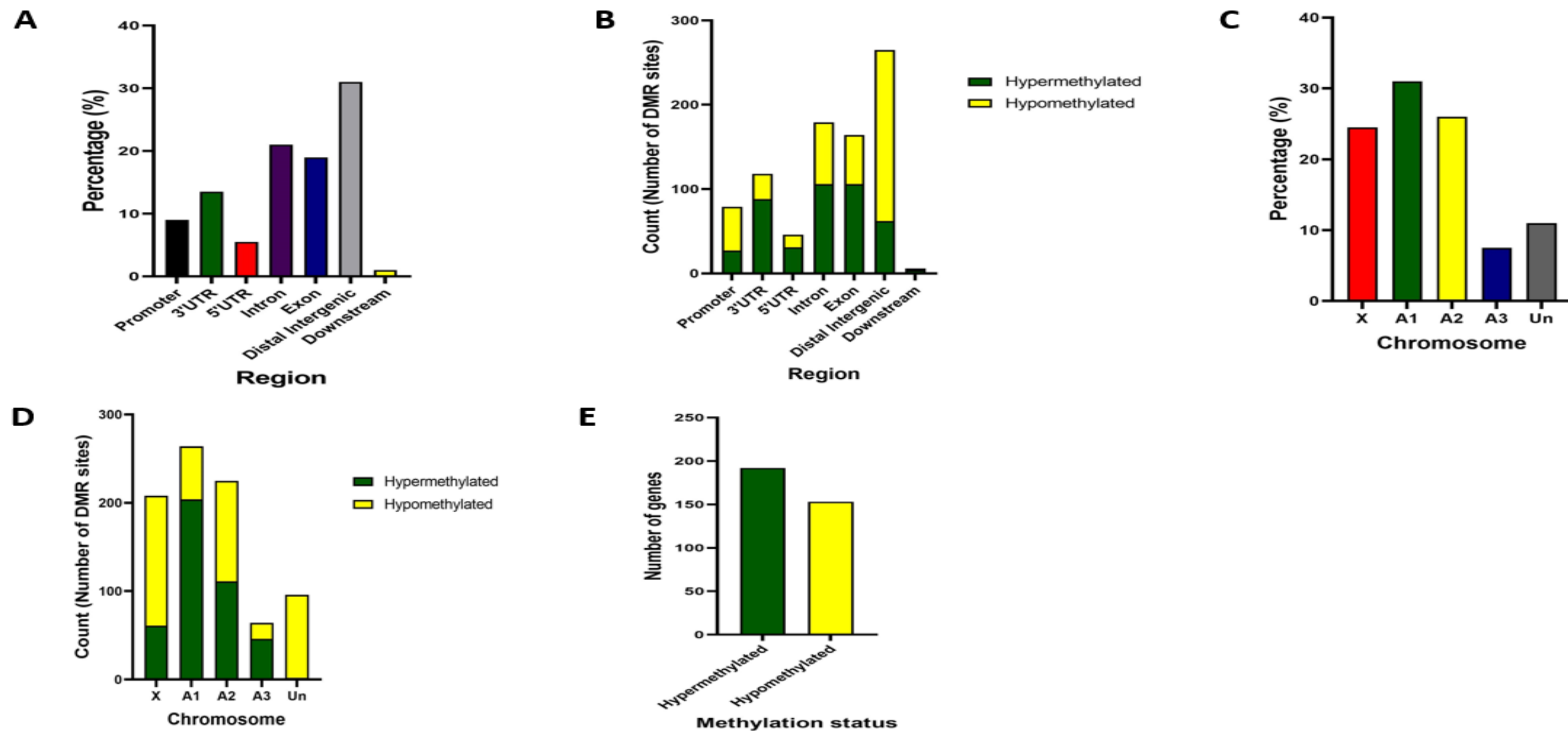


Figure 5.7. MBD-seq results for N127 pale and N127 red aphids. A) The DMR region between N127 pale aphid vs N127 red aphid. Black= promoter, green=3'UTR, red=5'UTR, purple=intron, blue=exon, grey=distal intergenic (DNA region between the coding genes), yellow=downstream. Y-axis (percentage)=The number of site for each region for both hypomethylation and hyper methylation together (Percentage calculated as the number of sites for region divided by the total number of sites). **B)** The methylation status of the DMR. Green=hypermethylated, yellow=hypomethylated. **C)** The distribution of the methylation sites by chromosome. Red=X chromosome, green=A1 chromosome, yellow=A2 chromosome, blue=A3 chromosome, grey=unplaced scaffold. **D)** The methylation status of DMR by chromosome. Green=hypermethylated, yellow=hypomethylated**. E)** The total number of genes that are differentially methylated. Green=hypermethylated, yellow=hypomethylated**.**

c) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N127 red and N116 green wingless aphids

Our DMR analysis showed that the gene body (exon + intron) (50%) is the most methylated region followed by distal intergenic (29%), UTR (16.5%), promoter (11.5%) and downstream (0.5%) (Figure 5.8A, B). Most of the methylated regions are located on the X chromosome (33.5%), followed by A1 chromosome (24%), A2 chromosome (21%), unplaced scaffold region (14.5%) and A3 chromosome (7%) (Figure 5.8C, D). We found 1210 genes to be hypermethylated and 807 hypomethylated in the N127 red morphs in comparison to the N116 wingless morphs (Figure 5.8E). The top 10 most hypermethylated genes have uncharacterized functions and with some involved in cellular signalling and development. In contrast, the top 10 most hypomethylated genes were mostly uncharacterized functions with some involved in protein sorting and translation (Table 5.6).

Next, the genes were subjected to GO analysis. Significant enrichments were obtained for 120 genes which all fall under three functional groups. The 120 genes were classified into two categories: 16 genes in biological processes (BP), and 104 genes in molecular function (MF) (Figure 5.9). In addition to the GO analysis, the genes were also subjected to KEGG analysis, and no significant enrichment was obtained for KEGG analysis.

Table 5.6. Top 10 most hypermethylated and hypomethylated in N127 red and N116 green wingless aphids.

Gene	Chromosome	Fold-change
Uncharacterized	A2	11.92
Uncharacterized	A2	11.22
Uncharacterized	A2	11.15
Uncharacterized	A2	10.92
kelch-like protein 2	A2	10.85
Uncharacterized	X	10.41
peroxidase	A3	10.40
Uncharacterized	A2	10.34
protein unc-119 homolog B	A1	10.32
Uncharacterized	A1	9.80
zinc finger protein OZF-like	A1	-12.10
Uncharacterized	A2	-11.27
mitochondrial ribosomal protein S18A	A1	-11.17
Dystonin	X	-10.96
f-box only protein 9-like	A2	-10.90
adapter molecule Crk	X	-10.68
tyrosine-protein phosphatase Lar	A1	-10.54
Uncharacterized	X	-10.50
Uncharacterized	A2	-10.47
coiled-coil domain containing 44	Unplaced scaffold	-10.38

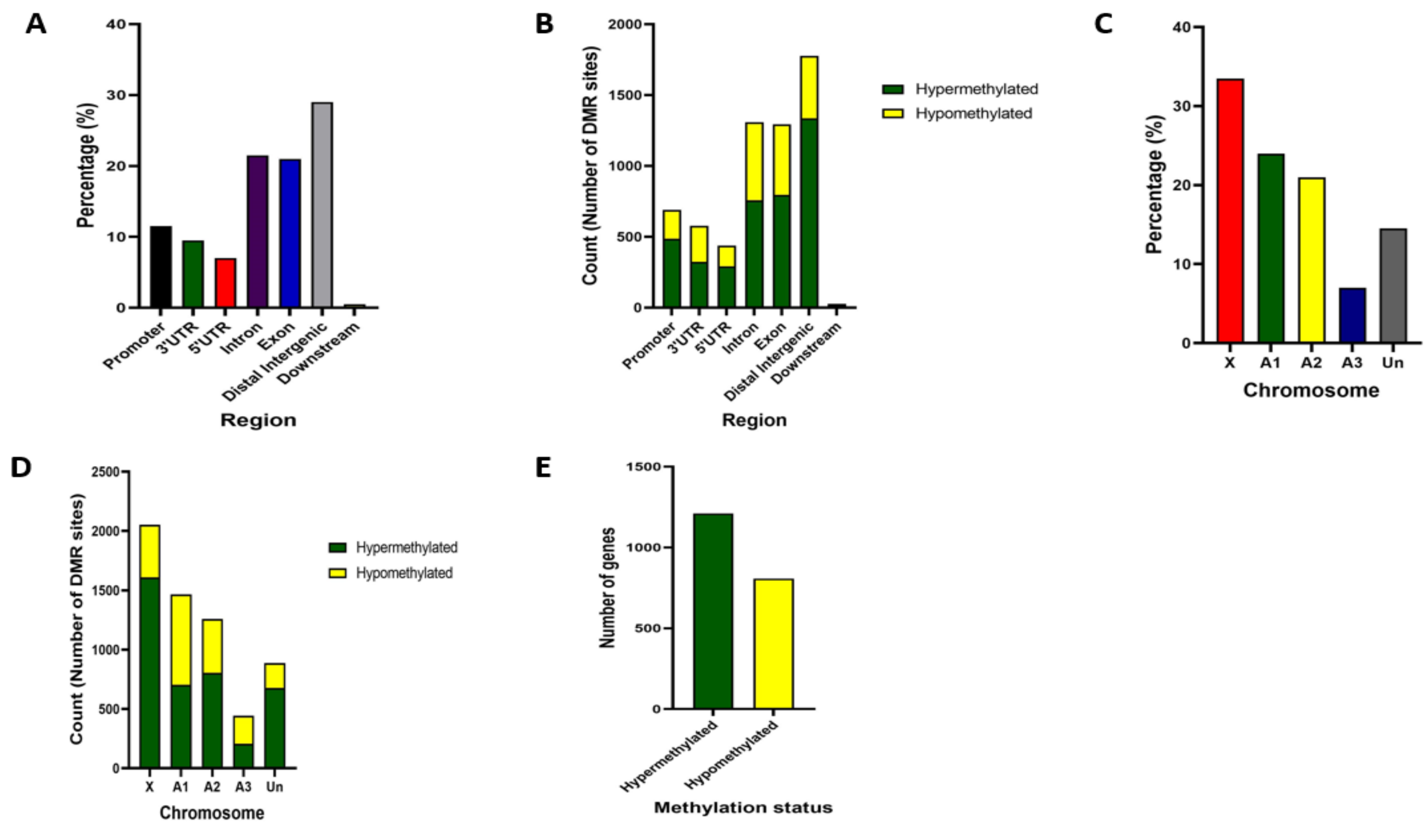


Figure 5.8. MBD-seq results for N127 red vs N116 green wingless aphids. A) The DMR region between N127 red aphid vs N116 wingless aphid. Black= promoter, green=3'UTR, red=5'UTR, purple=intron, blue=exon, grey=distal intergenic (DNA region between the coding genes), yellow=downstream, Y-axis (percentage)=The number of site for each region for both hypomethylation and hyper methylation together (Percentage calculated as the number of sites for region divided by the total number of sites).. **B)** The methylation status of the DMR. Green= hypermethylated, yellow= hypomethylated. **C)** The distribution of the methylation sites by chromosome. Red=X chromosome, green=A1 chromosome, yellow=A2 chromosome, blue=A3 chromosome, grey=unplaced scaffold. **D)** The methylation status of DMR by chromosome. Green= hypermethylated, yellow=hypomethylated**. E)** The total number of genes that are differentially methylated. Green=hypermethylated, yellow=hypomethylated.

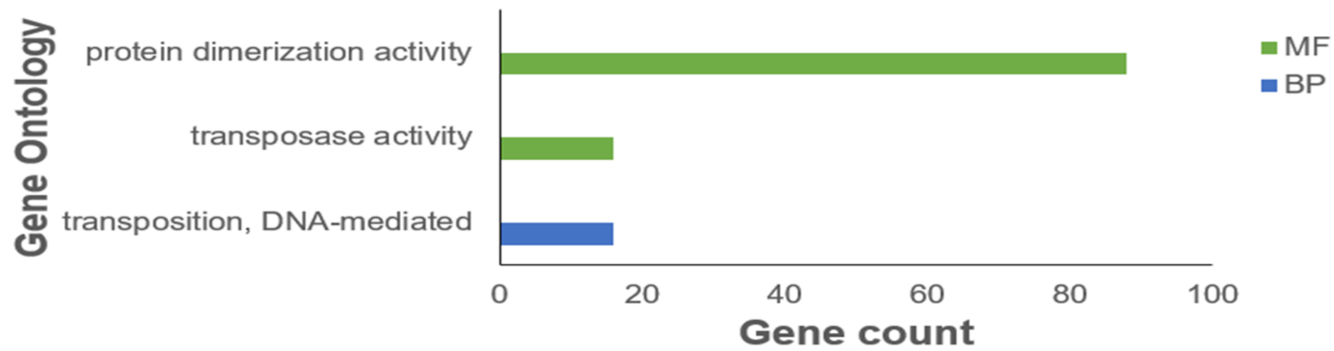


Figure 5.9. GO enrichment analysis between N127 red and N116 green wingless aphids. Gene ontology was classified into two main categories: BP=Biological process (blue), MF=Molecular Function (green).

d) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N127 pale and N116 winged aphids

Our DMR analysis showed that the gene body (exon + intron) (41.5%) is the most methylated region followed by distal intergenic (32%), UTR (14.5%), promoter (11%) and downstream (0.5%) (Figure 5.10A, B). Most of the methylated regions are located on the X chromosome (41.5%), followed by A1 chromosome (20.5%), A2 chromosome (18%), unplaced scaffold region (14%) and A3 chromosome (6%) (Figure 5.10C, D). We found 1109 genes to be hypermethylated and 615 hypomethylated in the N127 pale morphs in comparison to the N116 winged morphs (Figure 5.10E). The top 10 most hypermethylated genes have uncharacterized functions and with some involved in cellular signalling and development. In contrast, the top 10 most hypomethylated genes were mostly involved in catalytic process, cellular development and ion channel transportation (Table 5.7).

Next, the genes were subjected to GO analysis. Significant enrichments were obtained for 13 genes which all fall under one functional group. All the genes were classified into the biological process (BP) category (Figure 5.11). In addition to the GO analysis, the genes were also subjected to KEGG analysis, and no significant enrichment was obtained for KEGG analysis.

Table 5.7. Top 10 most hypermethylated and hypomethylated in N127 pale and N116 winged aphids.

Gene	Chromosome	Fold-change
kelch-like protein 2	A2	10.86
Uncharacterized	X	10.57
Uncharacterized	A2	10.44
DNA topoisomerase 2-binding protein 1-A	A1	10.18
kelch-like protein 2	A2	10.10
Uncharacterized	A2	10.04
Uncharacterized	A3	9.88
Peroxidase	A1	9.86
protein unc-119 homolog B	A1	9.82
Uncharacterized	X	9.59
Uncharacterized	A2	-10.29
Uncharacterized	X	-10.25
transient receptor potential cation channel trpm	A1	-9.98
Uncharacterized	A1	-9.93
enoyl-CoA hydratase domain-containing protein 3, mitochondrial	A2	-9.92
kelch-like protein 3	A1	-9.92
gem-associated protein 5	A1	-9.90
DNA polymerase delta catalytic subunit	A1	-9.84
kelch-like protein 3	A2	-9.59
AP-1 complex subunit gamma-1	A1	-9.58

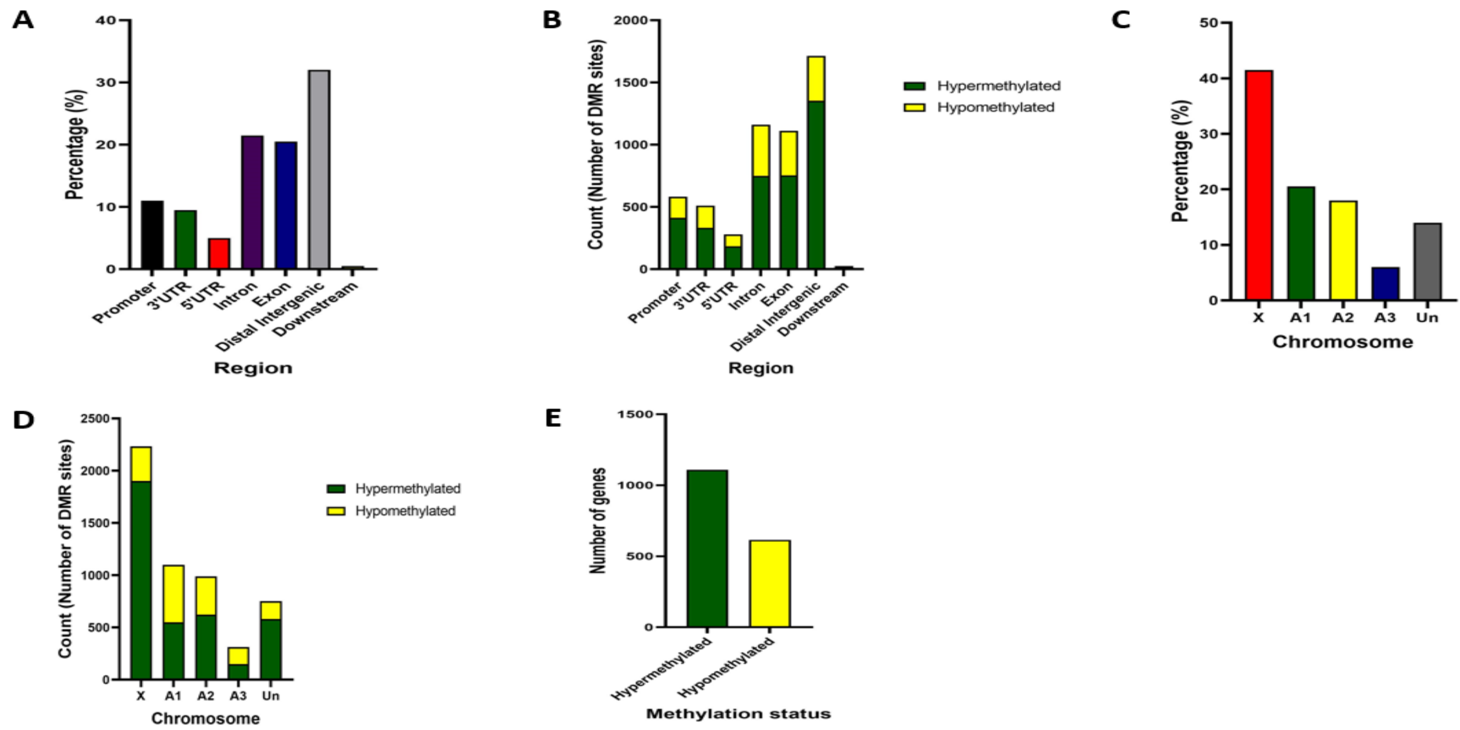


Figure 5.10. MBD-seq results for N127 pale and N116 winged aphids. A) The DMR region between N127 pale aphid vs N116 winged aphid. Black= promoter, green=3'UTR, red=5'UTR, purple=intron, blue=exon, grey=distal intergenic (DNA region between the coding genes), yellow= downstream, Y-axis (percentage)=The number of site for each region for both hypomethylation and hyper methylation together (Percentage calc ulated as the number of sites for region divided by the total number of sites). **B)** The methylation status of the DMR. Green=hypermethylated, yellow=hypomethylated. **C)** The distribution of the methylation sites by chromosome. Red=X chromosome, green=A1 chromosome, yellow=A2 chromosome, blue=A3 chr omosome, grey=unplaced scaffold. **D)** The methylation status of DMR by chromosome. Green=hypermethylated, yellow=hypomethylated**. E)** The total number of genes that are differentially methylated. Green=hypermethylated, yellow=hypomethylated.

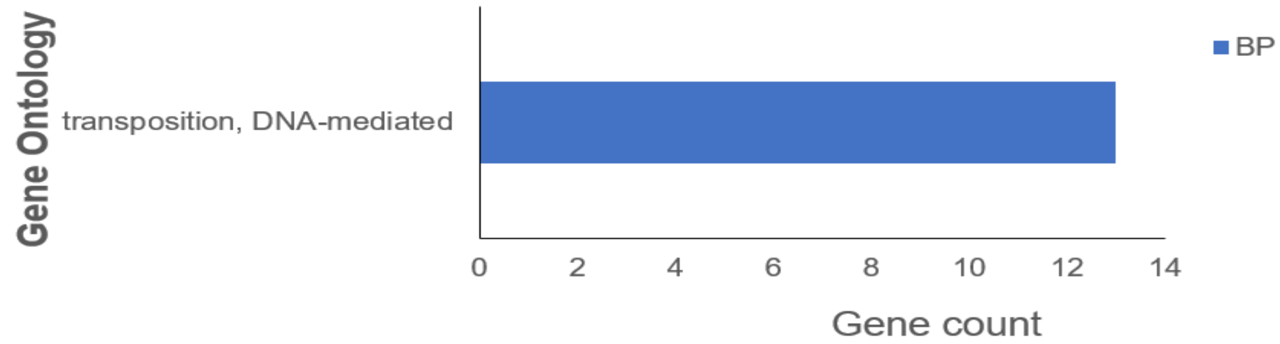


Figure 5.11. GO enrichments analysis between N127 pale and N116 winged morphs. Gene ontology was classified into one main category: BP=Biological process (blue).

e) GO and KEGG analysis of differentially methylated regions and differentially methylated genes in N127 (red+pale) and N116 (winged+wingless) aphids

Our DMR analysis showed that gene body (exon + intron) (42.5%) is the most methylated region followed by distal intergenic (30%), UTR (15.5%), promoter (11.5%) and downstream (0.5%) (Figure 5.12A, B). Most of the methylated regions are located on the X chromosome (38%), followed by A1 chromosome (22.5%), A2 chromosome (19.5%), unplaced scaffold region (14%) and A3 chromosome (6%) (Figure 5.12C, D). We found 1165 genes to be hypermethylated and 688 hypomethylated in the N127 genotype in comparison to the N116 genotype (Figure 5.12E). The top 10 most hypermethylated genes have uncharacterized functions and with some involved in cellular signalling and development. In contrast, the top 10 most hypomethylated genes were mostly involved in catalytic process, cellular development and uncharacterized function (Table 5.8).

Next, the genes were subjected to GO analysis. Significant enrichments were obtained for 30 genes which all fall under two functional groups. The 30 genes were classified into two categories: 15 in biological process (BP) and 15 in molecular function (MF) (Figure 5.13). In addition to the GO analysis, the genes were also subjected to KEGG analysis, and no significant enrichment was obtained for KEGG analysis.

Table 5.8. Top 10 most hypermethylated and hypomethylated in N127 (red+pale) and N116 (winged+wingless) aphids.

Gene	Chromosome	Fold-change
kelch-like protein 2	A2	11.13
Uncharacterized	A2	10.95
Uncharacterized	A2	10.91
Uncharacterized	X	10.76
Uncharacterized	A2	10.66
Uncharacterized	A3	10.41
protein unc-119 homolog B	A1	10.36
Uncharacterized	A2	10.23
Uncharacterized	A1	10.06
DNA topoisomerase 2-binding protein 1-A-like	A1	9.98
kelch-like protein 3	A1	-10.77
Uncharacterized	A1	-10.72
Uncharacterized	X	-10.61
enoyl-CoA hydratase domain-containing protein 3, mitochondrial	A2	-10.46
kelch-like protein 3	A2	-10.34
Uncharacterized	X	-10.10
protein angel	Unplaced scaffold	-10.07
gem-associated protein 5	A1	-10.04
AP-1 complex subunit gamma-1	A1	-10.03
DNA polymerase delta catalytic subunit	A1	-10.01

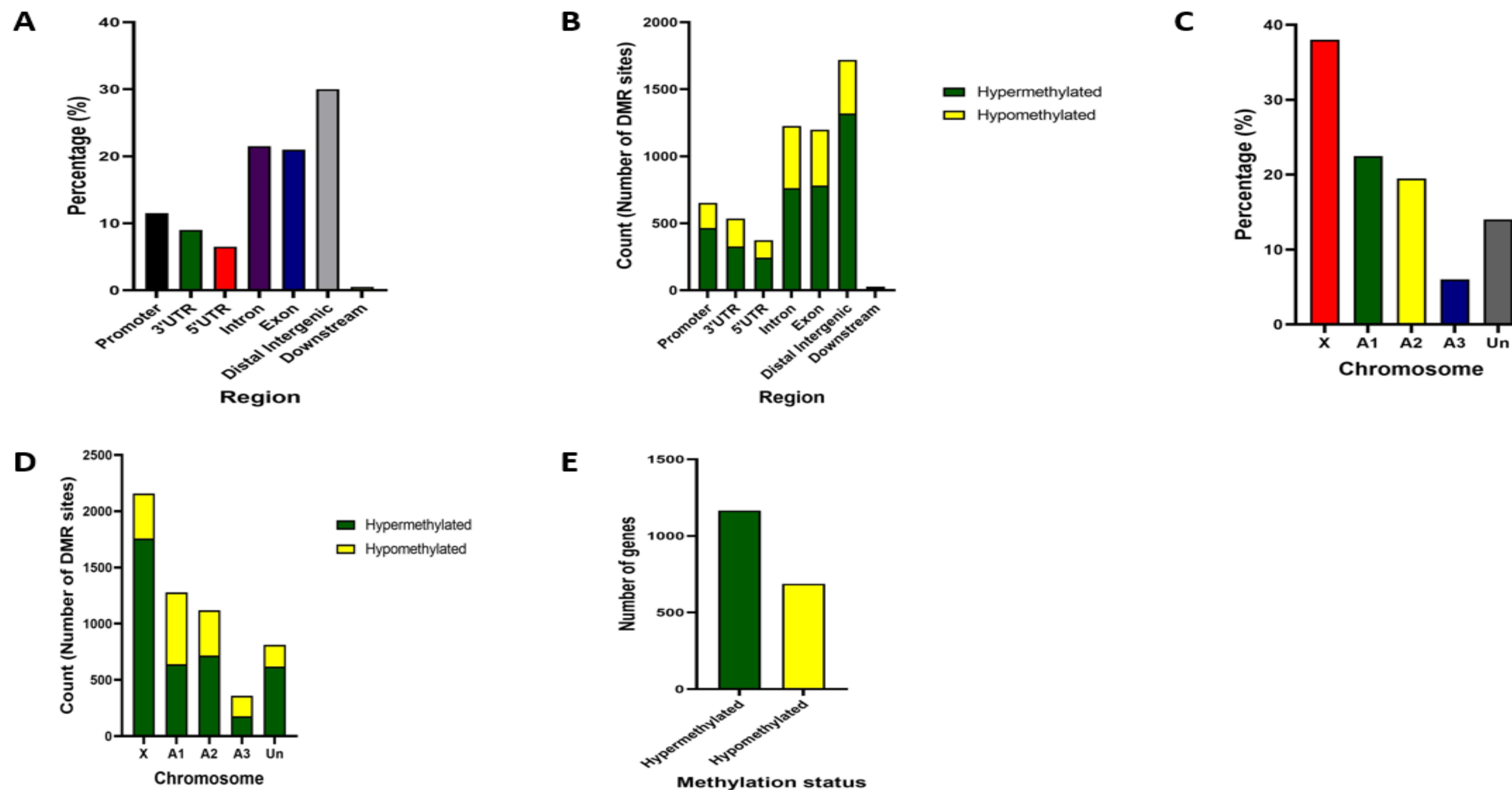


Figure 5.12. MBD-seq results for N127 (pale+red) and N116 (winged+wingless) aphids. A) The DMR region between N127 genotype vs N116 genotype. Black=promoter, green=3'UTR, red=5'UTR, purple=intron, blue=exon, grey=distal intergenic (DNA region between the coding genes), yellow= downstream, Y-axis (percentage)=The number of site for each region for both hypomethylation and hyper methylation together (Percentage calc ulated as the number of sites for region divided by the total number of sites). **B)** The methylation status of the DMR. Green=hypermethylated, yellow=hypomethylated. **C)** The distribution of the methylation sites by chromosome. Red=X chromosome, green=A1 chromosome, yellow=A2 chromosome, blue=A3 chromosome, grey=unplaced scaffold. **D)** The methylation status of DMR by chromosome. Green=hypermethylated, yellow=hypomethylated**. E)** The total number of genes that are differentially methylated. Green=hypermethylated, yellow=hypomethylated.

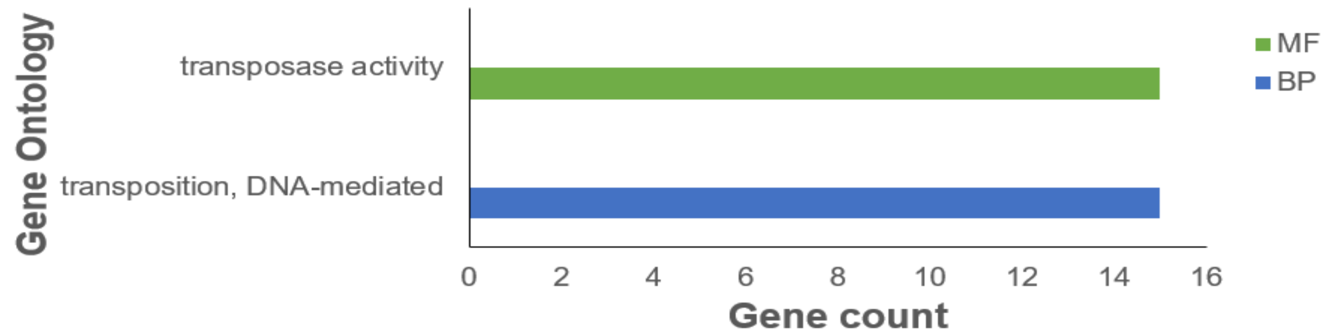


Figure 5.13. GO enrichment analysis between N127 (pale+red) and N116 (wingless+winged) aphids. Gene ontology was classified into two main categories: BP=Biological process (blue), MF=Molecular Function (green).

5.3.4 Correlation between DNA methylation and gene expression between pea aphid morphs and genotypes

The transcriptome profile of the four different pea aphid morphs (Chapter 4) was integrated with the methylome profile (by searching if the differentially methylated genes were also differentially expressed in the RNA-seq) to provide insight into the relationship between DNA methylation and gene expression. The overlapping gene were selected based on the criteria of (P-adjust<0.5 and <-0.3log2fold>0.3) for both differentially expressed gene (DEG) and differentially methylated gene (DMG). The correlation between DNA methylation and gene expression (for gene body only) is analysed using Pearson correlation coefficient for each group comparison (Figure 5.14A-E) as the function of promoter methylation in insects remain unclear. Among the group comparison only the N127red vs N116 wingless, N127 pale vs N116 winged and N127 vs N116 showed significant correlation between DNA methylation and gene expression (P-value <0.001). All three group showed a positive correlation between DNA methylation and gene expression ranging from moderate correlation to strong correlation.

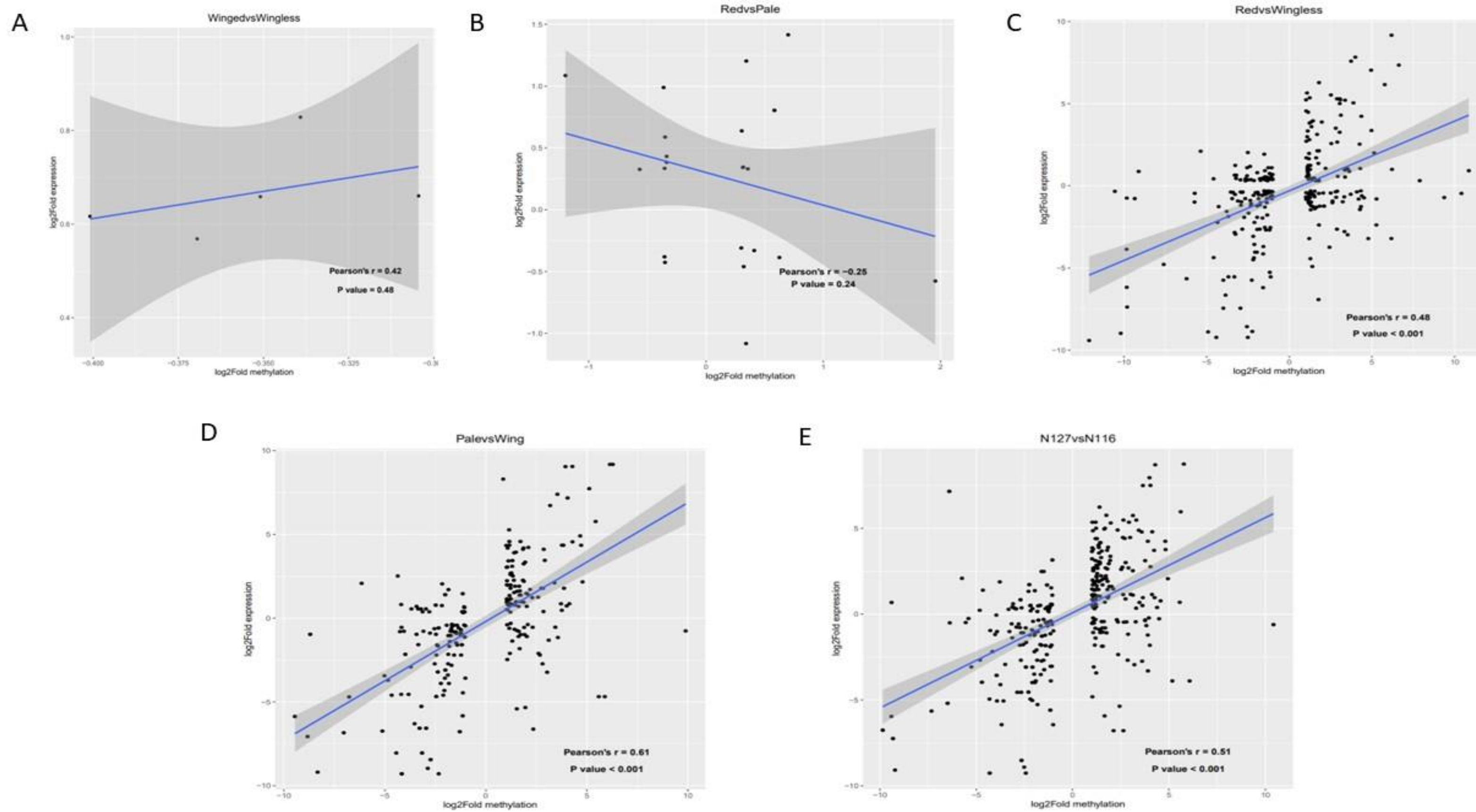


Figure 5.14. Correlation between DNA methylation and gene expression (gene body region only) for different group comparisons. Pearson' r value 0.1-0.3 (low correlation), 0.3-0.5 (moderate correlation), 0.5-1.0 (strong correlation). + Pearson's r= positive correlation, -Pearson's r= negative correlation. X-axis = any genes that has methylation difference of log2fold change between -0.3<log2fold>0.3 and the p-adjusted value of <0.05, Y-axis= any genes that has expression of log2fold between -0.3<log2fold>0.3 and the p-adjusted value of <0.05. Positive correlation = Increase methylation with increase gene expression in gene body. Negative correlation= Inrease methylation but decrease gene expression in gene body.

a) Differentially methylated genes and differentially expressed genes overlap between N116 winged and N116 wingless aphids

We integrated the results of MBD-seq with RNA-seq (Chapter 4) to provide insight into the correlation between DNA methylation and gene expression. For hypermethylated genes, there was no overlap with the RNA-seq data. Next, for hypomethylated genes we found a total of 9 genes overlap with RNA-seq data. Of the nine genes, two show a similar trend in DNA methylation and gene expression (increase methylation with higher gene expression for gene body and increase methylation with lower gene expression for promoter region and vice versa), five genes show the opposite trend (increase methylation with lower gene expression for gene body and increase methylation with higher gene expression for promoter region and vice versa). Further, two genes were not analysed due to having multiple regions methylated (e.g. 3'UTR, promoter, gene body all methylated) as we do not know about the effect of DNA methylation on multiple regions of genes. (Figure 5.15A-C).

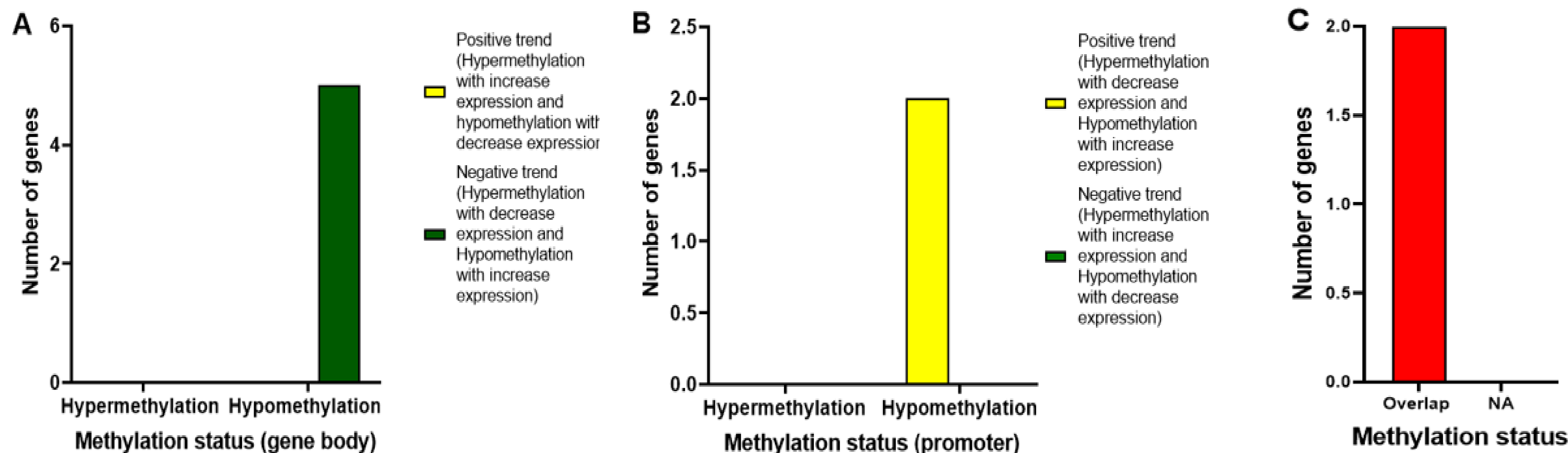


Figure 5.15. MBD-seq and RNA-seq for N116 winged and N116 wingless aphids. **A)** Integration of results for MBD-seq and RNA-seq for the gene body region. Green=genes following the same trend for DNA methylation and gene expression in gene body (increased methylation associated with increased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in gene body (increased methylation with decreased expression and decreased methylation with increased expression). **B) I**ntegration of results for MBD-seq and RNA-seq for the promoter region. Green= genes following the same trend for DNA methylation and gene expression (increased methylation with decreased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in the promoter region (increased methylation with increased expression and decreased methylation with decreased expression). **C)** Genes with multiple regions methylated or unclear relationship between methylated regions and gene expression. Red=overlap (genes with different region methylated such as promoter, gene body and 3'UTR). Blue=NA (genes with only 3'UTR or 5'UTR methylated).

b) Differentially methylated genes and differentially expressed genes overlap between N127 pale and N127 red aphids

For hypermethylated genes, we found 16 genes that overlap with RNA-seq data. Of the 16 genes, eight showed a similar trend for DNA methylation and gene expression and eight genes showed the opposite trend. For hypomethylated genes we found a total of nine genes overlapping with the RNA-seq data. Two of the genes show a similar trend in DNA methylation and gene expression (increase methylation with higher gene expression for gene body and increase methylation with lower gene expression for promoter region and vice versa), and seven genes show the opposite trend (increase methylation with lower gene expression for gene body and increase methylation with higher gene expression for promoter region and vice versa). Further, 13 genes across both hypermethylated and hypomethylated were not analysed due to multiple methylated regions (e.g. 3'UTR, promoter, gene body all methylated) as we do not know about the effect of DNA methylation on multiple regions of genes or unclear relationship between the methylated region and gene expression (e.g. 3'UTR and 5'UTR were methylated only but we do not know the effect of methylation on these two region in insects) (Figure 5.16A-C).

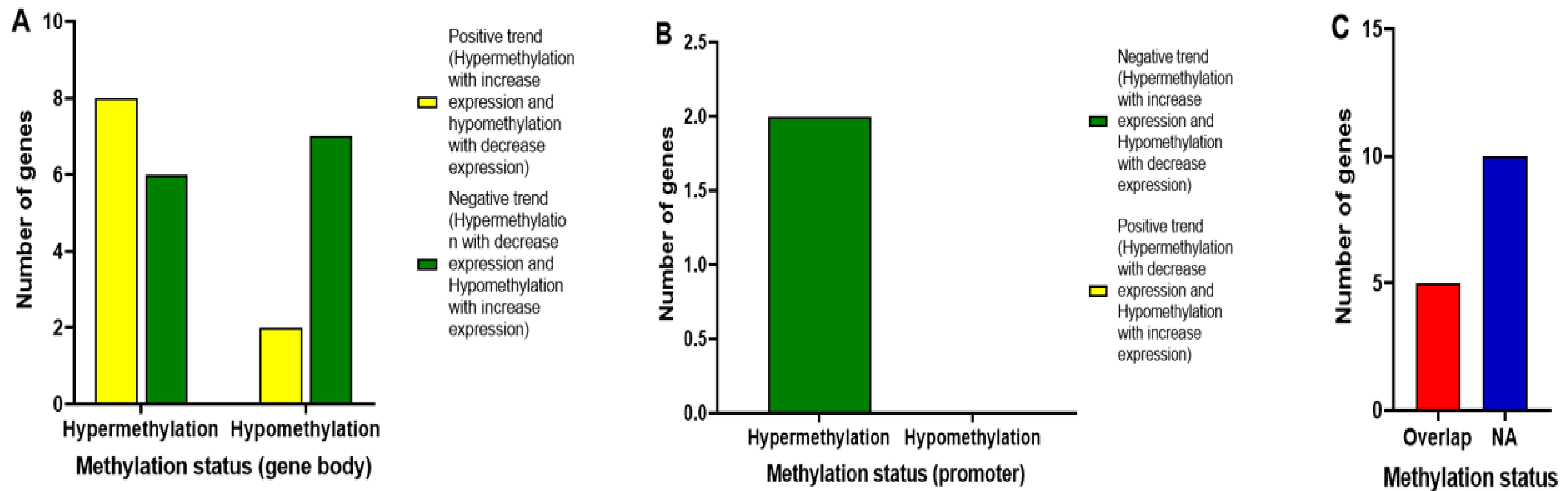


Figure 5.16. MBD-seq and RNA-seq for N127 pale vs N127 red aphids. **A)** Integration of results for MBD-seq and RNA-seq for the gene body region. Green=genes following the same trend for DNA methylation and gene expression in gene body (increased methylation associated with increased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in gene body (increased methylation with decreased expression and decreased methylation with increased expression). **B)** Integration of results for MBD-seq and RNA-seq for the promoter region. Green=genes following the same trend for DNA methylation and gene expression (increased methylation with decreased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in the promoter region (increased methylation with increased expression and decreased methylation with decreased expression). **C)** Genes with multiple regions methylated or unclear relationship between methylated regions and gene expression. Red=overlap (genes with different region methylated such as promoter, gene body and 3'UTR). Blue=NA (genes with only 3'UTR or 5'UTR methylated).

c) Differentially methylated genes and differentially expressed genes overlap between N127 red and N116 green wingless aphid

For hypermethylated genes, we found 200 genes that overlap with RNA-seq data. Of the 200 genes, 103 showed similar trend for DNA methylation and gene expression and 97 genes showing the opposite trend. For hypomethylated genes we found a total of 156 genes that overlap with RNA-seq data. Of the 156 genes, 101 show a similar trend in DNA methylation and gene expression (increase methylation with higher gene expression for gene body and increase methylation with lower gene expression for promoter region and vice versa) and 55 genes showed the opposite trend (increase methylation with lower gene expression for gene body and increase methylation with higher gene expression for promoter region and vice versa). Further, 520 genes across both hypermethylated and hypomethylated were not analysed due to multiple methylated regions (e.g. 3'UTR, promoter, gene body all methylated) as we do not know about the effect of DNA methylation on multiple region of genes or unclear relationship between methylated region and gene expression (e.g. 3'UTR and 5'UTR were methylated only but we do not know the effect of methylation on these two region in insects) (Figure 5.17A-C).

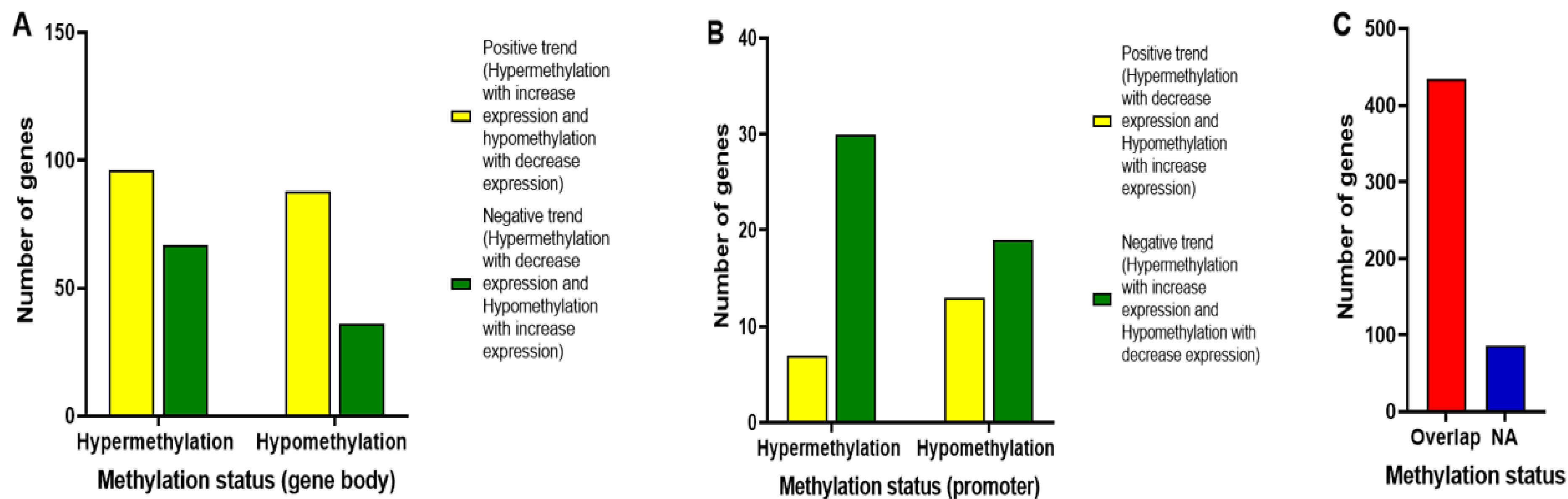


Figure 5.17. MBD-seq and RNA-seq for N127 red and N116 wingless aphids. **A)** Integration of results for MBD-seq and RNA-seq for the gene body region. Green=genes following the same trend for DNA methylation and gene expression in gene body (increased methylation associated with increased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in gene body (increased methylation with decreased expression and decreased methylation with increased expression). **B)** Integration of results for MBD-seq and RNA-seq for the promoter region. Green= genes following the same trend for DNA methylation and gene expression (increased methylation with decreased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in the promoter region (increased methylation with increased expression and decreased methylation with decreased expression). **C)** Genes with multiple regions methylated or unclear relationship between methylated regions and gene expression. Red=overlap (genes with different region methylated such as promoter, gene body and 3'UTR). Blu e=NA (genes with only 3'UTR or 5'UTR methylated).

d) Differentially methylated genes and differentially expressed genes overlap between N127 pale and N116 winged aphids

For hypermethylated genes, we found 129 genes that overlap with RNA-seq data. Of the 129 genes, 86 showed a similar trend for DNA methylation and gene expression and 43 genes showed the opposite trend. For hypomethylated genes we found a total of 94 genes overlapped with RNA-seq data. Of the 94 genes, 65 showed a similar trend in DNA methylation and gene expression (increase methylation with higher gene expression for gene body and increase methylation with lower gene expression for promoter region and vice versa) and 29 genes show the opposite trend (increase methylation with lower gene expression for gene body and increase methylation with higher gene expression for promoter region and vice versa). Further, 362 genes across both hypermethylated and hypomethylated were not analysed due to multiple methylated regions (e.g. 3'UTR, promoter, gene body all methylated) as we do not know about the effect of DNA methylation on multiple regions of genes or unclear relationship between the methylated region and gene expression (e.g. 3'UTR and 5'UTR were methylated only but we do not know the effect of methylation on these two region in insects) (Figure 5.18A-C).

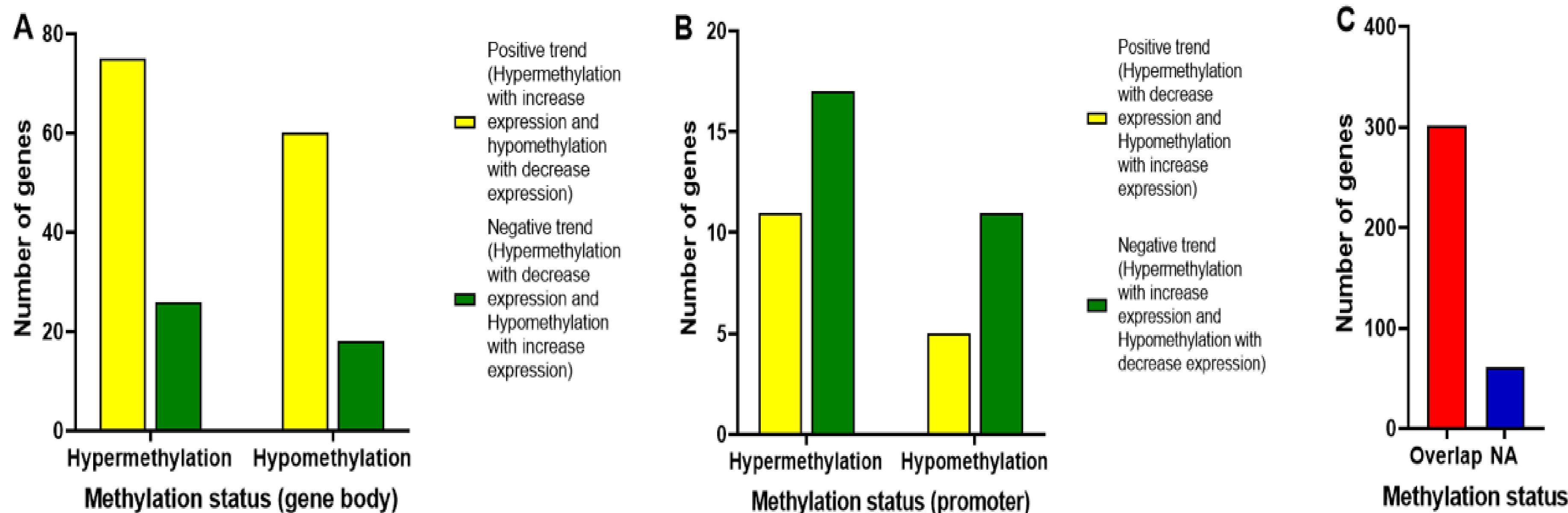


Figure 5.18. MBD-seq and RNA-seq for N127 pale vs N116 winged aphids. **A)** Integration of results for MBD-seq and RNA-seq for the gene body region. Green=genes following the same trend for DNA methylation and gene expression in gene body (increased methylation associated with increased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in gene body (increased methylation with decreased expression and decreased methylation with increased expression). **B)** Integration of results for MBD-seq and RNA-seq for the promoter region. Green= genes following the same trend for DNA methylation and gene expression (increased methylation with decreased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in the promoter region (increased methylation with increased expression and decreased methylation with decreased expression). **C)** Genes with multiple regions methylated or unclear relationship between methylated regions and gene expression. Red=overlap (genes with different region methylated such as promoter, gene body and 3'UTR). Blue=NA (genes with only 3'UTR or 5'UTR methylated).

e) Differentially methylated genes and differentially expressed genes overlap between N127 (red+pale) and N116 (winged+wingless) aphids

For hypermethylated genes, we found 288 genes that overlap with RNA-seq data. Of the 288 genes, 200 showed a similar trend for DNA methylation and gene expression and 88 genes showed the opposite trend. For hypomethylated genes we found a total of 189 genes overlapped with RNA-seq data. Of the 189 genes, 129 showed a similar trend in DNA methylation and gene expression (increase methylation with higher gene expression for gene body and increase methylation with lower gene expression for promoter region and vice versa) and 60 genes showed the opposite trend (increase methylation with lower gene expression for gene body and increase methylation with higher gene expression for promoter region and vice versa). Further, 476 genes across both hypermethylated and hypomethylated were not analysed due to having multiple regions methylated (e.g. 3'UTR, promoter, gene body all methylated) as we do not know about the effect of DNA methylation on multiple regions of genes or unclear relationship between the methylated region and gene expression (e.g. 3'UTR and 5'UTR were methylated only but we do not know the effect of methylation on these two region in insects) (Figure 5.19A-C).

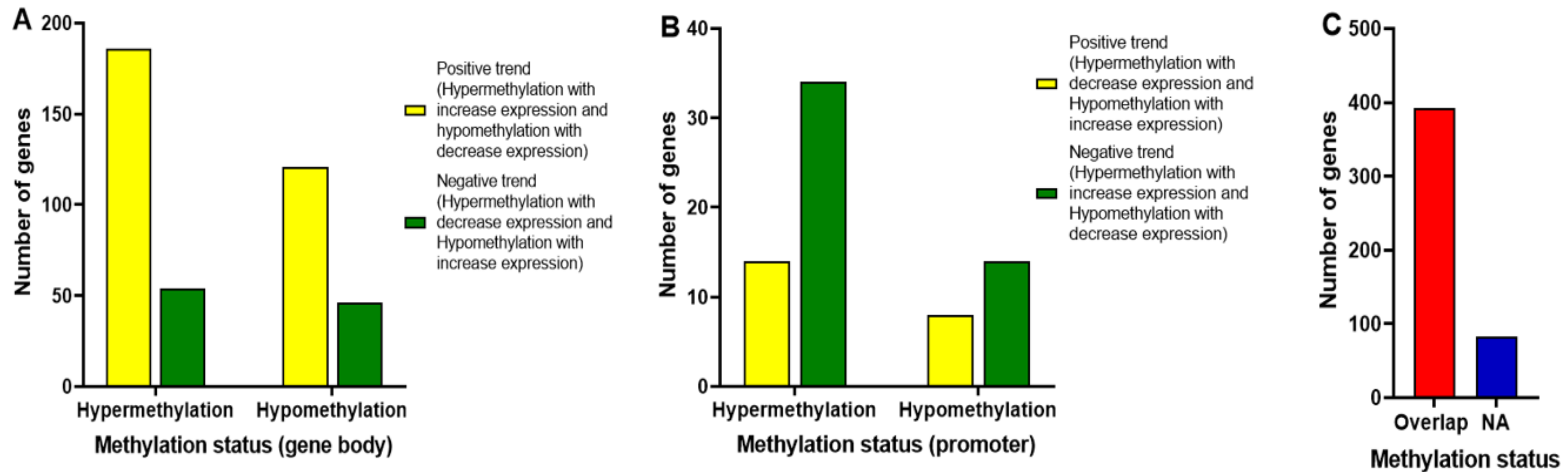


Figure 5.19. MBD-seq and RNA-seq for N127 (pale+red) and N116 (winged+wingless) aphids. **A)** Integration of results for MBD-seq and RNA-seq for the gene body region. Green=genes following the same trend for DNA methylation and gene expression in gene body (increased methylation associated with increased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in gene body (increased methylation with decreased expression and decreased methylation with increased expression). **B)** Integration of results for MBD-seq and RNA-seq for the promoter region. Green=genes following the same trend for DNA methylation and gene expression (increased methylation with decreased expression and decreased methylation with decreased expression). Yellow=genes with opposite trend in the promoter region (increased methylation with increased expression and decreased methylation with decreased expression). **C)** Genes with multiple regions methylated or unclear relationship between methylated regions and gene expression. Red=overlap (genes with different region methylated such as promoter, gene body and 3'UTR). Blue=NA (genes with only 3'UTR or 5'UTR methylated).

5.4 Discussion

In this Chapter, the DNA methylation levels of candidate genes involved in pea aphid development between different pea aphid morphs was investigated using whole aphid tissues from Chapter 3. In addition, the methylome profile of aphid morphs was investigated using tissue from Chapter 3. The specific aims were to investigate the role of epigenetic mechanisms, specifically DNA methylation, in regulating polyphenism in insects. Our results reveal that most of the genes were lowly methylated and were not differentially methylated between aphid morphs or genotype. Second, I aimed to investigate the methylome profile of different pea aphid morphs in two genotypes. Our MBD-seq reveals that the methylome profile differs between aphid genotype and morphs. Further, we also found that at a chromosome level, the X-chromosome was the most highly methylated. Lastly, we found that only 1 overlapping gene were observed between the N116 winged vs N116 wingless group compared to N127 pale vs N127 red group. The gene that was overlap in the Bruton tyrosine kinase (Btk). Btk has very diverse role in insects function including cellularization, morphogenesis and also germ cell patterning.

5.4.1 DNA methylation level of candidate genes

DNA methylation level was assessed by pyrosequencing of whole adult aphids of each morph (N116 winged, N116 wingless, N127 red, N127 pale) from Chapter 3, with specific primers for genes involved in different aspects of aphid development. Of the 15 genes analysed only four genes (Hsp83, Mad, Rpd3 and DMAP1) showed significant differences in methylation between morphs. In addition, of the four significantly differentially methylated genes only one gene (Rpd3) showed a positive relationship between DNA methylation and gene expression. Even though DNA methylation has been speculated to play an important role in regulating insect polyphenism, the evidence to date remains unclear and no definite conclusion can be drawn from these results regarding the exact role of epigenetic regulation in insect polyphenism. For example, a study by Ferreira et al. (2013) showed no significant

differences in DNA methylation of genes involved in caste differentiation of *Polistes canadensis* wasps. Further, juvenile hormone-associated genes have been speculated to play an important role in regulating trans-generational polyphenism and wing development in insects. However, a study by Walsh et al. (2010) reported no significant differences in the methylation of JH-associated genes between winged and unwinged pea aphids. A study of methylation profiles in different insects such as *Apis mellifera* and *Bombyx mori* by Hunt et al. (2010) suggested that genes that are ubiquitously expressed are usually more heavily methylated. In contrast, genes differentially expressed between different developmental stages or involved in regulating different morphs usually exhibit lower methylation levels. Our results agree with this study as the genes that are involved in regulating wing development and ecdysone signalling are lowly methylated between all the morphs (1 – 3%). A possible explanation for the low level of methylation observed in these genes could be related to the flexibility in regulating gene expression between alternative morphs. For example, studies have proposed that lowly methylated gene allows a wider flexibility in transcription by allowing access to alternative transcription start sites, increasing mutation and exon skipping (Gavery and Roberts 2010, Roberts and Gavery 2012).

One of the reasons that we did not find any differences in methylation level in genes involved in the endocrine system (e.g., ecdysone pathway) or wing development between morphs might be due to the fact these genes are tissue-specific. For example, Mukherjee et al. (2019) found tissue-specific differences in DNA methylation in the greater wax moth *Galleria mellonella* between the parasitic resistance and susceptible line. Apart from that, another study by Xu et al. (2018) reported that DNA methylation act in a tissue-specific manner in regulating silkmoth *Bombyx mori* wing development. Keller et al. (2016) also reported similar tissue-specific DNA methylation in the sea vase *Ciona Intestinalis*. In our analyses the whole aphid organism is used, therefore the methylation level for the gene might be underestimated such that the methylation differences in genes involved in wing development might be higher if only the wing tissues are used for the analysis.

An alternative explanation for the low methylation level could be due to the developmental stage. For example, Feliciello et al. (2013) showed that the red flour beetle *Tribolium castaneum* showed a higher level of DNA methylation in the embryonic stage in comparison to other developmental stages. Another study by Kronforst et al. (2008) also showed that the DNA methylation pattern in social Hymenoptera varies between developmental stages with young larvae to young pupae stage showing higher methylation levels in comparison to the adult stage. A study by Morandin et al. (2019) showed that many genes involved in caste differentiation in the ant *Formica exsecta* are dependent on the developmental stage. In aphids, all the newly born nymphs possess wing buds and it is only until the 3rd instar stage that the wing-destined nymph will continue to grow their wing bud while the wingless destined nymph will degenerate their wing buds. Since in our experiment only adult aphids are used for the analysis, therefore the methylation level of the genes might be underestimated as they might show developmental specific methylation differences.

5.4.2 Methylome profile expression

The methylome profile of different pea aphid morphs was investigated in this Chapter using MBD-seq. First, we analysed the differently methylated regions in the aphid morphs to determine the most highly methylated regions in the aphid genome. Our result suggests that across all the group comparisons the gene body (exon + intron) is the most methylated region. In contrast, most methylation in mammals is located in the promoter region. For example, a study by Wang et al. (2013) on the parasitoid wasp *Nasonia vitripennis* reveals that most of the methylated regions were found in the gene body with most found specifically in the coding region. Another study by Flores et al. (2012) also reported a similar methylation pattern in honeybee *Apis mellifera*.

Next, we found that the intron methylation is slightly higher in comparison to the exon region in the pea aphid across all different group comparisons. Although the methylation in insects is mostly observed in the gene body (exon + intron), the methylation location varies across

different insect species. For example, studies have shown that in *Drosophila melanogaster* and *Tribolium castaneum*, DNA methylation is found to be higher in the intron region rather than the exon (Song et al. 2017; Guan et al. 2019). In contrast, studies on social insects such as the Florida carpenter ant *Camponotus floridanus*, Nevada termite *Zootermopsis nevadensis*, the carpenter bee *Ceratina calcarata*, and the honeybee *Apis mellifera* showed that exon methylation is more common in comparison to intron methylation (Bonasio et al. 2012; Glastad et al. 2016; Rehan et al. 2016; Wang et al. 2020). In addition to social insects, some other non-social insects such as the parasitoid wasp *Nasonia vitripennis* and silkworm *Bombyx mori* also showed higher methylation in exon compared to intron (Xiang et al. 2010; Beeler et al. 2014).

One of the biggest differences between gene body methylation and promoter methylation is that CpG in the promoter region is largely free of methylation. In contrast, gene body methylation is usually associated with genes that show medium to high levels of expression across both vertebrates, invertebrates and plants (Suzuki and Bird. 2008; Zemach et al. 2010; Bonasio 2014; Zilberman et al. 2007). Some recent studies have suggested a link between gene body methylation and alternative splicing (Yan et al. 2015). For example, Herb et al. (2012) reported around 51% of the changes in gene body methylation are associated with alternative splicing and are involved in the transition from foraging to nursing in honeybees. Marshall et al. (2019) also reported the link between gene body methylation and alternative splicing in regulating reproduction between sterile and reproductive bumblebee workers *Bombus terrestris*. Furthermore, Harrison et al. (2022) also reported the link between gene body methylation and alternative splicing in regulating the caste differentiation in drywood termite *Zootermopsis nevadensis*. However, contradicting results have also been reported by many studies as well whereby no relationship was found between gene body methylation and alternative splicing (Arsenault et al. 2018; Standage et al. 2016; Harris et al. 2019). Therefore, further research across

more and different insect species is needed to elucidate the link between gene body methylation and alternative splicing.

Apart from alternative splicing, gene bodies have also been reported to be correlated with chromatin accessibility. For example, Gatzmaan et al. (2018) reported that in marbled crayfish, low-methylated genes were often found in chromatin and are associated with an increase in chromatin accessibility and increased expression variations. Gene body methylation might also regulate insect gene expression through histone modification. For example, Xu et al. (2021) reported that in *Bombyx mori*, the methyl binding domain protein 2/3 binds specifically to the intragenic region and is involved in promoting the methylation of histone 3 at position 27 (H3K27) and gene expression.

After gene body methylation, we found that the distal intergenic region is the second most methylated region in pea aphids across all the group comparisons. An intergenic region is defined as the region that is located between genes and may contain important functional elements and junk DNA. Interestingly, the distal intergenic region has been speculated to be inconsequential to gene expression (Jaenisch and Bird 2003; Jones 2012). However, some studies in humans have shown that distal intergenic methylation plays an important role in cancer progression (Yegnasubramanian et al. 2011). Weber et al. (2016) also reported the importance and possibility of distal intergenic methylation in regulating plant traits. However, to date, there is no reported association between distal intergenic region and gene expression in insects.

As mentioned above, promoter methylation is usually associated with gene repression and is usually more common in mammals than in insects. Our results showed a substantial level of promoter methylation (7-12%) in pea aphids across all group comparisons. Recently, Keller et al. (2016) reported promoter methylation in the sea vase *Ciona intestinalis*, and the promoter methylation was tissue and cell-type specific. Another study by Zhang et al. (2019) reported that the gene Kr-h1 was able to regulate the activity of steroidogenic enzymes through binding to their promoter sites and inducing promoter methylation. Xu et

al. (2018) also showed that promoter methylation plays an important role in *Bombyx mori* wing development but is tissue and developmental stage specific. Together the results suggest that promoter methylation might play an important role in regulating insects' development. Further research on manipulating promoter methylation in aphid developmental genes is needed to elucidate their role in regulating aphid polyphenism.

Next, we investigated the distribution of methylation at the chromosome level. Our result showed that the X-chromosome is the most methylated chromosome across all groups, apart from the group N127 pale vs N127 red where the most highly methylated chromosome was the A1 chromosome. In aphids, the methylation at a chromosome level has mostly been studied in the context of sexual polyphenism. For example, a recent study by Mathers et al. (2019) showed that high levels of hypermethylation in male aphids were located at the X-chromosome in comparison to the autosomes. Apart from that the function of X-chromosome methylation and its function in regulating wing polyphenism remain largely unknown.

We filtered out the top ten most significantly methylated genes (both hypomethylation and hypermethylation) to determine if any of them are involved in regulating polyphenism in pea aphids. Across most of the group comparisons, our results showed that the top ten most significantly differentially methylated genes have uncharacterized function or are involved in the cellular and translation process. Our DMR results also did not detect any genes that could potentially be responsible for regulating the polyphonic morph in aphids. One of the possible explanations for our result is that genes that are responsible for polyphenism are lowly methylated compared to those of housekeeping genes. For example, a study by Foret et al. (2009) showed that most differentially methylated genes in honeybees were those that were ubiquitously expressed in comparison to tissue or condition specific. A study by Hunt et al. (2010) also reveals that genes that are responsible for polyphenism all have lowly methylation in comparison to those that are ubiquitously expressed. Arsenault et al. (2018) also reported only small changes in DNA methylation between genes responsible

for developmental plasticity in the small carpenter bee, *Ceratina calcarata*. Our result agrees with these studies as no genes that are responsible for regulating the polyphenism in pea aphids are found to be differentially methylated.

Another possible explanation is that genes responsible for plasticity usually have low methylation to allow more flexibility in their expression. Another study by Entrambasaguas et al. (2021) reported that in the seagrass *Cymodocea nodosa*, genes with lower methylation usually showed higher flexibility in their gene expression and plasticity under changing environmental conditions. Gatzmann et al. (2018) show that in marbled crayfish the low methylated genes usually have higher chromatin accessibility compared to those that are heavily methylated. Therefore, the genes responsible for polyphenism in aphids may be lowly methylated to allow for higher chromatin accessibility that could allow for a wider range of expression and more flexible plasticity when exposed to stressful conditions. Many of the DMR in our results are uncharacterized and these genes could potentially play an important role in regulating polyphenism in aphids. Lastly, since the MBD-seq usually only captures the highly methylated region fragments, genes with low methylation might not be detected in the study.

5.3.3 DNA methylation and gene expression

In this study, we also aimed to identify the relationship between DNA methylation and gene expression to understand their role in regulating pea aphid polyphenism. Firstly, we quantified the changes in gene expression in different pea aphid morphs using RNA-seq (Chapter 4). Then we characterized the DMR in the pea aphid genome using MBD-seq. Finally, we integrated the RNA-seq data and MBD-seq to provide insight into the complex relationship between DNA methylation and gene expression in different pea aphid morphs. Across all our group comparisons, we found no clear association between gene body methylation and gene expression. However, it is important to note that in some group comparisons there is a stronger positive relationship between gene body methylation and

gene expression (increase methylation with increased expression). The complex relationship between DNA methylation and gene expression has been reflected by many studies across a wide range of different species. For example, Cunningham et al. (2019) found no association between gene expression and cytosine methylation in the burying beetle *Nicrophorus vespilloides* that are exposed to different parental care conditions. Apart from that, Morandin et al. (2019) reported some overlap between methylation and gene expression at a functional level between different ant-caste *Formica exescta* and there was no clear association between DNA methylation and gene expression. Further, another study by Cardoso-Júnior et al. (2021) also reported a high number of differentially expressed genes but a similar methylome profile for honeybees raised with and without a queen. Furthermore, a similar observation has also been reported in the silkmoth *Bombyx mori*, whereby many genes were differentially methylated but only a small portion of these genes were also differentially expressed with no consistent upregulation or downregulation patterns (Li et al. 2020). Similar observations were also reported by Herb et al. (2018) whereby no association between gene expression and methylation in honeybees upon exposure to an intruder. Taken together, our results agree with most studies and suggest no clear association between gene expression and methylation in regulating pea aphid polyphenism.

Although most studies have reported no direct correlation between DNA methylation and gene expression, some studies support the association between DNA methylation and gene expression. For example, Mashoodh et al. (2021) found that around 51% of the differentially methylated gene in the burying beetle, *Nicrophorus vespilloides* brain were associated with gene expression. However, it is important to note that in the study, the association between gene expression and methylation in these genes was absent after more than 30 generations of selection (Mashoodh et al. 2021). Apart from that, some studies have suggested the possibility of DNA methylation in regulating gene expression in the honeybee, but no direct

association between methylation and gene expression has been shown in these studies (Li-Byarlay et al. 2020; Herb et al. 2018).

One of the most direct ways to investigate the causal relationship between DNA methylation and gene expression is to use the functional genetics approach by knocking down the essential enzyme that is required for the DNA methylation process. For example, some studies have shown that knocking down the Dnmt1a gene in the jewel wasp, *Nasonia vitripennis* results in a global reduction of DNA methylation and this reduction is usually associated with a reduction in gene expression (Arsala et al. 2022). Hou et al. (2020) reported a similar observation in the reduction of gene expression and DNA methylation in the migratory locust *Locusta migratoria* when Dnmt3 was knocked down. However, another study of knocking down Dnmt1 in milkweed bug *Oncopeltus fasciatus* showed a significant change in methylation patterns of genes but these changes are not associated with gene expression (Bewick et al. 2019). Taking together our studies and these results, we suggest a complex relationshipp between DNA methylation and gene expression.

Apart from that, some studies suggest that DNA methylation does not influence gene expression directly but instead acts on the transcription factor affinity for cis-regulatory elements (Bludau et al. 2019). The modification could affect multiple aspects of transcription such as increasing the transcription stability or improving the fidelity of transcription through a reduction in gene expression variability. Other studies have also suggested that DNA methylation might have other functions unrelated to genomic function such as maintaining the genome structure and integrity (Bewick et al. 2019) and Harris et al. (2019) reported only an association between DNA methylation and gene expression in a tissue-specific manner in the honeybee. Based on this observation, Harris suggested that methylation might not be regulating gene expression in honeybees but rather acting as a homeotic function to help maintain normal gene expression through suppressing intergenic transcript initiation. Therefore, further investigation on other epigenetic mechanism such as

histone modification, miRNA could help provide insight into the role of epigenetic regulation in gene expression.

5.5 Summary

In this chapter the DNA methylation levels of genes involved in wing development, stress response and metabolite regulation were investigated between the different pea aphid morphs. We found that most genes were not significantly differentially methylated between the aphid morphs. Our results agree with many previous studies that suggest that in the pea aphid genes responsible for polyphenism are usually lowly methylated. The methylome profile of different pea aphid morphs was also investigated in this chapter. We found that, as observed in most insect species, the pea aphid also has the highest methylation in the gene body region (exon + intron) across all group comparisons. Next, we found that in most group comparisons most of the methylated regions are located on the X-chromosome. Further, this chapter also reveals many DMG genes among pea aphid morphs. However, most of the genes have an uncharacterized function with some involved in cellular development and signalling. Lastly, we integrated the transcriptome results from Chapter 4 to provide insight into the relationship between DNA methylation and gene expression. Our results reveal no clear association between DNA methylation and gene expression in the pea aphid. However, it is important to note that our result did show that gene body methylation and gene expression has a more tightly correlated positive relationship. In contrast, promoter methylation seems to have a more tightly correlated negative relationship with gene expression. Our results also reveal that many genes have multiple regions methylated which provides additional complexity in understanding the relationship between DNA methylation and gene expression. The results from this Chapter suggest that DNA methylation by itself might not be the most important epigenetic mechanisms in regulating polyphenism in pea aphids, and further research including investigation of different epigenetic mechanisms such as mIRNA, histone modification and crosstalk between these epigenetic mechanisms with DNA methylation provide further insight into the role of epigenetic regulation in insects polyphenism. Lastly, as discussed above due to the limitation of the MBD-seq, future research on capturing these lowly methylated genes

might help to further understand the role of DNA methylation in regulating aphid polyphenism.

Chapter 6: miRNA-profile of dispersal and non-dispersal morphs in two aphid genotypes

6.1 Introduction

The non-coding region in genomes has always been considered junk DNA and speculated to be inconsequential in regulating gene expression (Asgari 2013). However, with the advancement in scientific technology and molecular approaches, many studies have started reporting the significant role of this non-coding region in regulating gene expression. miRNA is defined as small non-coding RNA, usually around 22 nucleotides long. The presence of miRNAs was first reported in the roundworm *Caenorhabditis elegans* (Lee et al. 1993; Wightman et al. 1993). To date, microRNAs have been reported across both vertebrates, invertebrates, plants and viruses (Kozomara and Griffiths-Jones 2011). The discovery of miRNA has added another layer of complexity to the role of post-transcriptional modification in regulating gene expression. miRNA has been reported to be essential in ensuring optimise gene expression level and target tuning (Bartel and Chen 2004).

miRNA can be encoded by different regions including introns, non-coding transcripts and coding regions. In insects, RNA polymerase II will first transcribe the miRNA loci, which results in the formation of the primary miRNA transcript (pri-miRNA) (Lee et al. 2004). Next, the primary miRNA transcript will undergo a series of processing and sorting events before becoming mature miRNA and acting on target genes. miRNA has been reported to regulate multiple aspects of insect development ranging from growth and development, embryogenesis, and sex determination to behaviour (Lee et al. 2015; Yang et al. 2014; Kugler et al. 2013; Fagegaltier et al. 2014), see Chapter 1 (Section 1.3) for more details.

Recently, miRNA has been reported to play an important role in regulating the development of polyphenism in aphids and other insects (Fan et al. 2020; Xu et al. 2020). In Chapter 5, we investigated the role of DNA methylation in regulating polyphenism and found no significant correlation between DNA methylation and polyphenism. Therefore, in this chapter, we investigate the miRNA profile in different pea aphid morphs and genotypes. The overall aim of this chapter was to investigate the potential role of miRNA in regulating aphid polyphenisms. Recent work has also focused on understanding the difference in

aphid genotypes and the differences in the degree of phenotypic plasticity between aphid genotypes (Kanvil et al. 2014; Parker et al. 2019; Sentis et al. 2019). Because of the critical role of the ability to disperse in aphid evolutionary success, this chapter also aimed to assess the miRNA profile between the aphid morphs and genotype to provide insight into different and novel miRNAs that could regulate the difference in life-history traits between the morphs and genotype. Specifically, this Chapter aimed to:

- Investigate the miRNA profile of different pea aphid morphs in two genotypes (N116 winged vs N116 wingless and N127 red vs N116 wingless)
- Determine the possibility of novel miRNA in pea aphids that could play an important role in pea aphid polyphenism
- Integrate the results of transcriptome profile (Chapter 4) and miRNA to provide insight into the role of miRNA in polyphenism.

6.2 Methods

6.2.1. Materials

All samples used in the experiment were obtained from the mesocosm experiment described in section 3.3.3. The number of samples used in the miRNA analysis is outlined in (Table 6.1), stratified by genotype and morphs.

Table 6.1. Summary of morph numbers used for molecular analysis in Chapter 6.

Assay	Genotype	Morphs	Replicates
miRNA-seq	N116	wingless	x3(pooled samples)
	N116	winged	x3 (pooled samples)
	N127	red	x3 (pooled samples)

6.2.2 miRNA library preparation

Total RNA was extracted from whole adult aphids obtained from Chapter 2 (section 2.4.2) and was submitted to RealSeqbiosciences (USA) for miRNA library preparation and sequencing. The quality and integrity of the RNA samples were assessed using a 2200 TapeStation (Agilent Technologies) and only RNA with an Integrity Number (RIN) >7 will be used for the library preparation and sequencing using the RealSeq®-AC library kit (Realseqbiosciences, USA). Briefly, total RNA (100ng) was used as the input material. First, the total RNA sample will undergo the adapter ligation process. Total RNA will be mixed with the Realseq adapter and RNA buffer followed by heating for 2min at 70°C and transferred immediately to ice for another 2min. Then, samples were mixed with ligation buffer followed by incubating at 25°C for 60min and 65°C for 5min. After the adapter ligation, the samples proceeded to adapter blocking steps immediately. Briefly, blocking agents were added to the sample followed by incubating at 65°C for 5min and a step down from 65°C to 37°C at the rate of 0.1°C per second. Next, samples were added with blocking enzymes and buffer followed by incubation at 37°C for 60min and 65°C for 20min. After ligation blocking, samples proceeded to circularization steps immediately. Realseq enzyme

and buffer were added to the samples followed by incubation at 37°C for 60min. After ciruclarization, samples proceeded to the dimer removal steps. Dimer removal agents (DRA) were added to the samples and incubated at 37°C for 10min followed by resuspending the samples with the SPRIselect beads and incubation for a further 10min at 37°C. Then, the samples with beads were transferred to a magnetic rack and the supernatant was transferred to a new tube for reverse transcription. The samples were added with RT primers and dNTPs followed by incubation at 65°C for 5min and transferred to ice for another 2min. Then, RT enzymes were added to the samples with Rnase inhibitor and buffer solution followed by incubation at 42°C for 60min and 65°C for 20min. Then, the samples were mixed with PCR master mix and reverse primer index followed by PCR at the following conditions: 94°C at 30sec, 15 cycles of cDNA denaturation at 94°C for 15sec, followed by primers annealing at 62°C for 30s and lastly extension at 70°C for 15s and holding at 70^{0}C for 5min. Lastly, the samples underwent size selection and were quantified with Tapestation 4200 (Agilent, USA) and Qubit ((Thermo Fisher, UK) and sequenced on the Illumina NextSeq 500v2 (1 x 75 cycles).

6.2.3 Sequence assembly and miRNA annotation

The raw fastq files were processed using Cutadapt to remove adapter sequences and any reads that were shorter than 5bp to determine degraded RNAs. After trimming, the reads with a minimum length of 15bp were aligned to the miRNA sequence of pea aphid from the miRbase (miRNA database) and novel miRNAs from the miRDeep2 (Friedländer et al. 2012). miRNAbase contains all the known miRNA sequences and annotations with information on the location and sequence of the mature miRNA. Next, the differential expression miRNA was determined using the DeSeq2 software using the default setting (PMID:25516281). Only miRNA with a p-value adjusted of <0.05 is considered to have a significant change of expression. Then, the differentially expressed miRNAs between groups were analysed for target gene prediction using the RNAhybrid software on the

BiBiServ2 using the default settings. RNAhybrid is a tool that is usually used to calculate the minimum free energy hybridization between the long and short RNA. (Rehmsmeier et al. 2004). Blast2GO software (http://www.geneontology.org) was used for gene ontology (GO) annotations. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was used to identify significantly enriched metabolic pathways or signal transduction pathways in DEGs based on the database.

6.3 Results

6.3.1 miRNA-seq analysis

Three different morphs from two genotypes were selected for miRNA profiling: genotype N116 (wingless and winged) and genotype N127 (red); the pale morph was not included in the analysis due to lack of resources (budget). The two genotypes have very different body morph colours: N116 is usually green, while N127 is usually red. Apart from that, they respond differently to crowded conditions, with N116 producing winged offspring and N127 changing their body colour from red to pale (see also Chapter 2 Section 2.1).

6.3.2 Differentially expressed genes, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis between genotype and morphs

The miRNAs of three different pea aphid morphs were compared based on the gene expression results. We identified many known miRNAs and novel miRNAs that could potentially be important in regulating aphid polyphenism. 13.6 million reads were obtained for N116 (winged), 15.8 million reads for N116 (wingless), and 10.5 million reads for N127 (red) (Table 6.2). After filtering, 13 million, 15 million, and 9.4 million high-quality reads were obtained for each morph, respectively. Interestingly, only a low number of reads were uniquely mapped to the database with most of the remaining reads being tRNA and rRNA (Figure 6.1). Our microRNA-seq reveals 13 differentially expressed miRNA between N116 winged vs N116 wingless and 19 differentially expressed mIRNA between N127 pale vs N127 red (Figure 6.2). Then, there are 3 overlap miRNA there wree found between the N116winged vs N116wingless group vs N127red vs n116wingless group.

Table 6.2. Number of raw reads and uniquely mapped reads obtained from miRNA seq.

Strain	Number of input reads	Uniquely mapped reads
N116 winged	78,611,329	57,104
N116 wingless	80,283,756	80,667
N127 red	82,289,189	169,824

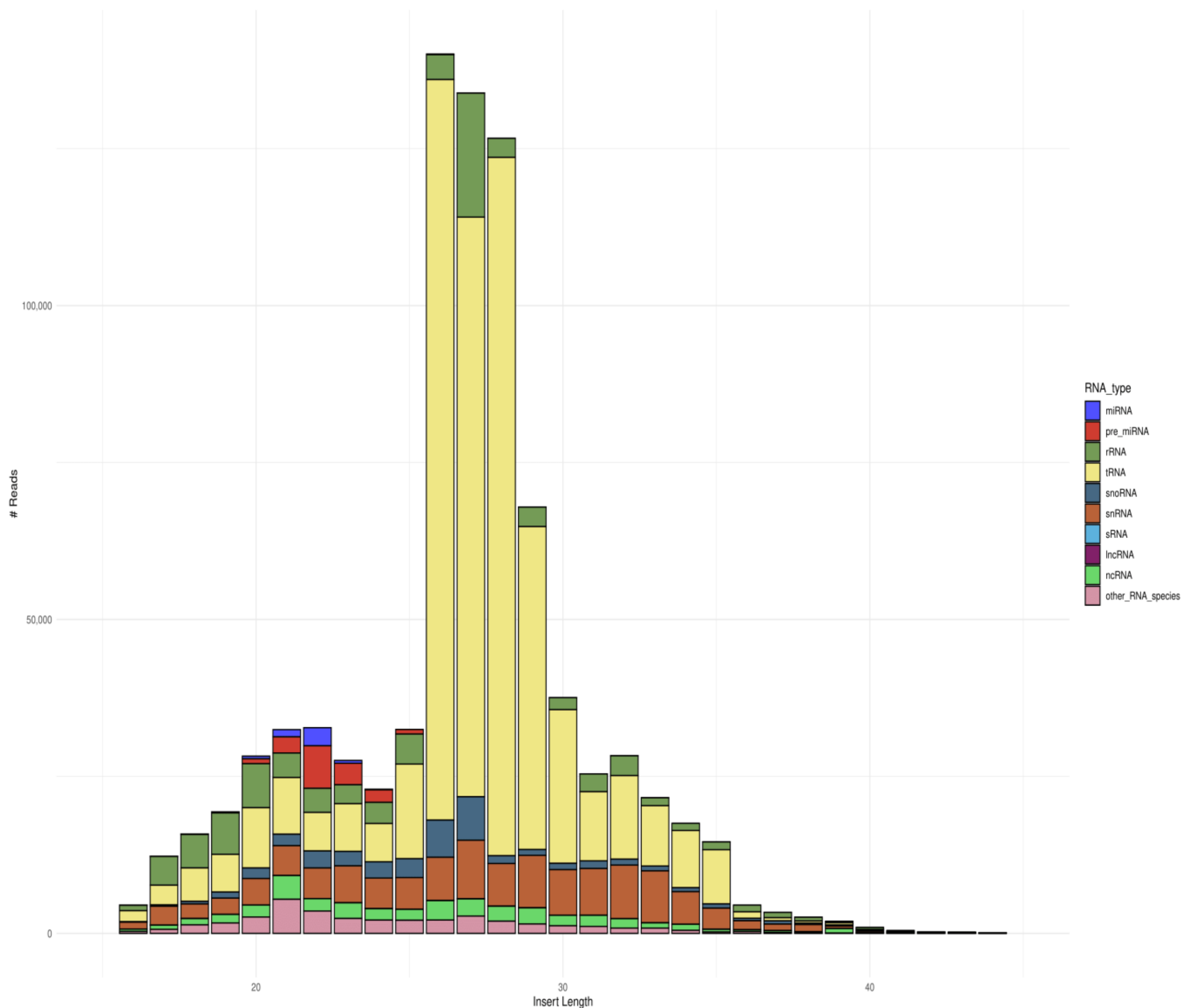


Figure 6.1. Distribution of different type of RNA in the microRNA-seq. miRNA are usually between 18-25 in nucleotide length with and average of 22 nucleotides. The figures shows that between these range most of the read were composed of tRNAs and rRNA.

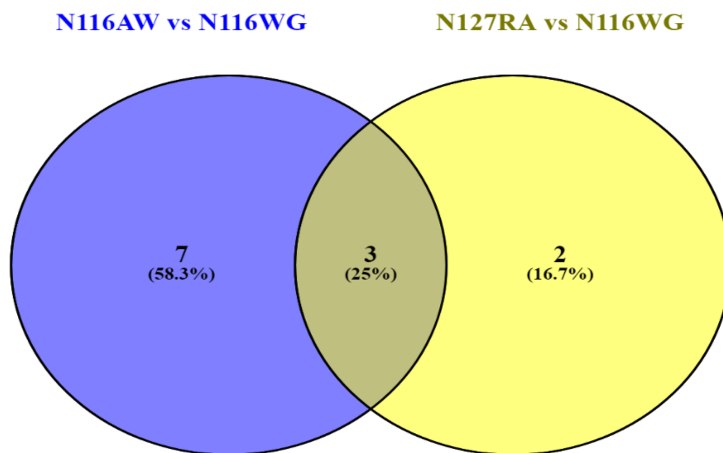


Figure 6.2. Venn diagram comparing number of differentially expressed miRNA identified by microRNA-seq between different aphid morphs. Blue=N116 winged +N116wingless morphs, Yellow=N127red +N116winglesss**.** Each sub-category represents the overlapping genes that were found between the two or more different group comparisons.

For miRNAs expression measurements, a range of 0.5-2% of the total reads was uniquely mapped to the genes in the reference database. The significant differentially expressed miRNAs for all groups were then further annotated for GO by using the Blast2Go parameters. The functions of the DEG by Blast2Go (Fisher's exact test, FDR<0.05) were then classified using GO assignments. For further functional categorisation, KEGG pathway analysis was performed using the pea aphid KEGG database. The DEG was then classified into different KEGG Ontology (KO) terms.

a) Differentially expressed miRNAs, GO and KEGG between N116 winged and N116 wingless aphids

We identified 13 different miRNA that were significantly differentially expressed (p<0.05, -0.3<log2fold>0.3) between the N116 winged and N116 wingless. Ten of the miRNAs were identified from the known miRNA database (miRbase) while the three remaining miRNAs were novel miRNAs detected from miRDeep2 (Table 6.3). The differentially expressed miRNA were subjected to RNAhybrid for target gene predictions. 1131 target genes were found to be regulated by four different miRNAs with 154 genes in miR-184a, 449 genes in miR-2765, 459 genes in miR-252a and 69 genes in miR-100. (Figure 6.2). Interestingly, some of the genes that were predicted to be regulated by miRNA-100 were involved in wing development, metabolite regulation and cuticular formation such as apterous(ap1), protein Wnt2(Wnt2), Trehalase (Treh), RR1 cuticle (RR1), glucose dehydrogenase (Gld). Some genes that were regulated by the miR-184a were involved in wing development, methylation process and stress response such as daschous (ds), N6-adenosine-methyltransferase catalytic subunit (METTL3), Notch (N), heat shock protein 70(Hsp70). Then, genes that were regulated by miR-252a were involved in histone modification, stress response, and carotene production such as histone deacetylase rpd3-like (Rpd3), histone deacetylase complex subunit SAP18 (SAP18), heat shock protein 68-like (Hsp68), phytoene synthase (crtYB). Lastly, some of the genes that were regulated by the mIR-2765 were involved in wing development such as bursicon (burs), juvenile hormone binding protein-like (JHBP), protein decapentaplegic (Dpp), protein distaless (Dll), protein daschous(ds), ecdysone receptor (Ecr).

In addition to the target gene predictions, the predicted target genes regulated by different miRNAs were also subjected to gene ontology (GO) analysis. Significant enrichment was obtained for 45 genes that fall under three functional groups. The 45 genes were classified into the molecular function (MF) categories: Of the genes in the cellular components category, 42% were involved in RNA polymerase II transcription factor activity, 42% in calcium ion binding and 16% in fatty-acyl-CoA reductase (alcohol-forming) activity (Figure 6.4).

Next, the list of the targeted gene was also subjected to KEGG analysis, but no significant enrichments were obtained.

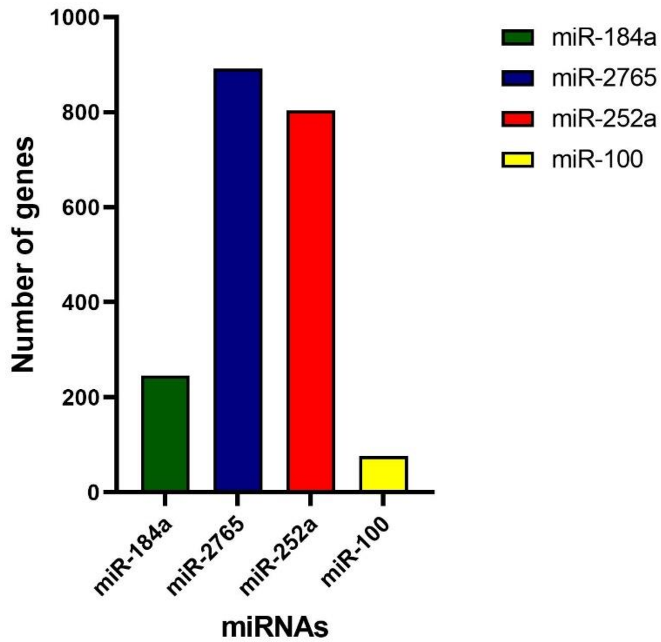


Figure 6.3. Number of genes regulated by each miRNA. Green=miR-184a, blue=miR-2765, red=miR-252a, yellow=miR-100. The number of genes in the graphs represent individual unique genes with duplicated genes removed.

Table 6.3. Differentially expressed conserved and novel miRNA in N116 winged vs wingless aphids . miR= miRNA that were obtained from the miRNA database (miRbase). NC/NW= novel miRNA from the study and not reported in the database yet.

miRNA	**Log2fold**	**p-adj**
api-miR-277	2.29	<0.001
api-miR-100	2.05	<0.001
NC_042494.1_chromosome_A1_13821	1.40	<0.001
NW_021769689.1_unplaced_genomic_scaffold_39370	1.35	<0.001
api-miR-14	-0.92	<0.001
api-let-7	1.23	<0.001
NC_042493.1_chromosome_X_3747	1.09	<0.001
api-miR252a	1.40	<0.001
api-miR184b	1.04	0.018
api-miR-184a	1.04	0.018
api-miR-8	-0.76	0.033
api-miR-3027	-1.43	0.033
api-miR-2075	-0.58	0.039

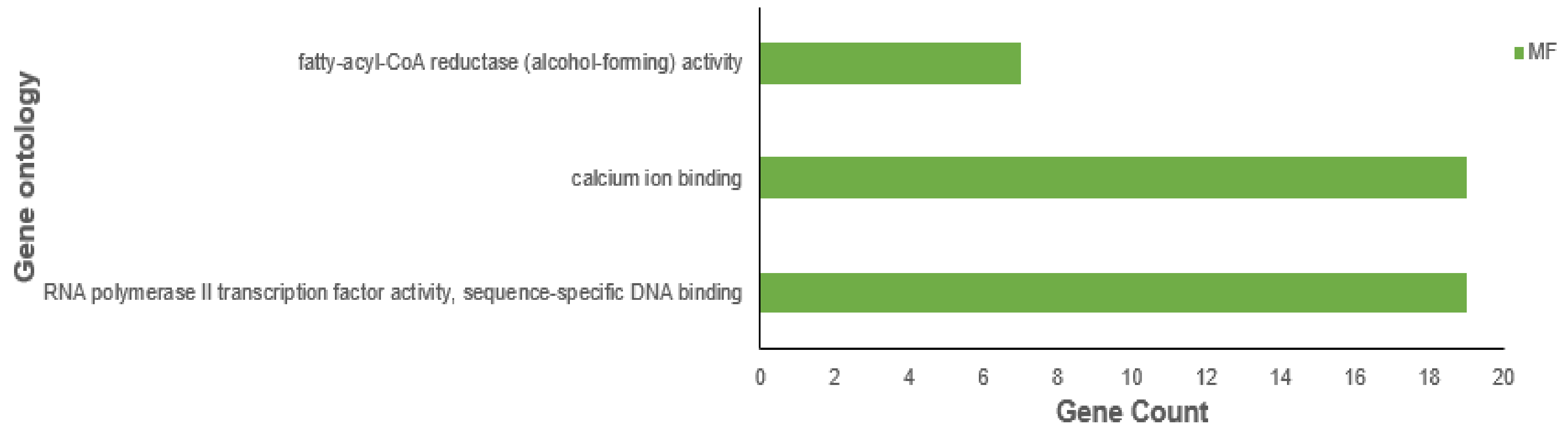


Figure 6.4. GO enrichment analysis between N116 winged and N116 wingless aphids. All of the gene ontology falls under the category molecular function (MF)

b) Differentially expressed miRNAs, GO and KEGG between N127 red and N116 green wingless aphid

We identified 19 different miRNAs that were significantly differentially expressed ($p<0.05$, $-0.3<\log2fold>0.3$) between the N127 red aphid and N116 wingless aphids. Six of the miRNAs were identified from the known miRNA database (miRbase) while the 13 remaining miRNAs were novel miRNAs detected from miRDeep2 (Table 6.3). Next, the differentially miRNA were subjected to RNAhybrid for target gene predictions. 1495 target genes were found to be regulated by four different miRNAs with 103 genes by miR-278, and 1392 genes by NW_021769971.1_unplaced_genomic_scaffold_39500(Figure 6.5). Interestingly, some of the genes that were predicted to be regulated by miRNA-278 were involved in eye development and external sensory organs development, actin filament formation, oocyte formation and neuronal development such as Homeobox protein BH-1 (B-H1), ataxin 2-homolog (Atx2), titin (sls), protein sax-3 (sax-3), protein scabrous (sca). The novel miRNA NW_021769971.1_unplaced_genomic_scaffold_39500 was involved in the regulation of many genes including those involved in wing development, hormone regulation, ecdysis process, sugar detection, fecundity, metabolite regulation, carotene regulation, stress response and histone modification such as Krueppel homolog 1(Kr-h1), homeotic protein distal-less(Dll), gustatory receptor for sugar taste-64f-like (Gr64f), apterous (ap1), ecdysone-induced protein 75 (Eip75), homeotic protein ultrabithorax(Ubx), ecdysone receptor (EcR), homeotic protein homothorax (hth), protein Gustavus (gus), protein Wnt-1 (wg). G-protein coupled receptor Mth2 (mth2), neurogenic locus Notch protein (N), protein Wnt-2 (Wnt-2), phytoene desaturase (PDS), histone acetyltransferase p300(P300), mothers against decapentaplegic (Mad), trehalase (Treh), heat shock protein 70kDa protein 14 (Hsp70), protein dacshous(ds), protein decapentaplegic (dpp), histone deacetylase Rpd3(Rpd3), homeobox protein engrailed-2 (en), heat shock protein 68-like (Hsp68).

In addition to the target gene predictions, the predicted target genes regulated by different miRNAs were also subjected to gene ontology (GO) analysis. Significant enrichment was

obtained for 206 genes that fall under nine functional groups (Figure 6.6). The 206 genes were classified into three categories: 51 genes in biological process (BP), 133 in molecular function (MF) and 22 in cellular components (CC). Of the genes in the biological processes category, 45% were involved in signal transduction, 39% in the regulation of transcription, and the remaining in cell surface receptor signalling pathway. In the molecular function category, 26% of genes were involved in the transmembrane transporter activity, 24% RNA polymerase II transcription factor activity, sequence-specific DNA binding, and the remaining in transcription factor activity, sequence-specific DNA binding sequence-specific DNA binding and G-protein coupled receptor activity. Lastly, all genes in cellular components were involved in extracellular activity (Figure 6.6).

Next, the target genes were also subjected to the Kyoto Encyclopedia of Genes and Genome (KEGG) analysis. Significant enrichment was obtained for 29 genes that fall under two different KEGG function descriptions, namely MAPK signalling pathway-fly, and the TGF-beta signalling pathway (Figure 6.7).

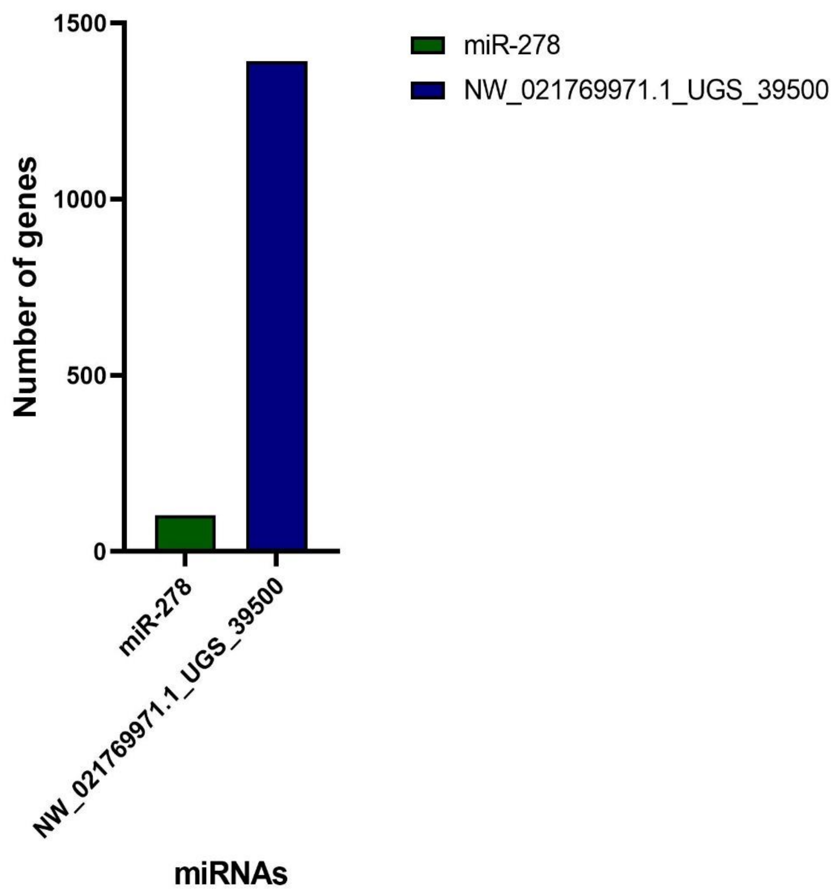


Figure 6.5. Number of genes regulated by each miRNA. Green=miR-278, blue= NW_021769971.1_unplaced_genomic_scaffold_39500. The number of genes in the graphs represent each individual unique gene with duplicated genes removed

Table 6.4. Differentially expressed conserved and novel miRNA in N127 red and N116 green wingless aphids. miR= miRNA that were obtained from the miRNA database (miRbase). NC/NW= novel miRNA from the study and not reported in the database yet.

miRNA	**Log2fold**	**p-adj**
NC_042494.1_chromosome_A1_13821	-2.74	<0.001
NW_021769971.1_unplaced_genomic_scaffold_39500	-3.25	<0.001
NC_042493.1_chromosome_X_3747	-2.65	<0.001
NC_042494.1_chromosome_A1_8551	-2.96	<0.001
NC_042493.1_chromosome_X_4914	-3.04	<0.001
api-miR-100	1.98	<0.001
api-miR-278	1.26	<0.001
NC_042493.1_chromosome_X_19	-1.35	<0.001
NC_042494.1_chromosome_A1_16979	2.00	<0.001
api-let-7	1.37	0.001
NW_021769689.1_unplaced_genomic_scaffold_39370	-0.88	0.018
NC_042494.1_chromosome_A1_18069	0.57	0.039
api-miR-14	-0.56	0.039
NC_042495.1_chromosome_A2_29278	-0.64	0.039

api-miR-993	-0.74	0.039
api-miR-2765	0.52	0.044
NC_042493.1_chromosome_X_4087	3.85	0.046
NC_042493.1_chromosome_X_4091	3.85	0.046
NC_042493.1_chromosome_X_4081	3.85	0.046

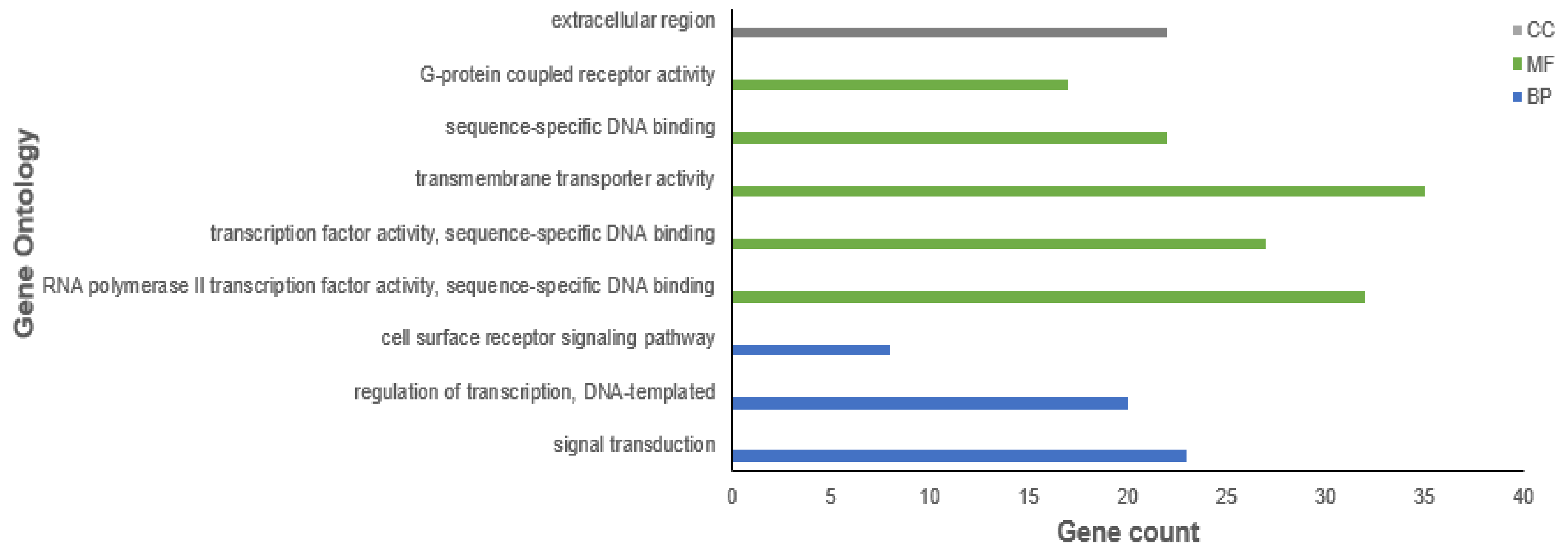


Figure 6.6. GO enrichment analysis between N127 red and N116 green wingless aphids. The gene ontology was classified into three main categories: BP=Biological process (blue), MF=Molecular Function (green), CC=Cellular Components (grey).

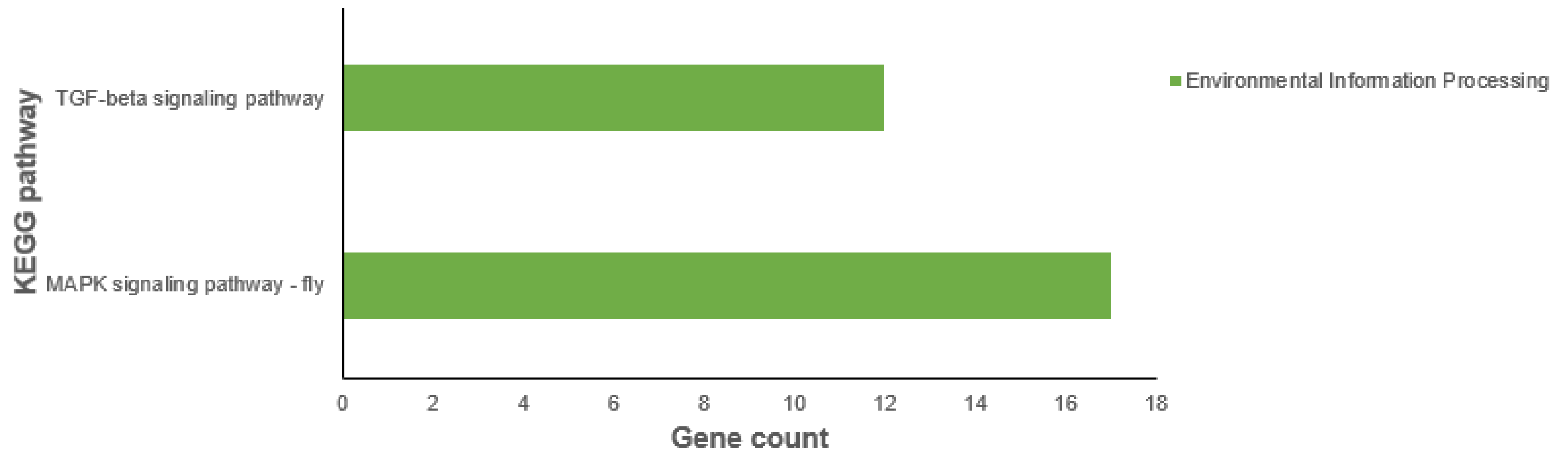


Figure 6.7. KEGG classification between N127 red and N116 green wingless aphids. X-axis represents the number of unigenes in the corresponding functional class. The genes were classified into three different wider categories, green=environmental information processing.

6.3.3 Integration of miRNA-seq and RNA-seq

The transcriptome profile of the four different pea aphid morphs (Chapter 4) was integrated with the miRNA target gene prediction profile to provide insight into the relationship between miRNA and gene expression.

a) miRNA targets genes that overlap with RNA-seq (differentially expressed genes) in N116 winged and N116 wingless aphids.

Based on the miRNA target gene predictions obtained from RNAhybrid, we found a total of 145 genes that overlap with RNA-seq data across four different miRNAs. Of the 69 target genes regulated by miRNA-100, nine genes overlap with the RNA-seq data with seven genes downregulated and two upregulated. Of the 154 genes regulated by miRNA-184a, 24 genes overlap with RNA-seq data with 20genes downregulated and four upregulated. Further, of the 459 genes regulated by miRNA-252a, 61 genes overlap with the RNA-seq data with 44 genes downregulated and 17 genes upregulated. Finally, of the 449 genes regulated by miRNA-2765, 66 genes overlap with the RNA-seq data with 51 genes downregulated and 15 genes upregulated (Figure 6.8A-B).

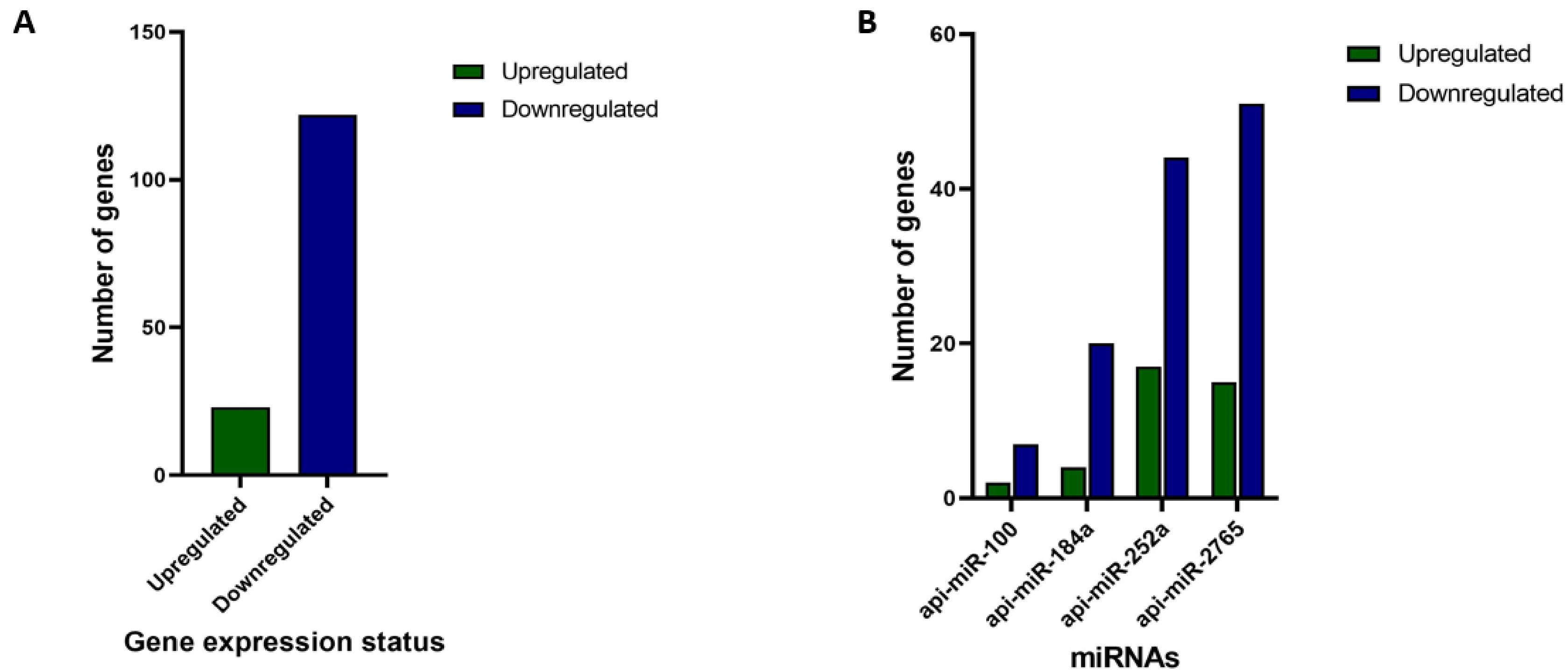


Figure 6.8. Integration of miRNA with RNA-seq data between N116 winged and N116 wingless aphids. A) Total number of genes that overlapped between miRNA target gene prediction and RNA-seq, Green=upregulated (increase expression in N116 winged), blue=downregulated (decrease expression in N116 wingless)**. B)** Total number of genes that overlapped between individual miRNA target gene prediction and RNA-seq. Green=upregulated (increase expression in N116 winged), blue=downregulated (decrease expression in N116 winged). miRNA can both upregulate and downregulate the gene expression depending on the region they bind to (e.g miRNA bind in promoter region can upregulate expression, miRNA binding in 5'UTR and coding region and repress gene expression).

b) miRNA targets genes that overlap with RNA-seq (differentially expressed genes) in N127 red and N116 green wingless aphids.

Based on the miRNA target gene predictions obtained from RNAhybrid, we found a total of 145 genes that overlap with RNA-seq data across four different miRNAs. Of the 103 target genes regulated by miRNA-278, 26 genes overlap with the RNA-seq data with 20 genes downregulated and six upregulated. Of the 1392 genes regulated by NW_021769971.1_unplaced_genomic_scaffold_39500, 478 genes overlap with RNA-seq data with 378 genes downregulated and 100 upregulated (Figure 6. 9A-B).

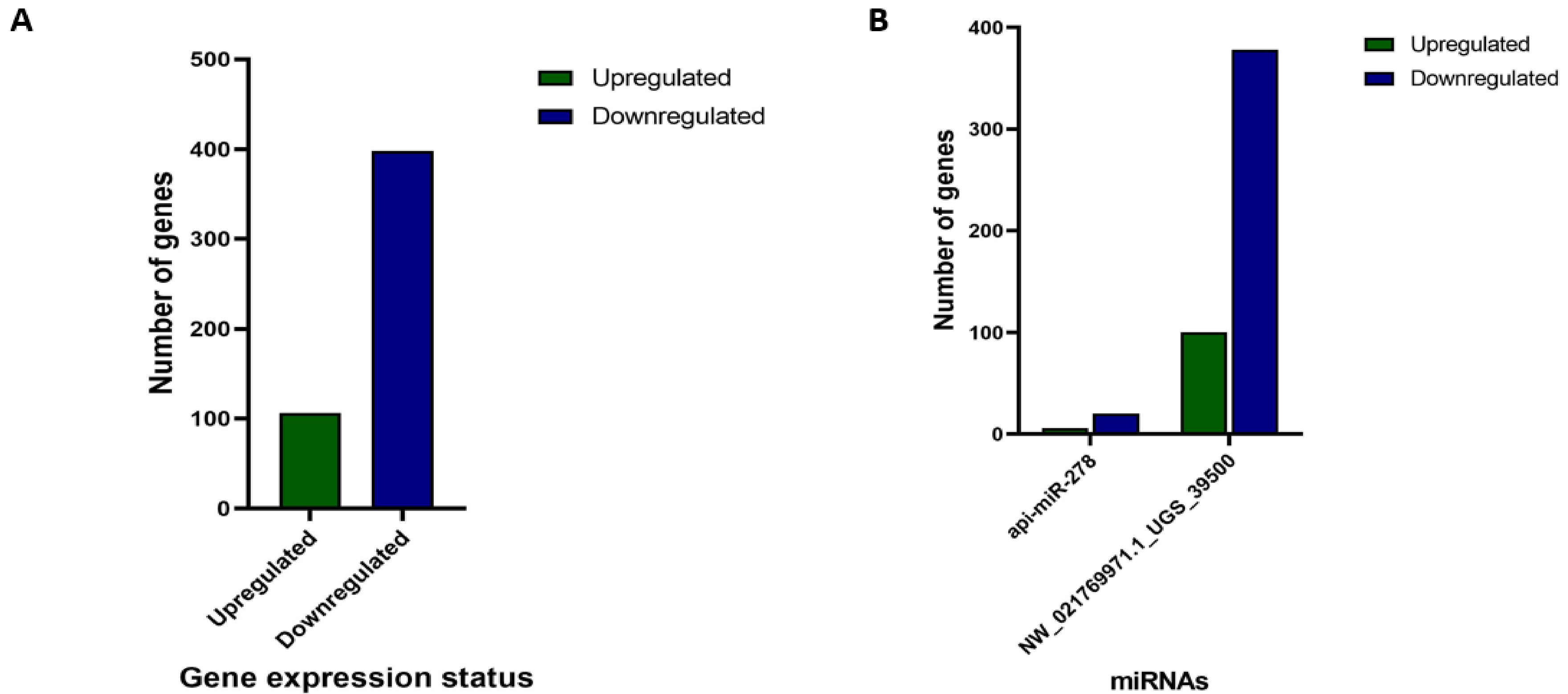


Figure 6.9. Integration of miRNA with RNA-seq data between N127 red and N116 green wingless aphids. A) Total number of genes that overlapped between miRNA target gene prediction and RNA-seq, Green=upregulated (increase expression in N127 red aphid), blue=downregulated (decrease expression in N127 red aphid). **B)** Total number of genes that overlapped between individual miRNA target gene prediction and RNA-seq. Green=upregulated (increase expression in N116 winged), blue=downregulated (decrease expression in N116 winged). miRNA can both upregulate and downregulate the gene expression depending on the region they bind to (e.g miRNA bind in promoter region can upregulate expression, miRNA binding in 5’UTR and coding region and repress gene expression).

6.4 Discussion

Wing polyphenism plays an essential role in the evolutionary success of many different insect species. In aphids, wing polyphenism seems particularly important as they allow aphids to quickly escape from stressful environmental conditions such as increased population density, predators and poor host-plant quality. Since aphids usually reproduce asexually, all offspring are clones of the mother which makes them an excellent model for studying the role of epigenetic regulating in wing polyphenism. The most widely studied epigenetic mechanism in regulating insect polyphenism is DNA methylation, which has been discussed above (Chapter 5).

However, the role of other epigenetic mechanisms in regulating wing polyphenism has been understudied, especially the role of miRNAs in insect polyphenism. By sequencing small RNA libraries, we identified 13 miRNAs including ten conserved and three novel miRNAs that were significantly differentially expressed between N116 winged morphs and N116 wingless morphs. Most miRNA are usually found between the range of (18-25) nucleotide length. Further, we found that most miRNAs were represented by the 22nt reads which is the common average size after Dicer digestion as observed in *Drosophila melanogaster* and *Spodoptera litura* (Lai et al. 2003, Ge et al. 2013). Interestingly, our sequencing results showed that the majority of the libraries were composed of tRNA across all the different nucleotides length (18-30). Next, Sattar et al. (2012) reported similar observations in the small RNA libraries for the greenfly *Aphis gossypii* where most of the reads between 18-30 were also composed of tRNA.

N116 winged vs N116 wingless

Among the differentially expressed miRNAs, miR-277 and miR-100 showed the highest fold change between N116 winged morphs and N116 wingless (fold change>2) with both miR being upregulated in N116 winged morphs. Interestingly, miR-277 has been shown to play an essential role in regulating insect lifespans. For example, Esslinger et al. (2013) showed

that *Drosophila* with higher increased expression of miR-277 showed a shorter lifespan and is usually accompanied by a reduction in insulin signalling. Our results showed that the miR-277 was upregulated in the N116 winged morphs. This observation suggests that miR-277 might play a vital role in regulating lifespan in different pea aphid morphs as the winged aphids usually showed a longer lifespan in comparison to the wingless aphid. Next, a similar observation is also reported by Li et al. (2016) in the English grain aphid *Sitobian avenae* whereby the winged morphs showed higher expression for this miRNA in comparison to the wingless morph. Apart from their role in lifespan, miRNA-277 has also been reported to play a role in lipid storage and ovarian development. For example, Ling et al. (2017) showed that depletion of miR-277 in the mosquito *Aedes aegypti* results in impairment of lipid storage and ovarian development. In aphids, the winged aphids usually showed higher lipid storage and lower reproduction in comparison to the wingless morphs, therefore miR-277 plays a key role in regulating the lipid regulation and reproduction in different aphid morphs. Next, miRNA-277 is also reported to play a key role in wing vein development. For example, Shen et al. (2020) showed that the cotton ballworm *Helicoverpa armigera* with depleted miR-277 showed a defect in wing veins development. Our results reported a higher level of miR-277 in the winged morphs, which suggests that miR-277 might play an essential role in aphid wing veins formation.

The miR-100 and let-7 are also highly expressed in the N116 winged morphs in comparison to N116 wingless morphs. miR-100 has been reported to play a significant role in wing morphogenesis in insects. For example, Rubio et al. (2013) reported that depletion of miR-100 and let-7 results in a reduction in the wing size of the german cockroach *Blattela germanica*. Soares et al. (2021) reported an elevated level of miR-100 and let-7 at the wing disk of honeybee workers and suggested that miR-100 could potentially regulate the wing formation in the honeybee. In *Drosophila*, let7, miR-100 and miR-125 are usually expressed together as a single transcript and have been reported to play an essential role in regulating *Drosophila* wing development (Caygill et al. 2008). Furthermore, let-7 has been reported to

regulate larval pupal in insects by targeting the ecdysone pathway. For example, Peng et al. (2019) reported that silencing the let-7 led to a downregulation of the expression of ecdysone in the oriental fruit fly *Bactrocera dorsalis*. Song et al. 2018 also reported that let-7 -plays a role in regulating the expression of Kruppel-homolog 1. Ecdysone has been reported to play an essential role in transgenerational wing polyphenism in pea aphids. For example, Vellichirammal et al. (2017) reported that an increase in ecdysone signalling usually results in more wingless progeny. In contrast, Krueppel homolog 1 has been reported to inhibit the synthesis of ecdysone (Zhang et al. 2018). In our results, we found that let-7 was highly expressed in the N116 winged morph in comparison to the N116 wingless morphs. Therefore, let-7 may regulate the wing polyphenism in aphids by regulating the expression of the ecdysone pathway gene and Krueppel homolog-1.

Next, our result found that miRNA-252a is also upregulated in the N116 winged morphs. miR252a has been reported to play a role in regulating the shell color of the pacific oyster *Crassostrea gigas* (Li et al. 2021). Apart from that, mIR-252a has been reported to play a role in seasonal polyphenism. For example, Mukherjee et al. (2020) reported that miR-252a was differentially expressed between larvae and pupae in the European map butterfly *Araschinia levana* that were raised under different daylight regimes. Together these results suggest that miR-252a could potentially play a role in regulating wing polyphenism and body colour changes in aphids and further research is needed to elucidate this. Next, miR-184a and miR-184b is also upregulated in the N116 winged morphs. miR184 has been reported to play a role in wing polyphenism and survival in aphids. For example, Li et al. (2022) reported that miR-184a were differentially expressed between different developmental stages of winged and wingless destined English grain aphid nymph with no differences between the adult stage. Our results disagree with this study as we found that miR-184a was differentially expressed between N116 winged and wingless morphs in the adult stage and future research is needed to elucidate the role of this miRNA in regulating wing polyphenism and survival in pea aphids. Furthermore, miR-184a and miR184b has

also been reported to play a role in immune response in the pea aphid. For example, Ma et al. (2020) reported the downregulation of both miRNAs upon infection by *P.aeruginosa* and *M.luteus*. Ma proposes that both miRNA negatively downregulate the JNK signalling pathway, which is involved in a major event such as phagocytosis and PPO activation. Although in our study, no infection induction was involved, the winged morphs usually have lower immune system function than the wingless morphs and these two miRNAs could play a role in regulating the immune response observed between the morphs.

Our results reveal that the miR-3027 was downregulated in the N116 winged morphs in comparison to N116 wingless morphs. This miRNA has only been reported only in aphids and it is likely that this miRNA is Hemiptera-specific (Legeai et al. 2010). However, the role of miRNA in aphids remains unknown and has not been studied till the present. Then, our results also reveal that miR-14 was downregulated in the N116 winged morphs in comparison to the wingless morphs. miR14 has been reported to play a role in regulating the ecdysone level in insects. For example, Liu et al. (2018) reported that miR-14 regulates the ecdysone titre in the silkworm *Bombyx mori* with overexpression and depletion resulting in delayed development and precocious wandering stage respectively. Next, similar observations were also reported in the Asiatic rice borer *Chilo suppressalis* (He et al. 2019). As discussed above, ecdysone plays a key role in aphid wing polyphenism and it is possible that this miRNA together with the other miRNA discussed above (miR-let7 and miR-100) helps regulate wing polyphenism in aphids by modulating the ecdysone titre. Furthermore, Chen et al. (2021) also reported that miR-14 plays a role in regulating the reproduction of honeybees by targeting the ecdysone pathway. Apart from that, some study has shown that ecdysone plays a role in the trade-off. For example, Wang et al. (2022) reported that ecdysone plays a crucial role in mediating the trade-off between reproduction and immunity in mosquito *Aedes aegypti*. Since winged aphids usually have lower reproduction than wingless aphids, it is possible that this trade-off might be regulated by the miR through regulating the ecdysone signalling.

Our results have shown that the miR-8 is downregulated in the N116 winged morphs. miR-8 has been reported to play a role in regulating the juvenile hormone regulation in insects. For example, Zhang et al. (2021) showed that the reduction of miR-8 in *Drosophila* leads to a decrease in the juvenile hormone gene expression. Further, Lucas et al. (2015) also reported similar observations in the mosquito *Aedes aegypti* whereby reduction of miR-8 expression leads to a decrease in reproduction through targeting the wingless signalling pathway. In aphids, the juvenile hormone has been reported to regulate fecundity and adult size weight (Gao et al. 2018). The winged aphid usually has lower reproduction compared to the wingless morphs; therefore, it is possible that miR-8 might play a role in regulating these fecundity differences between the morphs. Then, miR-8 has also been reported to play other functions in insect development, such as regulating pigmentation and eclosion (Kennell et al. 2012).

Then, our results also showed that the miR-2765 were downregulated in the N116 winged morphs in comparison to the wingless morph. The role of miR-2765 in insects remains widely unknown. However, some research has reported that this miRNA was differentially expressed between the female and males across a few different insect species (Jain et al. 2015; Li et al. 2021).

Our target gene prediction results showed that among the ten conserved and differentially expressed miRNAs, four miRNAs (miR-100, miR184, miR252a, miR2765) showed significant results. We found that many of the genes that were targeted by these miRNAs were involved in many different aspects of aphid development. For example, Wnt-2, Dll, ap1, Dpp, ds, and burs are known to play an essential role in many roles in insects wing development (Brisson et al. 2010; Elias-Neto et al. 2022; Zhang et al. 2022; Long et al. 2022). Further, some of the target genes were also involved in histone modification such as Rpd3. Sap18. These genes have been reported to play an important role in regulating starvation resistance and reproduction in *Drosophila* and pea aphids respectively (Nakajima

et al. 2016; Kirfel et al. 2019). Some of the target genes were also involved in stress response in insects such as Hsp68, and Hsp70 (Kausar et al. 2020; Beasley-Hall et al. 2022). Finally, many of the target genes were also involved in other aphid development processes such as metabolite regulation, carotene production and others. Together the results suggest that miRNA might be an important epigenetic regulation in modulating gene expression in aphids.

N127 red vs N116 green wingless

Among the differentially expressed miRNAs between N127 red and N116 wingless, miR-100 and let-7 had the highest fold change between N127 red aphid and N116 wingless. Interestingly, both miRNAs were upregulated in the N127 red morph in comparison to the N116 wingless. As discussed above, let-7 and miR-100 play an essential role in wing development. However, the N127 red aphid usually does not produce winged offspring but changes their body colour when exposed to environmental stress, therefore it is surprising to see that both miRNAs were upregulated in N127 rather than N116 which usually produces winged offspring under stress. However, it is also possible that this miRNA has other functions in aphids that have not been discovered yet. Next, we found that miRNA-278 was also upregulated in N127 red morphs compared to N116 wingless morphs. miRNA-278 has been reported to work together with let-7 in regulating metamorphosis in insects. For example, Song et al. (2018) reported that let-7 and miRNA-278 work together to suppress the expression of Kr-h1 which leads to precocious metamorphosis in locusts *L.migratoria*. As mentioned above, Kr-h1 is usually involved in suppressing ecdysone synthesis. Since miR-278 and let-7 reduce the Kr-h1 expression, this could result in higher ecdysone titre which usually results in more wingless progeny (Vellichirammal et al. 2017). Further, miR-278 has been reported to play a role in metabolite regulation. For example, Teleman et al. (2006) reported that *Drosophila* lacking miR-278 show a defect in energy homeostasis and result in an insulin resistance phenotype (increase insulin production but

reduced sensitivity). Insulin plays an important role in regulating insects' reproduction (Silva-Oliveria et al. 2021; Pan et al. 2022). The N127 red aphid usually has a lower reproduction rate in comparison to N116 wingless, therefore it is possible that miR-278 might play an important role in the reproduction difference observed between these two genotypes by manipulating the insulin sensitivity.

Our results also showed that miR-2765 was upregulated in the N127 red morph in comparison to the N116 wingless morphs. The miRNA-2765 is usually not found in *Drosophila* and the function of this miRNA remains unknown. However, a recent study by Matsunami et al. (2018) has suggested that this miRNA is differentially expressed between different castes in termites and might be important in regulating caste differentiation. Further, a study has also reported that miRNA-2765 was a group of miRNAs that were highly conserved and specific to Hymenoptera (Søvik et al. 2015). Next, the miR-2765 has been shown to be differentially expressed in different developmental stages across different insect species such as the predatory bug *Arma chinesis* and Colorado potato beetle *Leptinotarsa decemlineata* (Yin et al. 2021; Weibe et al. 2020). However, the precise function of the miRNA in these insects across the developmental stage remains unknown. Our study is the first to report the differential expression of this miRNA between different aphid genotypes and further research is needed to elucidate the function of this miRNA in aphids.

Our results reveal that the miR-993 was downregulated in the N127 red morphs compared to the N116 wingless morphs. The precise function of miR-993 remains unclear. Furthermore, the miR-993 is specific to invertebrates only and was not present in any vertebrate organism. A recent study by Quah et al. (2015) reported that miR-993 was upregulated in the ovary of the speckled wood butterfly *Pararge aegeria* in comparison to the other body part. Although the function of these miR-993 in the ovary remains unclear, it might be involved in regulating reproduction. In aphids, the N127 red morph usually has a lower reproductive rate compared to the N116 wingless, therefore it is possible that this

miRNA might be responsible for regulating the reproduction between these two genotypes. Furthermore, target gene prediction in the silk *Bombyx mori* suggests that miR-993 could also target ecdysone which suggests that they could play an important role in metamorphosis in insects (He et al. 2019).

Lastly, our results reveal that miR-14 is also downregulated in the N127 red morph in comparison to the N116 green morph. As discussed above, the miR-14 plays an important role in regulating the ecdysone level in insects by acting on the ecdysone receptor (EcR) and Spook (He et al. 2019). Ecdysone plays an important role in insect development ranging from metamorphosis to reproduction. For example, Zhou et al. (2020) reported that in the rice planthopper *Nilaparvata lugens* with ecdysone enzyme knockdown showed a lower reproduction rate and a defect in embryonic development. Similar observations in reproduction rate were also observed in the blood-gorging insect *Rhodnius prolixus* when the ecdysone receptor was knockdown (Benrabaa et al. 2022). Taken together, this suggests that the role of ecdysone in manipulating insect reproduction might be conserved across insects. In aphids, the N127 genotype usually has lower reproduction compared to the N116 genotype and ecdysone might play a role in regulating the reproduction between these two genotypes.

Then, our target gene prediction results showed that among the differentially expressed miRNA, two miRNAs showed significant results (miR-278 and NW_021769971.1_unplaced_genomic_scaffold_39500). Some of the target genes regulated by miR-278 were involved in metabolism regulation, and insect development such as Facilitated trehalose transporter Tret 1 (Tret1), and glucose dehydrogenase (Gld). These genes have been reported in many studies in regulating insect reproduction, regulation of energy balance, and metamorphosis (Leyria et al. 2021; Liu et al. 2021; Li et al. 2022). Next, the novel miRNA NW_021769971.1_unplaced_genomic_scaffold_39500 seems to be involved in regulating many genes such as apterous1 (ap1), Krueppel homolog 1 (Kr-h1), ultrabithorax (Ubx), homothorax (hth), ecdysone receptor (EcR), wingless (Wnt-1),

spalt-major (salm). Ecdysone induced protein 75 (E75). Gustatory receptor for sugar taste 64f-like (Gr64f). These genes have remarkably diverse functions ranging from wing development, ecdysone regulation, metabolism regulation to sugar sensing. Taken together, this suggests that miRNA might play an important role in regulating different development aspects of aphids and might be responsible for the difference we observed between genotypes.

Correlation between miRNA and DEG

To provide some insight into the correlation between miRNA and DEG we integrated the results from the target gene prediction by miRNA with DEG from RNA-seq. In the group comparison of N116 winged and wingless, most genes were by three upregulated miRNAs (miRNA-100, miR-184a, miR-252a) and one downregulated miRNA (miRNA-2765). Most of the genes that overlap with DEG were downregulated. In the group comparison of N127 red vs N116 wingless, the gene was mostly targeted by one regulated miRNA (miR-278) and the downregulated miRNA (NW_021769971.1_unplaced_genomic_scaffold_39500). We observed a similar trend with the previous group comparison and found that most genes that overlap with DEG were downregulated. Our results are concordant with the consensus of a negative relationship between miRNA and gene expression. However, it is also important to note that recent studies have shown that miRNA is also able to induce gene expression and further studies are needed to elucidate the role of miRNA in insects (Xiao et al. 2017).

6.5 Summary

In this chapter, the potential role of miRNA in regulating pea aphid polyphenism and polymorphism was investigated between the different pea aphid morphs. We found that a few conserved miRNAs were differentially regulated between different pea aphid morphs. We also discovered many novel miRNAs that could potentially play an important role in regulating pea aphid polyphenism and polymorphism. Apart from that, we also investigated the potential target genes that could be targeted by these differentially expressed miRNAs. We found that many genes were differentially regulated by these miRNAs and some of these genes were involved in important processes such as wing development, stress response, ecdysteroid pathway regulation and metabolite regulation. Lastly, we integrated the transcriptome results from Chapter 4 to provide insight into the relationship between miRNAs and gene expression. Our results showed that there is a consensus of a negative relationship between miRNA and gene expression whereby most overlapped genes between target gene prediction in miRNA and DEG in RNA-seq were downregulated. Lastly, due to the limitation of low reads we obtained from the miRNA sequencing, it is possible that we have missed some of the miRNAs that could play an important role in pea aphid polyphenism. However, this does not undermine that we found many conserved miRNAs and novel miRNAs that could be important in regulating pea aphid polyphenism and polymorphism.

Chapter 7: General discussion

7.1 General discussion

The ability to fly has been an essential part of the evolutionary success of many insects as this allows the insects to escape from stressful conditions such as predators, a decline in food resources and increased temperature. According to the life-history theory, an increase in one life history trait that is beneficial to the organisms always comes with a decrease in another life history trait that is detrimental to the organisms or otherwise known as a trade-off. One of the most common trade-offs in insects is between the ability to fly and reproduction (Chang et al. 2021). Trade-offs play an important role in helping our understanding of evolutionary trajectories. A specific focus of the experiments presented here was to provide insight into the potential trade-offs between the ability to fly and the trade-off between different pea aphid morphs. By using two genotypes of aphids (N116 and N127) that react differentially to environmental stress, a major finding of this thesis is that the dispersal morphs always showed lower reproduction in comparison to the non-dispersal morphs. Although our experiment did not directly measure the reproductive success through the whole lifespan of the aphid, the lower reproductive success of the dispersal morphs observed in our experiment suggests that a potential trade-off between flight and reproduction might be involved in the pea aphids. This agrees with existing literature on the trade-offs between flight and reproduction that have been reported across a wide range of insect species (Nasu et al. 2021; Ge et al. 2021; Chang et al. 2021). One of the limitations of most studies is that the trade-offs have only been investigated in one generation. My thesis extended the investigation of the reproductive output of the offspring born to either dispersal morph or non-dispersal mothers to determine the possibility of trade-offs in more than one generation. We found that offspring that came from dispersal mothers still shows lower reproduction in comparison to those that came from the non-dispersal mothers. As mentioned above we did not measure the reproduction of the aphid throughout its lifespan. Therefore, our results may also suggest that potential trade-offs between the ability to disperse and reproduction in aphid might occur over more than one generation. Some

studies have shown that the parent's phenotypes play an important role in influencing the offspring's performance through a process known as a non-genetic inheritance (Bonduriansky et al. 2012).

Furthermore, my thesis also investigated the degree of dispersal morph production in different pea aphid genotypes. Aphids respond to stress very differently, and, depending on the genotypes, the production of winged and wingless morphs can range from 0% to 100% (Parker et al. 2019). Our experiments show that the two aphid genotypes react very differently to environmental stress, with the N116 genotype producing winged offspring and the N127 genotype changing their body colour from red to pale. Apart from that, our experiment also revealed that the degree of dispersal morph production is vastly different between the two genotypes, with N127 producing a higher level of dispersal morphs (pale) than N116 genotypes. This result agrees with previous studies that suggest that the changes in body colour in red aphid only requires 10h of starvation and does not require the transgenerational signalling that is usually seen in winged production (Wang et al. 2019). The differences in genotypic responses to environmental stress have been reported across plants, and mammals as well. For example, the salt stress response in the sugar beet *Beta vulgaris* is dependent on the genotypes (Geng et al. 2019). Then, Rohde et al. (2021) showed that in *Drosophila* exposed to rapamycin treatment, the increase in heat stress tolerance and reduction in fecundity is genotype dependent. Furthermore, Hidalgo et al. (2019) also showed that parasitic wasp (*Nasonia vitripennis)* genotypes react differently towards thermal stress.

The major focus of chapter 4 was to provide insight into the potential underlying mechanisms for wing development in aphids. Brisson et al. (2010) deduced the network of aphid wing development pathways based on *Drosophila* but found no significant difference in any wing development gene across developmental stages (embryo - fourth instar) apart from ap1, which shows a significant level during the 1st and 2nd nymph. The data presented in Chapter 4 showed that some of the wing development genes such as (ap1, and vestigial)

did show a significant difference in expression between the adult winged morphs and wingless morphs.

Since N127 does not produce wings but changes its body colour we investigated the expression of carotenoid genes. Our results revealed no significant difference in the carotenoid gene (tor, Figure 4.18G) but the N127 red morph does show a higher expression than the pale morphs. However, we still do not clearly understand the mechanisms behind the carotenoid breakdown in aphids and it is possible that both red and pale morphs are still producing the red pigments but were being subsequently broken down to provide energy for the aphids. In contrast to the wing dimorphism in male aphids which is known to be controlled by a single locus on the X-chromosome known as aphicarus (api), the wing polyphenism in female aphids remains largely unclear (Braendle et al. 2005). Therefore, in Chapter 4, we also carried out transcriptomic profiling of all different pea aphid morphs to provide further insight into the possible genes that could play an important role in aphid development and polyphenism. Our transcriptomic data revealed many genes that could potentially play an important role in wing polyphenism and stress responses in aphids. Many genes that could potentially explain the difference in life-history traits were observed between the two genotypes of aphids used in the experiment. Further, our data revealed a major difference in gene expression between the dispersal morphs across the two genotypes, whereby the N127 genotype morphs do not show any difference in wing development gene expression. Although densovirus has been reported to play a role in this genotype's ability to produce winged offspring, we found no difference in expression between morphs. Therefore, it is possible that the wing developmental pathway was repressed in certain aphid genotypes resulting in non-wing-producing genotypes, but further experiments are needed to confirm this. Finally, the data in this thesis chapter reinforce that some of the major wing developmental genes that were indeed differentially expressed between pea aphid morphs can be responsible for regulating wing development in the aphid. Our data also provide many novel genes that were differentially expressed

between pea aphid morphs that were shown to regulate wing development in other insect species, and further in vivo manipulation of the expression of these genes could provide a clearer understanding of the mechanisms of wing polyphenism and stress response in the aphid.

Following the gene expression and transcriptomic profile data in chapter 4, the methylation level of the candidate genes and the methylome profile of the different pea aphid morphs was investigated in Chapter 5. DNA methylation provides an epigenetic mechanism that has been reported to play a role in regulating polyphenism in insects. Since aphids reproduce asexually, therefore, all the offspring are clones of the mother, which makes them excellent in studying the role of epigenetic regulating in insect polyphenism. Surprisingly, the pyrosequencing of candidate genes from Chapter 5 reveals that only four out of the 16 genes investigated were significantly differentially methylated between the morphs. Further, our results in Chapter 5 also showed that genes that were involved in wing development and signalling process such as Fl, vg, ap1, en (Figure 5.3A-D) have a low level of methylation (1%-3%) compared to those that are ubiquitously expressed such as Hsp83, DMAP1, rpd3 (Figure 5.2 F-G, Figure 5.3F)ffigur which has intermediate (30-40%) to high level of methylation (70-100%) . Our results in chapter 5 are consistent with other studies which reported higher methylation in genes involved in housekeeping functions and low methylation in genes responsible for cellular signalling and polyphenism (Sarda et al. 2012; Glastad et al. 2011). Next, another study showed that heavily methylated genes across different insect species showed enriched gene ontology for functions such as RNA splicing, and protein localization, which further suggest that housekeeping genes have been consistently heavily methylated throughout the invertebrate evolution (Sarda et al. 2012). The higher methylation seen in housekeeping genes could be due to the need to control and reduce transcriptional noise (Wedd et al. 2022). Interestingly, a study by Zeng and Yi (2010) showed a relationship between the level of methylation and gene length in honeybees with longer genes having lower DNA methylation. A further study by Sarda et

al. (2012) reported similar observations in the silkworm and suggest that the relationship between gene length and DNA methylation might be specific to insects. However, in our study we did not include any analysis of the correlation between gene length and DNA methylation level. Therefore, further study is needed to confirm if a similar pattern is observed in pea aphids.

In Chapter 5 we utilised the MBD-seq to provide insight into the genome-wide methylation profile of aphids. Our results show that the methylation pattern of aphids was similar to most insects with gene body (intron+exon) having the highest level of methylation in comparison to other regions. Our results reveal that the X-chromosome is the primary target of methylation in aphids although it varies depending on genotypes and morphs. A recent study by Mathers et al. (2021) showed that X-linked genes were preferentially hypermethylated in the male aphids and hypomethylated in autosomal-related genes. Interestingly, our results found that X-linked genes were hypomethylated in winged morphs compared to wingless morphs while all the other group comparisons showed hypermethylation for the X-linked gene. A study by Gatzmann et al. (2018) reported that different tissue in marble crayfish showed that gene bodies is usually highly hypomethylated. A study of salinity stress in *Daphnia magna* also reported hypomethylation of genes responsible for stress response (Jeremias et al. 2018). An explanation for this observation may be that hypomethylation allows for a broader range of gene expression variation. However, further studies are needed to determine why such methylation patterns are only observed between the morphs in N116 genotype and not N127. Since we know that the phenotype outcomes are usually based on the changes in gene expression, one of the most important questions is how the environmental signal is translated to the changes in gene expression. Therefore, we integrated our MBD-seq and RNA-seq to determine the relationship between DNA methylation and gene expression and the role of DNA methylation in aphid wing polyphenism. Overall, our results reveal no clear link between DNA methylation and gene expression. However, some group comparisons in our results

did suggest a tendency of a positive relationship between gene body methylation and gene expression. To date, no clear conclusion can be drawn about the link between DNA methylation and gene expression with a lot of studies reporting contradicting results regarding the link between DNA methylation and gene expression (Duncan et al. 2022). Then, a functional analysis by Xu et al. (2021) showed that DNA methylation in *Bombyx mori* acts on the histone modification (H3K27) which regulated gene expression. Therefore, it is possible that DNA methylation and histone modification both interact to modulate gene expression. However, further studies are needed to determine if such a link is conserved between different insect species or specialised for the silkworm only.

Based on the results we obtained from Chapter 5, we decided to investigate the possibility of the role of other epigenetic mechanisms known as small non-coding RNA (miRNA) in regulating wing polyphenism (Chapter 6). miRNA plays an essential role in regulating gene expression at the post-transcriptional level (Neshat et al. 2020). Our results from Chapter 6 reveal that several differentially expressed miRNA between winged and wingless pea aphid morphs. Among these differentially expressed miRNAs, some have been shown to regulate wing development across multiple insect species (Li et al. 2022; Soares et al. 2021). We conducted target gene prediction to determine the possibility of these miRNAs in regulating genes that were responsible for wing development and reproduction. To our surprise, the results show that many wing development genes and reproductive genes were targeted by these differentially expressed miRNAs. These results suggest, in contrast to the general belief of DNA methylation as the main epigenetic mechanism in regulating insect polyphenism, miRNA might also play a very important role in regulating this polyphenism. In humans, some studies have shown that miRNAs are regulated by DNA methylation in regulating cancer progression and other diseases. (Aure et al. 2021; Mao et al. 2021). However, to date in insects, there is still no study that has integrated miRNA analysis and DNA methylation yet. Therefore, future work integrating miRNA analysis and DNA methylation might provide some new insight into the crosstalk of different epigenetic

mechanisms in regulating polyphenism in insects. One of the main functions of miRNA is to silence translation through destabilizing mRNA, resulting in repressing protein production. Therefore, we integrated the results from Chapter 4 (RNA-seq) with the miRNA-seq to determine the role of miRNA in insect gene expression. Our results reveal that most target genes that overlap with DEG in RNA-seq were downregulated. This agrees with most studies that show that miRNA is involved in gene repression (Mengistu et al. 2021).

7.2 Future directions

7.2.1 The trade-offs between different pea aphid morphs

The experiments presented in this thesis focus on determining the degree of alternative morph production between two different genotypes and the reproductive success of the alternative morphs and the subsequent generation. A major finding was that the aphid genotypes react differently to environmental stress and the reproductive output of the morphs varies between the morphs. These findings warrant further investigation of winged production of more different aphid genotypes and their associated reproductive output. Many studies have reported the variation of winged production in the aphid's genotype (Parker et al. 2019; Parker et al. 2021). However, the reproductive output of the winged morphs from different genotypes was not analysed. Therefore, it would be interesting to compare the reproductive output of dispersal morphs produced from different aphid genotypes to see if the variation in the number of winged morphs production affects the reproductive output of the dispersal morphs. A further major finding of this thesis reveals that the reproductive output of offspring from dispersal morphs was also reduced. In aphids' telescopic generations are very common, therefore the grandmaternal environment in our thesis (crowding) may cause the decrease in the reproduction output observed. The grandmaternal effect has been reported in other species such as *Physa acuta* whereby exposure of grand-parental exposure to predators causes changes in offspring shell thickness (Tariel et al. 2020). Further, similar observations were also reported in the

common lizard *Zootoca vivipara* whereby the grandmaternal age affects the grand offspring reproduction and survival irrespective of the maternal age (Bleu et al. 2021). Therefore, future studies in measuring the reproductive output throughout the aphid reproductive period and across genotypes are needed to elucidate if such a decrease in reproductive output is constant throughout the lifespan of an aphid and if the grandmaternal effect varies across genotypes.

7.2.2 The transcriptome profile and mechanisms of aphid development and polyphenism

The development of wings in *Drosophila* is well studied. In contrast, the underlying mechanism for regulating pea aphid wing development remains largely unclear. In this thesis, the qPCR and RNA-seq together show several known wing development genes and some novel genes that could play an important role in regulating pea aphid development and wing development (Chapter 4). For example, the master gene vestigial has been reported to play an important role in the wing development of many insect species (Zhang et al. 2021). In this thesis, our qPCR is the first to reveal vestigial was also differentially expressed in the winged aphid in comparison to other morphs and could play an important role in wing development in the aphid. However, future *in vivo* studies manipulating the expression of the genes are needed to elucidate their role in aphid wing development. Further, our RNA-seq also reveals some potential target genes such as Mad, Pburs that could potentially be important for wing development and further research is needed to determine their actual role in aphid wing development. Next, the ecdysone pathway was found to be important for wing development in some previous studies (Vellichirammal et al. 2017). In this thesis, we found ecdysone genes were differentially expressed between winged and wingless which agrees with the previous study. However, we also found some novel genes, for example, Kr-h1, a transcription factor, that can repress insect ecdysone signalling and therefore could play an important role in regulating transgenerational wing

polyphenism in aphids. The results of our RNA-seq provide a foundation for future studies in manipulating these target genes which can further our understanding of wing polyphenism in aphids.

7.2.3 The role of epigenetic mechanisms in regulating pea aphid wing polyphenism and development

Epigenetic mechanisms have been reported to play a major role in regulating gene expression across many organisms (Glastad et al. 2019). In this thesis, our data reveal that although some genes are differentially methylated between different aphid morphs, most of the genes, especially those involved in wing development, were not differentially methylated between aphid morphs. Further, the genome-wide results from my thesis (Chapter 5) suggest that there is no clear link between DNA methylation and gene expression. However, it is important to note that we did find a stronger positive trend in gene expression and DNA methylation, but this was dependent on genotype and morph type. However, a recent functional study by Xu et al. (2021) reported crosstalk between histone modification, DNA methylation and gene expression in the silkworm. Xu reported that insect gene body methylation regulates gene expression through histone acetylation specifically (H3K27). The results of Chapter 4 did reveal many histones that were differentially expressed between the winged and wingless morphs. Together the results of these studies and our thesis warrant further studies of histone modification together with DNA methylation, which might provide valuable insight into the crosstalk between DNA methylation, histone modification and gene expression. Uncovering more about the link between these epigenetic mechanisms and gene expression might help us to understand more about the role of methylation in insect polyphenism and development. Our MBD-seq results also reveal that the gene body is the most targeted region for methylation in aphids, which is similar to most insects (Xu et al. 2021; Pozo et al. 2021). Future studies of methylation studies on aphids should focus on these gene body regions to provide more

insight into the role of gene body methylation in regulating aphid development and polyphenism.

Based on the results we obtained from Chapter 5, we decided to further investigate other epigenetic mechanisms (miRNA) to determine their role in aphid development and wing polyphenism (Chapter 6). Our results reveal that a few conserved miRNAs such as miR-100 and let-7, which were differentially expressed between winged and wingless morphs, have also been shown to regulate wing development in other insects (Zhang et al. 2020; Soares et al. 2021). Further, we also identified some novel miRNAs that were highly expressed in the winged morphs compared to the wingless morphs. The target gene prediction in Chapter 6 also reveals many wing development target genes (e.g., Ubx, Dll, ap1) that were regulated by these miRNAs. These results together with those of previous studies further support the fact that miRNAs might play an important role in wing development in insects than previously thought. Therefore, future studies on manipulating the expression of these miRNAs could help to further our understanding of the role of miRNAs in aphid development and wing polyphenism.

7.3 Limitations

There are a few limitations of the data presented here in this thesis. In Chapter 3, we showed that the reproductive output of the dispersal morphs in both genotypes was lower compared to the non-dispersal morph. However, we did not measure reproduction of aphids throughout their lifespan nor the changes in energy or metabolites; therefore we can only suggest the possibility of trade-offs between reproduction and dispersal but are unable to demonstrate this directly. A similar limitation applies to the transgenerational experiment where we cannot clearly conclude that decreased reproduction in the offspring is due to the grand-parental effect nor that trade-offs in the dispersal morphs last more than one generation since most aphids have telescopic generations (Durak et al. 2016). Another limitation is that we did not collect the winged and wingless aphids once they emerge. We

could have included aphids of different reproductive ages, which may affect their reproductive output, noting that the mother's reproductive age has been reported to affect their reproduction in some organisms (Singh et al. 2020).

Further, there are also limitations that apply to Chapter 4, Chapter 5 and Chapter 6. For example, in all three chapters we only investigated gene expression, DNA methylation, and miRNA expression in whole adult aphids. This might result in a dilution effect whereby certain cells are over or under-represented in the samples, therefore, underestimating gene expression, DNA methylation and miRNA levels, as some of these might show cell/tissue-specific expression (Maleszka and Kucharski 2022). A further limitation is that only adult aphids were used in this thesis. Studies have shown that the critical period of wing development in aphids is between the 2nd and 3rd instar age whereby wingless destined nymph will histolyse their wing buds (Braendle et al. 2006). Therefore, the difference in gene expression, DNA methylation and miRNA might be more prominent during this nymph stage. However, it is important to bear in mind that it is physically impossible to distinguish the winged and wingless nymph during the 1st and 2nd instar. Therefore, future studies, which should incorporate aphids of different developmental stages and cell-type specific analysis might provide clearer information into the role of epigenetic and underlying mechanism in aphid development and wing polyphenism.

7.4 Summary

In summary, this thesis has established that the aphid genotype varies in its ability to trigger wing polyphenism. There might also be a possibility of trade-offs between different pea aphid morphs and this effect might last more than one generation. We also showed differential expression in candidate genes involved in aphid development and wing polyphenism between different aphid morphs and genotypes. Further, we found overall no evidence of a clear association between DNA methylation and gene expression in regulating aphid development and wing polyphenism. However, DNA methylation did show

a stronger positive relationship with increased gene expression in some cases, but this was dependent on genotype and morph type. Lastly, my thesis showed that other epigenetic mechanisms such as miRNA could be important in regulating wing polyphenism as shown by the target gene prediction results.

References

Ables ET, Bois KE, Garcia CA, Drummond-Barbosa D. 2015. Ecdysone response gene E78 controls ovarian germline stem cell niche formation and follicle survival in *Drosophila*. Dev Biol. 400:33-42

Adachi-Yamade T, Nakamura M, Irie K, Tomoyasu Y, Sano Y, Mori E, Goto S, Ueno N, Nishida Y, Matsumoto K. 1999. p38 Mitogen-activated protein kinase can be involved in transforming growth factor β superfamily signal transduction in *Drosophila* wing morphogenesis. Mol Cell Biol. 19:2322-2329

Adams D, Wilkinson TL, Douglas AE. 1996. The aphid–bacterial symbiosis: a comparison between pea aphids and black bean aphids. Ent. exp. appl. 80:275-278

Agrawal BK. 2007. Phenotypic plasticity in aphids (Homoptera: Insecta): Components of variation and causative factors. Curr Sci. 93:308-313

Akiyama T, Kamimura K, Firkus C, Takeo S, Shimmi O, et al. 2008. Dally regulates Dpp morphogen gradient formation by stabilizing Dpp on the cell surface. Dev Biol. 313:408–419

Ali I, Conrad RJ, Verdin E, Ott M. 2018. Lysine acetylation goes global: From epigenetics to metabolism and therapeutics. Chem Rev. 118:1216-1252

Altincicek B, Gross J, Vilcinskas A. 2008. Wounding-mediated gene expression and accelerated viviparous reproduction of the pea aphid *Acyrthosiphon pisum*. Insect Mol Biol. 17:711-716

Ambrosi C, Manzo M, Baubec T. 2017. Dynamics and context-dependent roles of DNA methylation. J Mol Biol. 10:1459-1475

Arakane Y, Li B, Muthukrishnan S, Beeman RW, Kramer KJ, Park Y. 2008. Functional analysis of four neuropeptides, EH, ETH, CCAP and bursicon, and their receptors in adult ecdysis behavior of the red flour beetle, *Tribolium castaneum*. Mech Dev 125:984-985

Arnold P.A, Nicotra AB, Kruuk LE. 2019. Sparse evidence for selection on phenotypic plasticity in response to temperature. Philos Trans R Soc Lond B Biol Sci. 374:20180185

Arsala D, Wu X, Yi SV, Lynch JA. 2022. Dnmt1a is essential for gene body methylation and the regulation of the zygotic genome in a wasp. PLoS Genet. 18:e1010181

Arsenault SV, Hunt BG, Rehan SM. 2018. The effect of maternal care on gene expression and DNA methylation in a subsocial bee. Nat Commun. 9:3468

Asgari S. 2012. MicroRNA functions in insects. Insect Biochem Mol Biol. 43:388-397

Auld JR, Relyea RA. 2010. Adaptive plasticity in predator-induced defenses in a common freshwater snail: altered selection and mode of predation due to prey phenotype. Evol Ecol. 25:189-202

Aure MR, Fleischer T, Bjørklund S, Ankill J, Castro-Mondragon JA, OSBREAC, Børresen-Dale A, Tost J, Sahlberg KK, Mathelier A, Tekpli X, Kristensen VN. 2021. Crosstalk between microRNA expression and DNA methylation drives the hormone-dependent phenotype of breast cancer. Genome Med. 13:72

Baena-López LA, García-Bellido A. 2003. Genetic requirements of vestigial in the regulation of *Drosophila* wing development. Development. 130:197-208

Baldwin JM. 1902. Development and evolution. New York (NY): The McMillan Company.

Barrett RDH, Hendry AP. 2012. Evolutionary rescue under environmental change? In: Candolin U Wong BBM, editors. Behavioural responses to a changing world: mechanisms and consequences. London (UK): Oxford University Press. p. 216–233

Bartel DP, Chen C. 2004. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat Rev Genet. 5:396-400

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. Cell. 136:215-233

Baumann P, Lai C, Baumann L, Rouhbakhsh D, Moran NA, Clark MA. 1995. Mutualistic Associations of Aphids and Prokaryotes: Biology of the Genus *Buchnera*. Appl Environ Microbiol. 61:1-7

Beasley-Hall PG, Bertozzi T, Bradford TM, Foster CSP, Jones K, Tierney SM, Humphreys WF, Austin AD, Cooper SJB. 2022. Differential transcriptomic responses to heat stress in surface and subterranean diving beetles. Sci Rep. 12:16194

Becam I, Rafel N, Hong X, Cohen SM, Milán M. 2011. Notch-mediated repression of bantam miRNA contributes to boundary formation in the *Drosophila* wing. Development. 138:3781-3789

Becker C, Hogmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D. 2011. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. Nature. 480:245-249

Beeler SM, Wong GT, Zheng JM, Bush EC, Remnant EJ, Oldroyd BP, Drewell RA. 2014. Whole-Genome DNA methylation profile of the jewel wasp (*Nasonia vitripennis*). G3(Betesda). 4:383-388

Benrabaa S, Orchard I, Lange AB. 2022. The role of ecdysteroid in the regulation of ovarian growth and oocyte maturation in *Rhodnius prolixus*, a vector of Chagas disease. J Exp Biol. 225:jeb.244830

Bento G, Ogawa A, Sommer RJ. 2010. Co-option of the hormone-signalling module dafachronic acid–DAF-12 in nematode evolution. Nature. 466:494-497

Bewick AJ, Sanchez X, Mckinney EC, Moore AJ, Moore PJ, Schmitz RJ. 2019. Dnmt1 is essential for egg production and embryo viability in the large milkweed bug, *Oncopeltus fasciatus*. Epigenetics Chromatin. 12:6

Bewick AJ, Vogel KJ, Moore AJ, Schmitz RJ. 2017. Evolution of DNA Methylation across Insects. Mol Biol Evol. 34:654-665

Bird AP. 1987. CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3:342-347

Bird AP. 2002. DNA methylation patterns and epigenetic memory. Genes Dev. 16:6-21

Biryukova I, Asmar J, Abdesselem H, Heitzler P. 2009. *Drosophila* mir-9a regulates wing development via fine-tuning expression of the LIM only factor, dLMO. Dev Biol. 327:487-496

Bleu J, Meylan S, Clobert J, Massot M. 2021. Grandmaternal age at reproduction affects grandoffspring body condition, reproduction and survival in a wild population of lizards. Funct Ecol. 36:826-836

Bludau A, Royer M, Meister G, Neumann ID, Menon R. 2019. Epigenetic regulation of the social brain. Trends Neurosci. 42:471-484.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 15:2114-2120

Bonasio R, Li Q, Lian J, Mutti NS, Jin L, Zhao H, Zhang P, Wen P, Xiang H, Ding Y, Jin Z, Shen SS, Wang Z, Wang W, Wang J, Berger SL, Leibig J, Zhang G, Reinberg D. 2012. Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Curr Biol. 22:1755-1764

Bonasio R. 2014. The role of chromatin and epigenetics in the polyphenisms of ant castes. Brief Funct Genomics. 13:235-245

Bonduriansky R, Crean AJ, Day T. 2012. The implications of nongenetic inheritance for evolution in changing environments. Evol Appl. 5:192-201

Bonduriansky R. 2012. Rethinking heredity, again. Trends Ecol Evol. 27:330-336

Bossdorf O, Richards CL, Pigliucci M. 2007. Epigenetics for ecologists. Ecol Lett. 11:106-115

Both C, Bouwhuis S, Lessells CM, Visser ME. 2006. Climate change and population declines in a long-distance migratory bird. Nature. 441:81-83

Botuyan MV, Lee J, Ward IM, Kim J, Thompson JR, Chen J, Mer G. 2006. Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. Cell. 127:1361-1373

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. 2001. Dnmt3L and the establishment of maternal genomic imprints. Science. 5551:2536-2539

Bradford MJ, Roff DA. 1993. Bet hedging and the diapause strategies of the cricket *Allonemobius Fasciatus*. Ecology. 74:1129-1135

Braendle C, Davis GK., Brisson JA, Stern DL. 2006. Wing dimorphism in aphids. Heredity. 97: 192-199

Braendle C, Friebe I, Caillaud MC, Stern DL. 2005. Genetic variation for an aphid wing polyphenism is genetically linked to a naturally occurring wing polymorphism. Proc Royal Soc B. 272:657-664

Braendle C, Miura T, Bickel R, Shingleton AW, Kambhampati S, Stern DL. 2003. Developmental origin and evolution of bacteriocytes in the aphid–*buchnera* symbiosis. PLoS Biol. 1:e21

Brisson JA, Ishikawa A, Miura T. 2010. Wing development genes of the pea aphid and differential gene expression between winged and unwinged morphs. Insect Mol Biol. 2:63-73

Bryant PJ. 1978. Pattern formation in imaginal discs. In the genetics and biology of *Drosophila* (Ashburner, M. and Twright, R.F., eds), pp. 230–236. Academic Press, New York.

Buck RJ, Fienberg J, Larkin DJ. 2022. The use of weighted averages of Hedges' d in meta-analysis: Is it worth it? Methods Ecol Evol. 5:1093-1105

Candolin U. 1998. Reproduction under predation risk and the trade–off between current and future reproduction in the threespine stickleback. Proc Royal Soc B. 265:1402

Cao T, Jin J. 2020. Evolution of flight muscle contractility and energetic efficiency. Front Physiol. 11:1038

Cardoso-Júnior CAM, Yagound B, Ronai I, Remnant EJ, Hartfelder K, Oldroyd BP. 2021. DNA methylation is not a driver of gene expression reprogramming in young honey bee workers. Mol Ecol. 30:4804-4818

Caruso CM. 2006. Plasticity of inflorescence traits in *Lobelia siphilitica* (Lobeliaceae) in response to soil water availability. Botany. 93:531-538

Caygill EE, Johnston LA. 2008. Temporal regulation of metamorphic processes in *Drosophila* by the let-7 and miR-125 heterochronic microRNAs. Curr Biol. 18:943-950

Cedar H, Bergman Y. 2009. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet. 10:295-304

Champagne DL, Bagot RC, van Hasselt F, Ramakers G, Meaney MJ, Ronald de Kloet E, Joëls M, Krugers H. 2008. Maternal care and hippocampal plasticity: evidence for experience-dependent structural plasticity, altered synaptic functioning, and differential responsiveness to glucocorticoids and stress. J Neurosci. 28:6037-6045

Chang H, Guo X, Guo S, Yang N, Huang Y. 2021. Trade-off between flight capability and reproduction in Acridoidea (Insecta: Orthoptera). Ecol Evol. 11:16849-16861

Chang M, Cheng H, Cai Z, Qian Y, Zhang K, Yang L, Ma N, Li D. 2022. miR-92a-1-p5 modulated expression of the flightin gene regulates flight muscle formation and wing extension in the pea aphid, *Acyrthosiphon pisum* (Hemiptera: Aphidoidea). J Insect Sci. 22:14

Chapuis M, Simpson SJ, Blondin L, Sword GA. 2011. Taxa-specific heat shock proteins are over-expressed with crowding in the Australian plague locust. J Insect Physiol. 57:1562-1567

Charles H, Balmand S, Lamelas A, Cottret L, Pérez-Brocal V, Burdin B, Latorre A, Febvay G, Colella S, Calevro F, Rahbé Y. 2011. A genomic reappraisal of symbiotic function in the aphid/*Buchnera* symbiosis: reduced transporter sets and variable membrane organisations. PLoS One. 6:e29096

Charmantier A, McCleery RH, Cole LR, Perrins C, Kruuk LEB, Sheldon BC. 2008. Adaptive phenotypic plasticity in response to climate change in a wild bird population. Science. 320:800-803

Chechenova MB, Maes S, Oas ST, Nelson C, Kiani KG, Bryantsev AL, Cripps RM. 2017. Functional redundancy and nonredundancy between two Troponin C isoforms in *Drosophila* adult muscles. Mol Biol Cell. 28:760-770

Chen B, Wagner A. 2012. Hsp90 is important for fecundity, longevity, and buffering of cryptic deleterious variation in wild fly populations. BMC Ecol. 12:25

Chen D, Montllor CB, Purcell A.H. 2000. Fitness effects of two facultative endosymbiotic bacteria on the pea aphid, Acyrthosiphon pisum, and the blue alfalfa aphid, *A. kondoi*. Entomol Exp Appl. 95:315-323

Chen D, Purcell AH. 1997. Occurrence and transmission of facultative endosymbionts in aphids. Curr Microbiol. 34:220-225

Chen K, Tang T, Song Q, Wang Z, He K, Liu X, Song J, Wang L, Yang Y, Feng C. 2019. Transcription analysis of the stress and immune response genes to temperature stress in *Ostrinia furnacalis*. Front Physiol. 10:1289

Chen W, Li D, Zhang M, Zhao Y, Wu W, Zhang G. 2015. Cloning and differential expression of five heat shock protein genes associated with thermal stress and development in the polyphagous predatory mite *Neoseiulus cucumeris* (Acari: Phytoseiidae). Exp Appl Acarol. 67:65-85

Chen X, Fu J. 2021. The microRNA miR-14 regulates egg-laying by targeting EcR in honeybees (*Apis mellifera*). Insects. 12:351

Cheng X, Blumenthal RM. 2008. Mammalian DNA methyltransferases: a structural perspective. Structure. 16:341-350

Chevin LM, Lande R, Mace GM. 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. PLoS Biol. 8:1-8

Chittka A, Wurm Y, Chittka L. 2012. Epigenetics: The making of ant castes. Curr Biol. 22:R835-R838

Chong S, Whitelaw E. 2004. Epigenetic germline inheritance. Curr Opin Genet Dev. 6:692-696

Chowdhury D, Keogh M, Ishii H, Peterson CL, Buratowski S, Lieberman J. 2005. Gamma-H2AX dephosphorylation by protein phosphatase 2A facilitates DNA double-strand break repair. Mol Cell. 20:801-809

Churchill GA. 2007. Recombinant inbred strain panels: a tool for systems genetics. Physiol Genomics. 31:174-175

Cohen B, McGuffin ME, Pfeifle C, Segal D, Cohen SM. 1992. Apterous, a gene required for imaginal disc development in *Drosophila* encodes a member of the LIM family of developmental regulatory proteins. Genes Dev. 6:715-729

Cohen SM. 1993. Imaginal disc development. In The development of *Drosophila melanogaster* (Bate, M. and Martinez-Arias, A., eds), pp. 747–841. Cold Spring Harbor Laboratory Press, Plainview, NY.

Conover D, Present TMC. 1990. Countergradient variation in growth rate: compensation for length of the growing season among Atlantic silversides from different latitudes. Oecologia. 83:316–324.

Corchete LA, Rojas EA, Alonso-López D, Rivas JDL, Gutiérrez NC, Burguillo FJ. 2020. Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. Sci Rep. 10:19737

Corchete LA, Rojas EA, Alonson- López, De Las Rivas J, Gutiérrez NC, Burguillo FJ. 2020. Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. Sci Rep. 10:19737

Costopoulos K, Kovacs JL, Kamins A, Gerardo NM. 2014. Aphid facultative symbionts reduce survival of the predatory lady beetle *Hippodamia convergens*. BMC Ecol. 14:5

Courtney MC, Linz DM, Tomoyasu Y. 2013. Insights into insect wing origin provided by functional analysis of vestigial in the red flour beetle, *Tribolium castaneum*. Proc Natl Acad Sci U S A. 110:16951-16956

Covic M, Hassa PO, Saccani S, Buerki C, Meier NI, Lombardi C, Imhof R, Bedford MT, Natoli G, Hottiger MO. 2005. Arginine methyltransferase CARM1 is a promoter-specific regulator of NF-kappaB-dependent gene expression. EMBO J. 24:85-96

Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K. 2011. Effects of abiotic stress on plants: a systems biology perspective. BMC Plant Biol. 11:163

Crispo E, DiBattista JD, Correa C, Thibert-Plante X, Mckellar AE, Schwartz AK, Berner D, De León IF, Hendry AP. 2010. The evolution of phenotypic plasticity in response to anthropogenic disturbance. Evol Ecol Res. 12:47–66

Crispo E. 2007. The baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. Evolution. 61:2469–2479

Cristino AS, Barchuk AR, Freitas FC, Narayana RK, Biergans SD, Zhao Z, Simoes ZL, Reinhard J, Caludianos C. Neuroligin-associated microRNA-932 targets actin and regulates memory in the honeybee. Nat Commun. 20:5529

Crossley AC, Crossley N, Crossley C. 1978. The morphology and development of the *Drosophila* muscular system, M. Ashburner, T.R.F. Wright (eds.). Academic Press, London.

Cruz J, Mané-Padrós D, Bellés X, Martín D. 2006. Functions of the ecdysone receptor isoform-A in the hemimetabolous insect *Blattella germanica* revealed by systemic RNAi in vivo. Dev Biol. 297:158-171

Cubas P, Vincent C, Coen E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. Nature. 401:157-161

Cunningham CB, Ji L, McKinney EC, Benotwitz KM, Schmitz RJ, Moore AJ. 2019. Changes of gene expression but not cytosine methylation are associated with male parental care reflecting behavioural state, social context and individual flexibility. J Exp Biol. 222:jeb188649

Dasch GA, Weiss E. 1992. The genera Rickettsia, Rochalimaea, Ehrlichia, Cowdaria, and Neorickettsia, p. 2407-2470. In A. Balows, H.G. Truper, M. Dworkin, W. Harder, and K.-H. Schleifer (ed.), The prokaryotes, vol. 3. Springer-Verlag, New York, N.Y.

Davidson AM, Jennions M, Nicotra AB. 2011. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. Ecol Lett. 14:419-431

Dawson KJ. 1996. Evolutionary consequences of a trade-off between parental effort and mating effort. J Theor Biol. 183:139-158

Day T, McPhail JD. 1996. The effect of behavioural and morphological plasticity on foraging efficiency in the threespine stickleback (*Gasterosteus sp*). Oecologia 108:380–388

Deans C, Maggert KA. 2015. What do you mean, "epigenetic"? Genetics. 4:887-896

Degnan PH, Moran NA. 2008. Diverse phage-encoded toxins in a protective insect endosymbiont. Appl Environ Microbiol. 74:6782-6791

DeWitt TJ, Robinson BW, Wilson DS. 2000. Functional diversity among predators of a freshwater snail imposes an adaptive trade-off for shell morphology. Evol Ecol Res. 2:129-148

Ding X, Liu J, Zheng L, Song J, Li N, Hu H, Tong X, Dai F. 2019. Genome-wide identification and expression profiling of Wnt family genes in the silkworm, *Bombyx mori*. Int J Mol Sci. 20:1221

Dingle H. 1996. Migration: the biology of life on the move. Oxford University Press, Oxford, United Kingdom

Dixon AFG. and Howard, MT. 1986. Dispersal in aphids, a problem in resource allocation. in insect flight: dispersal and migration (Danthanarayana, W., ed.), pp. 145–151. Springer-Verlag, Berlin.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2012. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 1:15-21

Dogulas AE. 1998. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. Annu Rev Entomol. 43:17-37

Doskocil J, Sorm F. 1962. Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids. Biochim Biophys Acta. 55:953-959

Dottermusch-Heidel C, Gärtner SMK, Tegeder I, Rathke C, Barckmann B, Bartkuhn M, Bhushan S, Steger K, Meinhardt A, Renkawaitz-Pohl R. 2014. H3K79 methylation: a new conserved mark that accompanies H4 hyperacetylation prior to histone-to-protamine transition in *Drosophila* and rat. Biol Open. 3:444-452

Dubrova YE, Plumb M, Gutierrez B, Boulton E, Jeffreys A.J. 2000. Transgenerational mutation by radiation. Nature. 405:37

Dufour HD, Koshikawa S, Finet C. 2020. Temporal flexibility of gene regulatory network underlies a novel wing pattern in flies. Proc Natl Acad Sci U S A. 117:11589-11596

Ehrentraut S, Weber JM, Dybowski JN, Ehrenhofer-Murray AE. 2010. Rpd3-dependent boundary formation at telomeres by removal of Sir2 substrate. Proc Natl Acad Sci U S A. 107:5522-5527

Elango N, Hunt BG, Goodisman MAD, Yi SV. 2009. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. Proc Natl Acad Sci U S A. 106:11206-11211

Elias-Neto M, Alvarez N, Ventos-Alfonso A, Belles X. 2022. Flight or protection: the genes Ultrabithorax and apterous in the determination of membranous and sclerotized wings in insects. Proc Royal Soc B. 289:20220967

Ellefsen KO. 2013. Balancing the Costs and Benefits of Learning Ability. ECAL 2013: The Twelfth European Conference on Artificial Life

Eller F, Brix H. 2012. Different genotypes of Phragmites australis show distinct phenotypic plasticity in response to nutrient availability and temperature. Aquat Bot. 103:89-97

Enders LS, Bickel RD, Brisson JA, Heng-Moss TM, Siegfried BD, Zera AJ, Miller NJ. 2014. Abiotic and biotic stressors causing equivalent mortality induce highly variable transcriptional responses in the soybean aphid. G3 (Bethesda). 5:261-270

Entrambasaguas L, Ruocco M, Verhoeven KJF, Procaccini G, Marín-Guirao L. 2021. Gene body DNA methylation in seagrasses: inter- and intraspecific differences and interaction with transcriptome plasticity under heat stress. Sci Rep. 11:14343

Esslinger SM, Schwalb B, Helfer S, Michalik KM, Witte H, Maier KC, Martin D, Michalke B, Tresch A, Cramer P, Förstemann K. 2013. *Drosophila* miR-277 controls branched-chain amino acid catabolism and affects lifespan. RNA Biol. 10:1042-1056

Ezard T.H.G, Prizak R, Hoyle R.B. 2014. The fitness costs of adaptation via phenotypic plasticity and maternal effects. Funct Ecol. 28:693-704

Fabbrizio E, Messaoudi SE, Polanowska J, Paul C, Cook JR, Lee J, Negre V, Rousset M, Pestka S, Cam AL, Sardet C. 2002. Negative regulation of transcription by the type II arginine methyltransferase PRMT5. EMBO Rep. 3:641-645

Fabian D, Flatt T. 2012. Life history evolution. Nature Education Knowledge 3:24

Fagegaltier D, König A, Gordon A, Lai EC, Gingeras TR, Hannon GJ, Shcherbata HR. 2014. A genome-wide survey of sexually dimorphic expression of *Drosophila* miRNAs identifies the steroid hormone-induced miRNA let-7 as a regulator of sexual identity. Genetics. 198:647-668

Fan Y, Li X, Mohammed AAAH, Liu Y, Gao X. 2020. miR-147b-modulated expression of vestigial regulates wing development in the bird cherry-oat aphid *Rhopalosiphum padi*. BMC Genom. 21:71

Farahani S, Bandani AR, Alizadeh H, Goldansaz SH, Whyard S. 2020. Differential expression of heat shock proteins and antioxidant enzymes in response to temperature, starvation, and parasitism in the carob moth larvae, *Ectomyelois ceratoniae* (Lepidoptera: Pyralidae). PLoS One. 15:e0228104

Febvay G, Rhabe Y, Rynkiewicz M, Guillaud J, Bonnot G. 1999. Fate of dietary sucrose and neosynthesis of amino acids in the pea aphid, *Acyrthosiphon pisum*, reared on different diets. J Exp Biol. 202:2639-2652

Feliciello I, Parazjder J, Akrap I, Ugarković D.2013. First evidence of DNA methylation in insect *Tribolium castaneum*. Epigenetics. 8:534-541

Felsenfeld G. 2014. The evolution of epigenetics. Perspect Biol Med. 1:132-148

Ferreira PG, Patalano S, Chauhan R, Ffrench-Constant R, Gabaldón T, Guigó Roderic, Sumner S. 2013. Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. Genome Biol. 14:R20

Field LM. 2000. Methylation and expression of amplified esterase genes in the aphid *Myzus persicae* (Sulzer). Biochem J. 349:863-868

Flores K, Wolschin F, Corneveaux JJ, Allen AN, Huentelman MJ, Amdam GV. 2012. Genome-wide association between DNA methylation and alternative splicing in an invertebrate. BMC Genet. 13:480

Fordcye JA. 2006. The evolutionary consequences of ecological interactions mediated through phenotypic plasticity. J Exp Biol. 209:2377-2383

Foret S, Kucharski R, Pittelkow Y, Lockett GA, Maleszka R. 2009. Epigenetic regulation of the honey bee transcriptome: unravelling the nature of methylated genes. BMC Genet. 10:472

Fox CW, Czesak ME. 2000. Evolutionary ecology of progeny size in arthropods. Annu Rev Entomol. 45:341-369

Freeman AS, Byers JE. 2006. Divergent induced responses to an invasive predator in marine mussel populations. Science. 313:831–833

Freitak D, Knorr A, Vogel H, Vilcinskas A. 2012. Gender- and stressor-specific microRNA expression in *Tribolium castaneum*. Biol Lett. 8:860-863

Fujinaga D, Gu J, Kawahara H, Orighara MH, Kojima I, Takeshima M, Katoaka H. 2020. Twenty-hydroxyecdysone produced by dephosphorylation and ecdysteroidogenesis regulates early embryonic development in the silkmoth, *Bombyx mori*. Insect Biochem Mol Biol. 127:103491

Fussmann GF, Loreau M, Abrams P.A. 2007. Eco-evolutionary dynamics of communities and ecosystems. Funct Ecol. 21:465-477

Ganassi S, Signa G, Mola L. 2005. Development of the wing buds in *Megoura viciae*: a morphological study. Bull Insect. 58:101-105

Gao J, Guo HJ, Sun CY, Ge F. 2018. Juvenile hormone mediates the positive effects of nitrogen fertilization on weight and reproduction in pea aphid. Pest Manag Sci. 74:2511-2519

Gao X, Fu Y, Ajayi OE, Guo D, Zhang L, Wu Q. 2019. Identification of genes underlying phenotypic plasticity of wing size via insulin signaling pathway by network-based analysis in *Sogatella furcifera*. BMC Genom. 20:396

Gao Y, Ren R, Peng J, Wang D, Shi X, Zheng L, Zhang Z, Zhu C, Liu Y, Dai L, Zhang D. 2021. The gustavus gene can regulate the fecundity of the green peach aphid, *Myzus persicae* (Sulzer). Front Physiol. 11:596392

Gatzmann F, Falckenhayn C, Gutekunst J, Hanna K, Raddatz G, Carneiro CV, Lyko F. 2018. The methylome of the marbled crayfish links gene body methylation to stable expression of poorly accessible genes. Epigenetics Chromatin. 11:57

Gavery MR, Roberts SB. 2010. DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). BMC Genom. 11:483

Ge S, Sun X, He W, Wyckhuys KAG, He L, Zhao S, Zhang H, Wu K. 2021. Potential trade-offs between reproduction and migratory flight in *Spodoptera frugiperda*. J Insect Physiol. 132:104248

Ge X, Zhang Y, Jiang J, Zhong Y, Yang X, Li Z, Huang Y, Tan A. 2013. Identification of MicroRNAs in *Helicoverpa armigera* and *Spodoptera litura* based on deep sequencing and homology analysis. Int J Biol Sci. 9:1-15

Gegner J, Gegner T, Vogel H, Vilcinskas A. 2020. Silencing of the DNA methyltransferase 1 associated protein 1 (DMAP1) gene in the invasive ladybird *Harmonia axyridis* implies a role

of the DNA methyltransferase 1-DMAP1 complex in female fecundity. Insect Mol Biol. 29:148-159

Geng G, Lv C, Stevanato P, Li R, Liu H, Yu L, Wang Y. 2019. Transcriptome analysis of salt-sensitive and tolerant genotypes reveals salt-tolerance metabolic pathways in sugar beet. Int J Mol Sci. 20:5910

Ghalambor CK, McKay JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Func Ecol. 21:394–407

Gianoli E, González-Teuber M. 2005. Environmental heterogeneity and population differentiation in plasticity to drought in *Convolvulus Chilensis* (Convolvulaceae). Evol Ecol. 19:603-613

Gill JA, Alves JA, Sutherland WJ, Appleton GF, Potts PM, Gunnarsson TG. 2013. Why is timing of bird migration advancing when individuals are not? Proc Royal Soc B. 281:20132161

Glastad KM, Gokhale K, Liebig J, Goodisman MAD. 2016. The caste- and sex-specific DNA methylome of the termite *Zootermopsis nevadensis*. Sci Rep. 6:37110

Glastad KM, Hunt BG, Yi S.V, Goodisman MAD. 2011. DNA methylation in insects: on the brink of the epigenomic era. Insect Mol Biol. 20:553-565

Goll GM, Bestor TH. 2005. Eukaryotic cytosine methyltransferases. Annu Rev Biochem. 74:481-514

Goll GM, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, Golic KG, Jacobsen SE, Bestor TH. 2006. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science. 311:395-398

Gómez-Valero L, Silva FJ, Simon JC, Latorre A. 2007. Genome reduction of the aphid endosymbiont *Buchnera aphidicola* in a recent evolutionary time scale. Gene. 389:87-95

Gonzalo S, Jaco I, Fraga MF, Chen T, Li E, Esteller M, Blasco MA. 2006. DNA methyltransferases control telomere length and telomere recombination in mammalian cells. Nat Cell Biol. 4:416-424

Grantham ME, Antonio CJ, O'Neil BR, Zhan YX, Brisson JA. 2016. A case for a joint strategy of diversified bet hedging and plasticity in the pea aphid wing polyphenism. Biol Lett. 12:20160654

Grardo NM, Altincicek B, Anselme C, Atamian H, Barribeau SM, de Vos M, Duncan EJ, Evans JD, Gabaldón T, Ghanim M, Heddi A, Kaloshian I, Latorre A, Moya A, Nakabachi A, Parker BJ, Pérez-Brocal V, Pignatelli M, Rahbé, Y, Ramsey JS, Spragg CJ, Tamames J, Tamarit D, Tamborindeguy C, Vincent-Monegat C, Vilcinskas A. 2010. Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. Genome Biol. 11:R21

Grether GE. 2005. Environmental change, phenotypic plasticity, and genetic compensation. Am Nat. 166:115-123

Griffith SJ, Mahler HR. 1969. DNA ticketing theory of memory. Nature. 223:580-582

Guan D, Ding R, Hu X, Yang X, Xu S, Gu W, Zhang M. 2019. Cadmium-induced genome-wide DNA methylation changes in growth and oxidative metabolism in *Drosophila melanogaster*. BMC Genet. 20:356

Gujar H, Palli SR. (2016a). Juvenile hormone regulation of female reproduction in the common bed bug, *Cimex Lectularius*. Sci Rep. 6:35546

Gündüz EA, Douglas AE. 2009. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. Proc Royal Soc B. 276:987-991

Hansen AK, Moran NA. 2011. Aphid genome expression reveals host–symbiont cooperation in the production of amino acids. Proc Nat Acad Sci U S A. 108:2849-2854

Hansford RG, Johnson RG. 1975. The nature and control of tricarboxylate cycle in beetle flight muscle. Biochem J. 14:389-402

Hardie J, Leckstein P. 2007. Antibiotics, primary symbionts and wing polyphenism in three aphid species. Insect Biochem Mol Biol. 37:886-890

Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature. 405:486-489

Harris KD, Lloyd JPB, Domb K, Zilberman D, Zemach A. 2019. DNA methylation is maintained with high fidelity in the honey bee germline and exhibits global non-functional fluctuations during somatic development. Epigenetics Chromatin. 12:62

Harrison MC, Dohmen E, George S, Sillam- Dussès D, Séité S, Vasseur-Cognet M. 2022. Complex regulatory role of DNA methylation in caste- and age-specific expression of a termite. Open Biol. 12:220047

Harshman LG, Zera AJ. 2006. The cost of reproduction: the devil in the details. Trends Ecol Evol. 22:80-86

He K, Xiao H, Sun Y, Ding S, Situ G, Li F. 2019. Transgenic microRNA-14 rice shows high resistance to rice stem borer. Plant Biotchnol J. 17:451-471

Heinrich B. 1975. Thermoregulation in bumblebees – II. Energetics of warm-up and free flight. J Comp Physiol. 96:155-166

Hendry AP 2015. Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. Heredity. 1:25-41

Henkin JA, Maughan DW, Vigoreaux JO. 2004. Mutations that affect flightin expression in *Drosophila* alter the viscoelastic properties of flight muscle fibers. Am J Physiol Cell Physiol. 286:C65-C72

Herb BR, Shook MS, Field CJ, Robinson GE. 2018. Defense against territorial intrusion is associated with DNA methylation changes in the honey bee brain. BMC Genom. 19:216

Herb BR, Wolschin F, Hansen KD, Aryee MJ, Langmead B, Irizarry R, Amdam GV, Feinberg AP. 2012. Reversible switching between epigenetic states in honeybee behavioral subcastes. Nat Neurosci. 15:1371-1373

Herman J, Spencer JHG, Donohue K, Sultan SE. 2014. How stable 'should' epigenetic modifications be? Insights from adaptive plasticity and bet hedging. Evolution. 68:632-643

Hidalgo K, Beaugeard E, Renault D, Dedeine F, Lécureuil C. 2019. Physiological and biochemical responses to thermal stress vary among genotypes in the parasitic wasp *Nasonia vitripennis*. J Insect Physiol. 117:103909

Hill K, Kaplan H. 1999. Life history traits in humans: theory and empiricial studies. Annu Rev Anthropol. 28:397-430

Hirata K, Jouraku A, Kuwazaki S, Shimomura H, Iwasa T. 2017. Studies on *Aphis gossypii* cytochrome P450s CYP6CY22 and CYP6CY13 using an in vitro system. J Pestic Sci. 42:97-104

Hoch J.A. 2000. Two-component and phosphorelay signal transduction. Curr Opin Microbiol. 3:165-170

Hollander J. 2008. Testing the grain-size model for the evolution of phenotypic plasticity. Evolution. 62:1381-1389

Holliday R, Pugh JE. 1975. DNA modification mechanisms and gene activity during development. Science. 187:226-232

Holliday R. 2006. Epigenetics: a historical overview. Epigenetics. 2:76-80

Hou L, Wang X, Yang P, Li B, Lin Z, Kang L, Wang X. 2020. DNA methyltransferase 3 participates in behavioral phase change in the migratory locust. Insect Biochem Mol Biol. 121:103374

Houk EJ, Griffiths GW, Hadjokas NE, Beck SD. 1977. Peptidoglycan in the cell wall of the primary intracellular symbiote of the pea aphid. Science. 198:401-403

Hu D, Luo W, Fan LF, Liu FL, Gu J, Deng HM, Zhang C, Huang LH, Feng QL. 2016. Dynamics and regulation of glycolysis-tricarboxylic acid metabolism in the midgut of *Spodoptera litura* during metamorphosis. Insect Mol Biol. 25:153-162

Hu K, Tian P, Yang L, Tang Y, Qiu L, He H, Ding W, Li Y. 2020. Molecular characterization of the Krüppel-homolog 1 and its role in ovarian development in *Sogatella furcifera* (Hemiptera: Delphacidae). Mol Biol Rep. 47:1099-1106

Hunt BG, Brisson JA, Yi SV, Goodisman MAD. 2010. A functional DNA methylation system in the pea aphid, *Acyrthosiphon pisum*. Genome Biol Evol. 2:719-728

Hussain M, Torres S, Schnettler E, Funk A, Grundhoff A, Pijlman GP, Khromykh AA, Asgari S. 2012. West Nile virus encodes a microRNA-like small RNA in the 3′ untranslated region which upregulates GATA4 mRNA and facilitates virus replication in mosquito cells. Nucleic Acid Res. 40:2210-2223

Hutvágner G, McLachlan J, Pasquinelli AE, Bálint E, Tuschl T, Zamore P.D. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science. 293:834-838

Ishikawa A, Hongo S, Miura T. 2008. Morphological and histological examination of polyphenic wing formation in the pea aphid *Acyrthosiphon pisum* (Hemiptera, Hexapoda). Zoomorphology 127:121–133.

Ishimaru Y, Tomonari S, Matsuoka Y, Watanabe T, Miyawaki K, Bando T, Tomioka K, Ohuchi H, Noji S, Mito T. 2016. TGF-β signaling in insects regulates metamorphosis via juvenile hormone biosynthesis. Proc Natl Acad Sci U S A. 113:5634-5639

Ishimaru Y, Tomonari S, Watanabe T, Noji S, Mito T. 2019. Regulatory mechanisms underlying the specification of the pupal-homologous stage in a hemimetabolous insect. Proc Royal Soc B. 374:20190225

Itoh T, Martin W, Nei M. 2002. Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. Proc Nat Acad Sci U S A. 99:12944-12948

Itoyama K, Morooka S, Tojo S. 1999. Triacylglycerol storage in the adults of two selected strains of the brown planthopper, *Nilaparvata lugens*. Appl Entomol Zool. 34:171-177

Jablonka E. 2013. Epigenetic inheritance and plasticity: The responsive germline. Prog Biophys Mol Biol. 111(2-3):99-107

Jaenisch R, Bird A. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 33:245-254

Jain S, Rana V, Tridibes A, Sunil S, Bhatnagar RK. 2015. Dynamic expression of miRNAs across immature and adult stages of the malaria mosquito *Anopheles stephensi*. Parasit Vectors. 8:179

Jeltsch A, Gowher H. 2019. Editorial—Role of DNA methyltransferases in the epigenome. Genes. 8:574

Jeremias G, Barbosa J, Marques SM, De Schamphelaere KAC, Nieuwerburgh FV, Deforce D, Gonçalves FJM, Pereira JL, Asselman J. 2018. Transgenerational inheritance of DNA hypomethylation in *Daphnia magna* in response to salinity stress. Environ Sci Technol. 52:10114-10123

Jeyaprakash A, Hoy MA. 2000. Long PCR improves *Wolbachia* DNA amplification: wsp sequences found in 76% of sixty-three arthropod species. Insect Mol Biol. 9:393-405

Jia S, Wan P, Zhou L, Mu L, Li G. 2013. Knockdown of a putative Halloween gene Shade reveals its role in ecdysteroidogenesis in the small brown planthopper *Laodelphax striatellus*. Gene. 531:168-174

Jiang J, Ge X, Li Z, Wang Y, Song Q, Stanley DW, Tan A, Huang Y. 2013. MicroRNA-281 regulates the expression of ecdysone receptor (EcR) isoform B in the silkworm, *Bombyx mori*. Insect Biochem Mol Biol. 43:692-700

Jin H, Kim VN, Hyun S. 2012. Conserved microRNA miR-8 controls body size in response to steroid signaling in *Drosophila*. Genes Dev. 26:1427-1432

Jin M, Xue J, Yao Y, Lin X. 2014. Molecular Characterization and Functional Analysis of Krüppel-homolog 1 (Kr-h1) in the Brown Planthopper, *Nilaparvata lugens* (Stål). J Integr Agric. 13:1972-1981.

Jin. H, Kim VN, Hyun S. 2012. Conserved microRNA miR-8 controls body size in response to steroid signaling in *Drosophila*. Genes Dev. 25:1427-1432

Jockusch EL, Ober KA. 2004. Hypothesis testing in evolutionary developmental biology: a case study from insect wings. Heredity. 95:382-396

Johannee F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuisson J, Heredia F, Audigier P, Bouchez D, Dillmann C, Guerche P, Hospital F, Colot V. 2009. Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genet. 5:e1000530

Johnson AA, Stolzing A. 2019. The role of lipid metabolism in aging, lifespan regulation, and age-related disease. Aging Cell. 18:e13048

Johnson B, Birks PR. 1960. Studies on wing polymorphism in aphids. The developmental process involved in the production of the different forms. Entomol Exp Appl. 3:327-339

Jones PA. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 13:484-492

Kanduri C, Pant V, Loukinov D, Pugacheva E, Qi C, Wolffe A, Ohlsson R, Lobanenkov VV. 2000. Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. Curr Biol. 14:853-856

Kangassalo K, Sorvari J, Nousiainen I, Pölkki M, Valtonen T.M, Krams I, Rantala MJ. 2020. Intra- and trans-generational phenotypic responses of the greater wax moth, *Galleria mellonella*, to a Low-Nutrition Larval Diet. Ann Zool Fenn. 57:99-114

Kanvil S, Powell G, Turnbull C. 2014. Pea aphid biotype performance on diverse medicago host genotypes indicates highly specific virulence and resistance functions. Bull Entomol Res. 104:689-701

Kareta M, Botello SZM, Ennis J.J, Chou C, Chédin F. 2006. Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. J Biol Chem. 36:25893-25902

Kausar S, Abbas MN, Yang L, Cui H. 2020. Biotic and abiotic stress induces the expression of Hsp70/90 organizing protein gene in silkworm, *Bombyx mori*. Int J Biol Macromol. 143:610-618

Kay S, Skowronski D, Hunt BG. 2018. Developmental DNA methyltransferase expression in the fire ant *Solenopsis invicta*. Insect Sci. 25:57-65

Kayukawa T, Murata M, Kobayashi I, Muramatsu D, Okada C, Uchino K, Sezutsu H, Kiuchi M, Tamura T, Hiruma K, Ishikawa Y, Shino T. 2014. Hormonal regulation and developmental role of Krüppel homolog 1, a repressor of metamorphosis, in the silkworm *Bombyx mori*. Dev Biol. 388:48-56

Keller TE, Han P, Yi S.V. 2016. Evolutionary transition of promoter and gene body DNA methylation across invertebrate-vertebrate boundary. Mol Biol Evo. 33:1019-1028

Kennell JA, Cadigan KM, Shakhmantsir I, Waldron EJ. 2012. The microRNA miR-8 is a positive regulator of pigmentation and eclosion in *Drosophila*. Dev Dyn. 241:161-168

Kilvitis H, Alvarez JM, Foust CM, Schrey AW, Robertson M, Richards CL. 2014. Ecological epigenetics. Adv Exp Med Biol. 781:191-210

Kim J, Sebring A, Esch JJ, Kraus ME, Vorwerk K, Magee J, Carroll SB. 1996. Integration of positional signals and regulation of wing formation and identity by *Drosophila* vestigial gene. Nature. 382:133-138

Kim JK, Samaranayake M, Pradhan S. 2009. Epigenetic mechanisms in mammals. Cell Mol Life Sci. 66:596-612

Kim K, Vinayagam A, Perrimon NA. 2014. A rapid genome-wide microRNA screen identifies miR-14 as a modulator of Hedgehog signaling. Cell Rep. 7:2066-2077

King EG, Roff DA, Fairbairn DJ. 2011. Trade-off acquisition and allocation in *Gryllus firmus*: a test of the Y model. J Evol Biol. 24:256-264

Kinnison MT, Hairston JR NG. 2007. Eco-evolutionary conservation biology: contemporary evolution and the dynamics of persistence. Funct Ecol. 21:444-454

Kirfel P, Skalijac M, Grotmann J, Kessel T, Seip M, Michaelis K, Vilcinskas A. 2020. Inhibition of histone acetylation and deacetylation enzymes affects longevity, development, and fecundity in the pea aphid (*Acyrthosiphon pisum*). Arch Insect Biochem Physiol. 103:e21614

Konopova B, Smykal V, Jindra M. 2011. Common and distinct roles of juvenile hormone signaling genes in metamorphosis of holometabolous and hemimetabolous insects. PLoS One. 6:e28728

Kopp Z, Park Y. 2019. Longer lifespan in the Rpd3 and Loco signaling results from the reduced catabolism in young age with noncoding RNA. Aging. 11:230-239

Kovach-Orr C, Fussmann GF. 2013. Evolutionary and plastic rescue in multitrophic model communities. Philos Trans R Soc B: Biol Sci. 368:20120084

Kozomara A, Griffiths-Jones S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. 39:D152-157

Kramer JM, Slade JD, Staveley BE. 2008. Foxo is required for resistance to amino acid starvation in *Drosophila*. Genome. 51:668-672

Kronforst MR, Gilley DC, Strassmann JE, Queller DC. 2008. DNA methylation is widespread across social hymenoptera. Curr Biol. 18:R287-288

Kucharski R, Maleszka J, Foret S, Maleszka R. 2008. Nutritional control of reproductive status in honeybees via DNA methylation. Science. 319:1827-1830

Kugler J, Chen Y, Weng R, Cohen SM. 2013. miR-989 Is required for border cell migration in the *Drosophila* ovary. PLoS One. 8:e67075

Lafont R, Dauphin-Villemant C, Warren JT, Rees H. 2012. Insect Endocrinology, L.I. Gilbert (Eds.), Ecdysone chemistry and Biochemistry, Elsevier Press, Amsterdam, pp. 106-176

Lago DC, Humann FC, Barchuk AR, Abraham KJ, Hartfelder K. 2016. Differential gene expression underlying ovarian phenotype determination in honey bee, *Apis mellifera* L., caste development. Insect Biochem Mol Biol. 79:1-12

Lahr EC, Dean D, Ewer J. 2012. Genetic Analysis of Ecdysis Behavior in *Drosophila* reveals partially overlapping functions of two unrelated neuropeptides. J Neurosci. 32:6819-6829

Lahti DC, Johnson NA, Ajie BC, Otto SP, Hendry AP, Blumstein DT, Coss R.G, Donohue K, Foster SA. 2009. Relaxed selection in the wild. Trends Ecol Evol. 24:487–496

Lai EC, Tomancak P, Williams RW, Rubin GM. 2003. Computational identification of *Drosophila* microRNA genes. Genome Biol. 4:R42

Lailvaux SP, Husak JF. 2014. The life history of whole-organism performance. Q Rev Biol. 89:285-318

Lan H, Hong X, Huang R, Lin X, Li Q, Li K, Zhou T. 2018. RNA interference-mediated knockdown and virus-induced suppression of Troponin C gene adversely affect the behavior or fitness of the green rice leafhopper, *Nephotettix cincticeps*. Arch Insect Biochem Physiol. 97

Lan H, Zhang Z, Wu J, Cao H, Liu T. 2021. Performance and transcriptomic response of the English grain aphid, *Sitobion avenae*, feeding on resistant and susceptible wheat cultivars. J Intergr Argic. 20:178-190

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

Law R. 1979. Optimal life histories under age-specific predation. Am Nat. 114:399-417

Leather SR. 1989. Do alate aphids produce fitter offspring? The influence of maternal rearing history and morph on life-history parameters of *Rhopalosiphum padi* (L.). Funct Ecol. 3:237-244

Ledón-Rettig C. 2013. Ecological Epigenetics: An introduction to the symposium. Integr Comp Biol. 53:307-318

Lee GE, Kim JH, Taylor M, Muller MT. 2010. DNA methyltransferase 1-associated protein (DMAP1) is a co-repressor that stimulates DNA methylation globally and locally at sites of double strand break repair. J Biol Chem. 285:37630-37640

Lee GJ, Jun JW, Hyun S. 2015. MicroRNA miR-8 regulates multiple growth factor hormones produced from *Drosophila* fat cells. Insect Mol Biol. 24:311-318

Lee RC, Feinbaum RL, Ambros V. 1993. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 75:843-854

Lee S, Lee S, Paik D, Min K. 2012. Overexpression of fatty-acid-β-oxidation-related genes extends the lifespan of *Drosophila melanogaster*. Oxid Med Cell Longev. 2012:854502

Lee Y, Kim M, Han J, Yeom K, Lee S, Baek S, Kim VN. 2004. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 23:4051-4060

Legeai F, Rizk G, Walsh T, Edward O, Gordon K, Lavenier D, Leterme N, Méreau A, Nicolas J, Tagu D, Jaubert-Possamai S. 2010. Bioinformatic prediction, deep sequencing of microRNAs and expression analysis during phenotypic plasticity in the pea aphid, *Acyrthosiphon pisum*. BMC Genom. 11:281

Lei L, Zhang Z, Groenewoud F, Kingma SA, Li J, van der Velde M, Komdeur J. 2020. Extra-pair mating opportunities mediate parenting and mating effort trade-offs in a songbird. Behav Ecol. 31:421-431

Leyria J, El-Mawed H, Orchard I, Lange AB. 2021. Regulation of a trehalose-specific facilitated transporter (TRET) by insulin and adipokinetic hormone in *Rhodnius prolixus*, a vector of chagas disease. Front Physiol. 12:624165

Li B, Hou L, Zhu D, Xu X, An S, Wang X. 2018. Identification and caste-dependent expression patterns of DNA methylation associated genes in *Bombus terrestris*. Sci Rep. 1:2332

Li B, Hu P, Zhu L, You L, Cao H, Wang J, Zhang S, Liu M, Toufeeq S, Huang S, Xu J. 2020. DNA methylation is correlated with gene expression during diapause termination of early embryonic development in the silkworm (*Bombyx mori*). Int J Mol Sci. 21:671

Li C, Zhang Y, yun X, Wang Y, Sang M, Liu X, Hu X, Li B. 2014. Methuselah-like genes affect development, stress resistance, lifespan and reproduction in *Tribolium castaneum*. Insect Mol Biol. 23:587-597

Li E. 2002. Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet. 9:662-673

Li H, Wang X, Zheng X, Dong Z, Yi X, Lu W. 2022. Transcriptome and proteome analysis of oviposition- and spermatogenesis-related genes of *Diaphorina citri*. Animal Gene. 23:200120.

Li KL, Yuan SY, Nanda S, Wang WX, Lai FX, Fu Q, Wan PJ. 2018. The roles of E93 and Kr-h1 in metamorphosis of *Nilaparvata lugens*. Front Physiol. 9:1677

Li L, Wang S, Huang K, Zhang Y, Li Y, Zhang M, Huang J, Deng Z, Ni X, Li X. 2021. identification and characterization of MicroRNAs in gonads of *Helicoverpa armigera* (Lepidoptera: Noctuidae). Insects. 12:749

Li X, Zhang F, Coates B, Wei C, Zhu X, Zhang Y, Zhou X. 2022. Temporal analysis of microRNAs associated with wing development in the English grain aphid, *Sitobion avenae* (F.) (Homoptera: Aphidiae). Insect Biochem Mol Biol. 142:103579

Li X, Zhang F, Coates B, Zhang Y, Zhou X, Cheng D. 2016. Comparative profiling of microRNAs in the winged and wingless English grain aphid, *Sitobion avenae* (F.) (Homoptera: Aphididae). Sci Rep. 6:35668

Li Y, Chen X, Wang S, Pan B, Wang S, Wang S, Tang B. 2019. Evaluation of the expression and function of the TRE2-like and TRE2 genes in ecdysis of *Harmonia axyridis*. Front Physiol. 10:1371

Li Y, Tian C, Wu Y, Niu T, Wang H, Fan W, Liu H. 2021. Enhanced expression of DNA methyltransferase 1-associated protein1 gene thermotolerance in a high-temperature acclimated predatory mite *Neoseiulus barkeri*. BioControl. 66:779-788

Li Z, Li Q, Liu S, Han Z, Kong L, Yu H. 2021. Integrated analysis of coding genes and non-coding rnas associated with shell color in the pacific oyster (*Crassostrea gigas*). Mar Biotechnol. 23:417-429

Li-Byarlay H, Boncristiani H, Howell G, Herman J, Clark L, Strand MK, Tarpy D, Rueppell O. 2020. Transcriptomic and epigenomic dynamics of honey bees in response to lethal viral infection. Front Genet. 11:566320

Lind MI, Johansson F. 2007. The degree of adaptive phenotypic plasticity is correlated with the spatial environmental heterogeneity experienced by island populations of *Rana temporaria*. J Evol Biol. 20:1288-1297

Lind MI, Johansson F. 2009. Costs and limits of phenotypic plasticity in island populations of the common frog *Rana temporaria* under divergent selection pressures. Evolution. 63:1508-1518

Lindhart YB, Grant MC. 1996. Evolutionary significance of local genetic differentiation in plants. Ann Rev Ecol Syst. 27:237–277

Ling L, Kokoza VA, Zhang C, Aksoy E, Raikhel AS. 2017. MicroRNA-277 targets insulin-like peptides 7 and 8 to control lipid metabolism and reproduction in *Aedes aegypti* mosquitoes. Proc Natl Acad Sci U S A. 114:E8017-E8024

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 462:315

Liu F, Li K, Li J, Hu D, Zhao J. He Y, Zou Y, Feng Y, Hua H. 2015. Apterous A modulates wing size, bristle formation and patterning in *Nilaparvata lugens*. Sci Rep. 5:10526

Liu K, Chen Z, Su Q, Yue L, Chen W, Zhang W. 2020. Comparative analysis of the ecological fitness and transcriptome between two genotypes of the brown planthopper *Nilaparvata lugens*. J Interg Argic. 19:1501-1511

Liu S, Li K, Gao Y, Liu X, Chen W, Ge W, Feng Q, Palli SR, Li S. 2018. Antagonistic actions of juvenile hormone and 20-hydroxyecdysone within the ring gland determine developmental transitions in *Drosophila*. Proc Natl Acad Sci U S A. 115:139–44

Liu X, Chen G, He J, Wan G, Shen D, Xia A, Chen F. 2021. Transcriptomic analysis reveals the inhibition of reproduction in rice brown planthopper*, Nilaparvata lugens*, after silencing the gene of MagR (IscA1). Insect Mol Biol. 30:253-263

Liu Z, Ling L, Xu J, Zeng B, Huang Y, Shang P, Tan A. 2018. MicroRNA-14 regulates larval development time in *Bombyx mori*. Insect Biochem Mol Biol. 93:57-65

Long G, Yang J, Kin D, Yang H, Zhou C, Wang Z. 2022. Silencing of Decapentaplegic (Dpp) gene inhibited the wing expansion in the white-backed planthopper, *Sogatella furcifera* (Horváth) (Hemiptera: Delphacidae). Arch Insect Biochem Physiol. 110:e21879

Lorenz MW. 2007. Oogenesis-flight syndrome in crickets: age-dependent egg production, flight performance, and biochemical composition of the flight muscles in adult female *Gryllus bimaculatus*. J Insect Physiol. 53:819-832

Losos JB, Schoener TW, Langerhans RB, Spiller DA. 2006. Rapid temporal reversal in predator-driven natural selection. Science. 314:1111

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550

Lozano J, Belles X. 2011. Conserved repressive function of Krüppel homolog 1 on insect metamorphosis in hemimetabolous and holometabolous species. Sci Rep. 1:163

Lu M, Liu S, Kumarsangaiah A, Zhou Y, Pan Z, Zuo Y. 2017. Nucleosome positioning with fractal entropy increment of diversity in telemedicine. IEEE Access. 99:1-1

Lucas KJ, Roy S, Ha J, Gervaise AL, Kokoza VA, Raikhel AS. 2015. MicroRNA-8 targets the Wingless signaling pathway in the female mosquito fat body to regulate reproductive processes. Proc Natl Acad Sci U S A. 112:1440-1445

Łukasik P, van Asch M, Guo H, Ferrari J, Godfray HCJ. 2012. Unrelated facultative endosymbionts protect aphids against a fungal pathogen. Ecol Lett. 16:214-218

Lundsgaard-Hansen B, Matthews B, Seehausen O. 2014. Ecological speciation and phenotypic plasticity affect ecosystems. Ecology. 95:2723–2735

Luo C, Dewey EM, Sudo S, Ewer J, Hsu SY, Honegger H, Hsueh AJ. 2005. Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G protein-coupled receptor LGR2. Proc Natl Acad Sci U S A. 102:2820-2825

Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, Maleszka R. 2010. The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol. 8:e1000506

Lyytinen A, Brakefield PM, Lindström L, Mappes J. 2004. Does predation maintain eyespot plasticity in *Bicyclus anyana*. Proc Royal Soc B. 271:279-283

Ma L, Chen F, Wang W, Xu L, Lu Z. 2020. Identification of two clip domain serine proteases involved in the pea aphid's defense against bacterial and fungal infection. Insect Sci. 27:735-744

Ma L, Liu L, Zhao Y, Yang L, Chen C, Li Z, Lu Z. 2020. JNK pathway plays a key role in the immune system of the pea aphid and is regulated by microRNA-184. PLoS Pathog. 16:e1008627

Ma Z, Guo W, Guo X, Wang X, Kang L. 2011. Modulation of behavioral phase changes of the migratory locust by the catecholamine metabolic pathway. Proc Natl Acad Sci U S A. 108:3882-3887

Machado-Assefh CR, Lopez-Isasmendi G, Tjallingii WF, Jander G, Alvarez AE. 2015. Disrupting *Buchnera aphidicola*, the endosymbiotic bacteria of *Myzus persicae*, delays host plant acceptance. Arthropod-Plant Interact. 9:529-541

Maksimoska J, Segura PD, Cole PA, Marmorstein R. 2014. Structure of the p300 Histone acetyltransferase bound to acetyl-coenzyme A and its analogues. Biochemistry. 53:3415-3422

Maleszka R, Kucharski R. 2022. Without mechanisms, theories and models in insect epigenetics remain a black box. Trends Genet. 38:1108-1111

Mandrioli, M. 2004. Epigenetic tinkering and evolution: is there any continuity in the role of cytosine methylation from invertebrates to vertebrates? Cell Mol Life Sci. 61:2425-2427

Manuck SB. 2010. The reaction norm in gene-environment interaction. Mol Psychiatry. 15:991-882

Mao Z, Li T, Zhao H, Qin Y, Wang X, Kang Y. 2021. Identification of epigenetic interactions between microRNA and DNA methylation associated with polycystic ovarian syndrome. J Hum Genet. 66:123-137

Marco A, Hui JHL, Ronshaugen M, Griffiths-Jones S. 2010. Functional shifts in insect microRNA evolution. Genome Biol Evol. 2:686-696

Marshall H, Lonsdale ZN, Mallon EB. 2019. Methylation and gene expression differences between reproductive and sterile bumblebee workers. Evol Lett. 3:485-499

Mashoodh R, Sarkies P, Westoby J, Kilner RM. 2021. Evolved changes in DNA methylation in response to the sustained loss of parental care in the burying beetle. bioRxiv 2021.

Mathers T.C, Mugford ST, Percival-Alwyn L, Chen Y, Kaithakottil G, Swarbreck D, Hogenhout SA, van Oosterhout C. 2019. Sex-specific changes in the aphid DNA methylation landscape. Mol Ecol. 28:4229-4241

Mathers TC, Wouters RHM, Mugford ST, Swarbreck D, van Oosterhout C, Hogenhout SA. 2020. Chromosome-scale genome assemblies of aphids reveal extensively rearranged autosomes and long-term conservation of the X chromosome. Mol Biol Evol. 38:856-875

Matsunami M, Nozawa M, Suzuki R, Toga K, Masuoka Y, Yagamuchi K, Maekwa K, Shigenobu S, Miura T. 2019. Caste-specific microRNA expression in termites: insights into soldier differentiation. Insect Mol Biol. 28:86-89

Matthews B, Narwani A, Hausch S, Nonaka E, Peter H, Yamamichi M, Sullam KE, Bird KC, Thomas MK, Hanley TC, Turner CB. 2011. Toward an integration of evolutionary biology and ecosystem science. Ecol Lett. 14:690–701

McCue MD, Guzman RM, Passement CA, Davidowitz G. 2015. How and when do insects rely on endogenous protein and lipid resources during lethal bouts of starvation? A new application for 13C-Breath testing. PLoS One. 10:e0140053

Mehrparvar M, Zytynska SE, Weisser WW. 2013. Multiple cues for winged morph production in an aphid Metacommunity. PLoS One. 8:e58323

Melnattur K, Zhang B, Shaw PJ. 2020. Disrupting flight increases sleep and identifies a novel sleep-promoting pathway in *Drosophila*. Sci Adv. 6:eaaz2166

Mengistu AA, Tenkegna TA. 2021. The role of miRNA in plant–virus interaction: a review. Mol Biol Rep. 48:2853-2861

Miao ZQ, Tu YQ, Guo PY, He W, Jing TX, Wang JJ, Wei DD. 2020. Antioxidant enzymes and heat shock protein genes from *Liposcelis bostrychophila* are involved in stress defense upon heat shock. Insects. 11:839

Mira A, Ochman H, Moran NA. 2001. Deletional bias and the evolution of bacterial genomes. Trends Genet. 17:589-596

Mishra D, Miyamoto T, Rezenom YH, Broussard A, Yavuz A, Slone J, Russell DH, Amrein H. 2013. The molecular basis of sugar sensing in *Drosophila* larvae. Curr Biol. 23:1466-1471

Mohan KN, Ding F, Chailet JR. 2011. Distinct roles of DMAP1 in mouse development. Mol Cell Biol. 31:1861-1869

Mondal T, Rasmussen M, Pandey GK, Isaksson A, Kanduri C. 2010. Characterization of the RNA content of chromatin. Genome Res. 7:899-907

Montllor CB, Maxmen A, Purcell AH. 2002. Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. Ecol Entomol. 27:189-195

Moran NA, Dunbar HE, Wilcox JL. 2005. Regulation of transcription in a reduced bacterial genome: nutrient-provisioning genes of the obligate symbiont *Buchnera aphidicola*. J Bacteriol. 187:4229-4237

Moran NA, Jarvik T. 2010. Lateral transfer of genes from fungi underlies carotenoid production in aphids. Science. 328:624-627

Moran NA, Mira A. 2001. The process of genome shrinkage in the obligate symbiont *Buchnera aphidicola*. Genome Biol. 2:RESEARCH0054

Moran NA, Plague GR, Sandström JP, Wilcox JL. 2003. A genomic perspective on nutrient provisioning by bacterial symbionts of insects. Proc Nat Acad Sci U S A. 100:14543-14548

Moran NA, Wernegreen JJ. 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. Trends Ecol Evol. 15:321-326

Morandin C, Brendel VP, Sundström L, Helanterä H, Mikheyev AS. 2019. Changes in gene DNA methylation and expression networks accompany caste specialization and age-related physiological changes in a social insect. Mol Ecol. 28:1975-1993

Morandin C, Brendel VP. 2021. Tools and applications for integrative analysis of DNA methylation in social insects. Mol Ecol Resour. 22:1656-1674

Morata G, Lawrence PA. 1975. Control of compartment development by the engrailed gene in *Drosophila*. Nature. 255:614-617

Mukherjee K, Baudach A, Vogel H, Vilcinskas A. 2020. Seasonal phenotype-specific expression of microRNAs during metamorphosis in the European map butterfly *Araschnia levana*. Arch Insect Biochem Physiol. 104:e21657

Mukherjee K, Dubovskiy I, Grizanova E, Lehmann R, Vilcinskas A. 2019. Epigenetic mechanisms mediate the experimental evolution of resistance against parasitic fungi in the greater wax moth *Galleria mellonella*. Sci Rep. 9:1626

Müller CB, Williams IS, Hardie J. 2001. The role of nutrition, crowding, and interspecific interactions in the development of winged aphids. Eco Ent. 26:330-340

Munson MA, Baumann P, Clark MA, Baumann L, Moran NA, Voegtlin DJ, Campbell BC. 1991. Aphideubacterial endosymbiosis: evidence for its establishment in an ancestor of four aphid families. J Bacteriol Res. 173:6321-6324

Murat S, Hopfen C, McGregor AP. 2010. The function and evolution of Wnt genes in arthropods. Arthropod Struct Dev. 39:446-452

Nahmad M, Glass L, Abouheif E. 2008. The dynamics of developmental system drift in the gene network underlying wing polyphenism in ants: a mathematical model. Evol Dev. 10:360-374

Nakajima E, Shimaji K, Umegawachi T, Tomida S, Yoshida H, Yoshimoto N, Izawa S, Kimura H, Yamaguchi M. 2016. The histone deacetylase gene Rpd3 Is Required for starvation stress resistance. PLoS One. 11:e0167554

Nasu S, Tokuda M. 2021. Dispersal–reproduction trade-off in the leaf beetle *Galerucella grisescens*. Entomol. Exp. Appl. 169:542-549

Neshat SY, Tzeng SY, Green JJ. 2020. Gene delivery for immunoengineering. Curr Opin Biotechnol. 66:1-10

Nettle D, Frankenhuis WE. 2020. Life-history theory in psychology and evolutionary biology: one research programme or two? Proc Royal Soc B. 375:1803

Nikoh N, McCutcheon JP, Kudo T, Miyagishima S, Moran NA, Nakabachi A. 2010. Bacterial genes in the aphid genome: absence of functional gene transfer from *Buchnera* to its host. PLoS Genet. 6:e1000827

Ning L, Wang L, Sang W, Liu C, Qiu B. 2018. Effects of endosymbiont disruption on the nutritional dynamics of the pea aphid *Acyrthosiphon pisum*. Insects. 9:161

Nirgianaki A, Banks GK, Frohlich DR, Veneti Z, Braig HR, Miller TA, Bedford ID, Markham PG, Savakis C, Bourtzis K. 2003. *Wolbachia* Infections of the Whitefly *Bemisia tabaci*. Curr Microbiol. 47:0093-0101

Nussey DH, Postma E, Gienapp P, Visser ME. 2005. Selection on heritable phenotypic plasticity in a wild bird population. Science. 310:304-306

Ogawa K, Ishikawa A, Kanbe T, Akimoto S, Miura T. 2012. Male-specific flight apparatus development in *Acyrthosiphon pisum* (Aphididae, Hemiptera, Insecta): comparison with female wing polyphenism. Zoomorphology. 131: 197-207

Ogawa K, Miura T. 2013. Two developmental switch points for the wing polymorphisms in the pea aphid *Acyrthosiphon pisum*. EvoDevo. 4:30

Ogawa K, Miura T. 2014. Aphid polyphenisms: trans-generational developmental regulation through viviparity. Front Physiol. 5:1

Okamura K, Matsumoto KA, Nakai K. 2010. Gradual transition from mosaic to global DNA methylation patterns during deuterostome evolution. BMC Bioinform. 11:S2

Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for denovo methylation and mammalian development. Cell. 99:247-257

Paim RMM, Araunjo RN, Leis M, Sant'anna MRV, Gontijo NF, Lazzari CR, Pereira MH. 2016. Functional evaluation of heat shock proteins 70 (HSP70/HSC70) on *Rhodnius prolixus* (Hemiptera, Reduviidae) physiological responses associated with feeding and starvation. Insect Biochem Mol Biol. 77:10-20

Pan X, Pei Y, Zhang C, Huang Y, Chen L, Wei L, Li C, Dong X, Chen X. 2022. Effect of Insulin Receptor on Juvenile Hormone Signal and Fecundity in *Spodoptera litura* (F.). Insects. 13: 701

Paoloni-Giacobino A, Chaillet JR. 2004. Genomic imprinting and assisted reproduction. Reprod Health. 1:6

Parker BJ, Brisson JA. 2019. A laterally transferred viral gene modifies aphid wing plasticity. Curr Biol. 12:2098-2103

Parker BJ, Spragg CJ, Altincicek B, Gerardo NM. 2013. Symbiont-mediated protection against fungal pathogens in pea aphids: a role for pathogen specificity? Appl Environ Microbiol. 79:2455-2458

Peabody NC, Diao F, Luan H, Wang H, Dewey EM, Honegger HW, White BH. 2008. Bursicon functions within the *Drosophila* CNS to modulate wing expansion behavior, hormone secretion, and cell death. J Neurosci. 28:14379–14391

Pener MP. Simpson SJ. 2009. Locust phase polyphenism: an update. Adv Insect Physiol. 36:1-286

Peng W, Zheng W, Tariq K, Yu S, Zhang H. 2019. MicroRNA Let-7 targets the ecdysone signaling pathway E75 gene to control larval-pupal development in *Bactrocera dorsalis*. Insect Sci. 26:229-239

Peng X, Liu L, Huang Y, Wang S, Li D, Chen S, Simon J, Qu M, Chen M. 2021. Involvement of chemosensory proteins in host plant searching in the bird cherry-oat aphid. Insect Sci. 28:1338-1353

Pers D, Hansen AK. 2019. The effects of different diets and transgenerational stress on *Acyrthosiphon pisum* development. Insects. 10:260

Peterson CL, Laniel M. 2004. Histones and histone modifications. Curr Biol. 14:546-551

Pfennig DW, McGee M. 2010. Resource polyphenism increases species richness: a test of the hypothesis. Philos Trans R Soc B Biol Sci. 365:577–591.

Pfennig DW, Murphy PJ. 2002. How fluctuating competition and phenotypic plasticity mediate species divergence. Evolution. 56:1217–1228

Pigliucci M, Murren CJ, Schlichting CD. 2006. Phenotypic plasticity and evolution by genetic assimilation. J Exp Biol. 209:2362-2367

Pigliucci M. 2005. Evolution of phenotypic plasticity: where are we going now? Trends Ecol Evol. 20:481-486

Porlier M, Charmantier A, Bourgault P, Perret P, Blondel J, Garant D. 2012. Variation in phenotypic plasticity and selection patterns in blue tit breeding time: between- and within-population comparisons. J Anim Ecol. 81:1041-1051

Pozo MI, Hunt BJ, Kemenade GV, Guerra-Sanz JM, Wäckers F, Mallon EB, Jacquemyn H. 2021. The effect of DNA methylation on bumblebee colony development. BMC Genom. 22:73

Pray-Grant MG, Daniel JA, Schieltz D, Yates III JR, Grant PA. 2005. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature. 433:434-438

Present TMC, Conover DO. 1992. Physiological basis of latitudinal growth differences in *Menidia menidia*: variation in consumption or efficiency? Func Ecol. 6:23–31

Price DRG, Wilson ACC, Luetje CW. 2015. Proton-dependent glutamine uptake by aphid bacteriocyte amino acid transporter ApGLNT1. Biochim Biophys Acta Biomembr. 1848:2085-2091

Probst A, Dunleavy VE, Almouzni G. 2009. Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol. 10:192-206

Provataris P, Meusemann K, Niehuis O, Grath S, Misof B. 2018. Signatures of DNA methylation across insects suggest reduced DNA methylation levels in holometabola. Genome Biol Evol. 10:1185-1197

Puthiyakunnon S, Yao Y, Li Y, Gu J, Peng H, Chen X. 2013. Functional characterization of three microRNAs of the Asian tiger mosquito, *Aedes albopictus*. Parasit Vectors. 6:230

Qu Z, Kenny NJ, Lam HM, Chan TF, Chu KH, Nemdema WG, Tobe SS, Jerome LHH. 2015. How did Arthropod sesquiterpenoids and ecdysteroids arise? Comparison of hormonal pathway genes in noninsect arthropod genomes. Genome Biol Evol. 7:1951-1959

Quah S, Breuker CJ, Holland PWH. 2015. A diversity of conserved and novel ovarian micrornas in the speckled wood (*Pararge aegeria*). PLoS One. 10:e0142243

Ragsdale EJ, Müller MR, Rödelsperger C, Sommer RJ. 2013. A developmental switch coupled to the evolution of plasticity acts through a sulfatase. Cell. 155:922-933

Raikhel AS, Dhadialla TS. 1992. Accumulation of yolk proteins in insect oocytes. Annu Rev Entomol. 37:217-251

Rank NE, Bruce DA, McMillan DM, Barclay C. Dahlhoff EP. 2007. Phosphoglucose isomerase genotype affects running speed and heat shock protein expression after exposure to extreme temperatures in a montane willow beetle. J Exp Biol. 210:750-764

Ratikainen I, Haaland TR, Wright J. 2018. Differential allocation of parental investment and the trade-off between size and number of offspring. Proc Royal Soc B. 285:1884

Reed TE, Waples RS, Schindler DE, Hard JJ, Kinnison MT. 2010. Phenotypic plasticity and population viability: the importance of environmental predictability. Proc Royal Soc B. 277:3391-3400

Rehan SM, Glastad KM, Lawson SP, Hunt BG. 2016. The genome and methylome of a subsocial small carpenter bee, *Ceratina calcarata*. Genome Biol Evol. 8:1401-1410

Richards CL, Verhoeven KJF, Bossdorf O. 2012. Evolutionary significance of epigenetic variation. Plant Genome. 1:257-274.

Riggs AD. 1975. X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet. 14:9-25

Rigoutsos I. 2009. New tricks for animal microRNAs: targeting of amino acid coding regions at conserved and nonconserved sites. Cancer Res. 69:3245-3248

Roberts SB, Gavery MR. 2012. Is There a relationship between DNA methylation and phenotypic plasticity in invertebrates? Front Physiol. 2:116

Robertson KD. 2005. DNA methylation and human disease. Nat Rev Genet. 8:597-610

Roessingh P, Bouaïchi A, Simpson SJ. 1998. Effects of sensory stimuli on the behavioural phase state of the desert locust, *Schistocerca gregaria*. J Insect Physiol. 44:883-893
Roesti M, Hendry AP, Salzburger W, Berner D. 2012. Genome divergence during evolutionary diversification as revealed in replicate lake–stream stickleback population pairs. Mol Ecol. 12:2852-2862

Roff DA. 1986. The evolution of wing dimorphism in insects. Evolution. 40:1009-1020

Rohde PD, Bøcker A, Jensen CAB, Bergstrøm AL, Madsen MIJ, Christensen SL, Villadsen SB, Kristensen TN. 2021. Genotype and trait specific responses to rapamycin intake in *Drosophila melanogaster*. Insects. 12:474

Romani P, Duchi S, Gargiulo G, Cavaliere V. 2017. Evidence for a novel function of Awd in maintenance of genomic stability. Sci Rep. 7:16820

Rubio M, Belles X. 2013. Subtle roles of microRNAs let-7, miR-100 and miR-125 on wing morphogenesis in hemimetabolan metamorphosis. J Insect Physiol. 59:1089-1094

Rubio M, Montañez R, Perez L, Milan M, Belles X. 2013. Regulation of atrophin by both strands of the mir-8 precursor. Insect Biochem Mol Biol. 43:1009-1014

Russell JA, Latorre A, Sabater-Muñoz B, Moya A, Moran A. 2003. Side-stepping secondary symbionts: widespread horizontal transfer across and beyond the Aphidoidea. Mol Ecol. 12:1061-1075

Ryabov EV, Keane G, Naish N. 2009. Densovirus induces winged morphs in asexual clones of the rosy apple aphid, *Dysaphis plantaginea*. Proc Natl Acad Sci U S A. 106:8465-9470

Sakurai M, Koga R, Tsuchida T, Meng X, Fukatsu T. 2005. *Rickettsia* symbiont in the pea aphid *Acyrthosiphon pisum*: novel cellular tropism, effect on host fitness, and interaction with the essential symbiont *Buchnera*. Appl Environ Microbiol. 71:4069-4075

Samietz J, Köhler G. 2012. A fecundity cost of (walking) mobility in an insect. Ecol Evol. 2:2788-2793

Sandström J, Pettersson J. 1994. Amino acid composition of phloem sap and the relation to intraspecific variation in pea aphid (*Acyrthosiphon pisum)* performance. J Insect Physiol. 40:947-955

Santos CG, Fernandez-Nicolas A, Belles X. 2016. Smads and insect hemimetabolan metamorphosis. Dev Biol. 417:104-113

Santos-Rosa H, Caldas C. 2005. Chromatin modifier enzymes, the histone code and cancer. Eur J Cancer. 41:2381-2402

Sarda S, Zeng J, Hunt BG, Yi SV. 2012. The evolution of invertebrate gene body methylation. Mol Biol Evol. 29:1907-1916

Sasaki T, Ishikawa H. 1995. Production of essential amino acids from glutamate by mycetocyte symbionts of the pea aphid, *Acyrthosiphon pisum*. J Insect Physiol. 41:41-46

Sattar S, Addo-Quaye C, Song Y, Anstead J.A, Sunkar R, Thompson GA. 2012. Expression of small RNA in *Aphis gossypii* and its potential role in the resistance interaction with melon. PLoS One. 7:e48579

Scarborough CL, Ferrari J, Godfray HCJ. 2005. Aphid protected from pathogen by endosymbiont. Science. 310:1781

Schaefer M, Lyko F. 2010. Lack of evidence for DNA methylation of Invader4 retroelements in *Drosophila* and implications for Dnmt2-mediated epigenetic regulation. Nat Genet. 11:920-921

Schlichting CD, Wund MA. 2014. Phenotypic plasticity and epigenetic marking: an assessment of evidence for genetic accommodation. Evolution. 68:656–672

Schmitz OJ. 2013. Global climate change and the evolutionary ecology of ecosystem functioning. Ann NY Acad Sci. 1297:61–76

Schuh R, Aicher W, Gaul U, Côté S, Preiss A, Maier D, Seifert E, Nauber U, Schröder C, Kemler R, Jäckle H. 1986. A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Krüppel, a *Drosophila* segmentation gene. Cell. 47:1025-1032

Scoville AG, Barnett LL, Bodbyl-Roels S, Kelly JK, Hileman LC. 2011. Differential regulation of a MYB transcription factor is correlated with transgenerational epigenetic inheritance of trichome density in *Mimulus guttatus*. New Phytol. 191:251-263

Seger J, Brockmann HJ. 1987. What is bet-hedging? In Oxford surveys in evolutionary biology, vol. 4 (eds Harvey P. H., Partridge L.), pp. 182–211 Oxford University Press Oxford, UK

Sentis A, Bertam R, Dardenne N, Ramon-Portugal F, Louit I, Trionnaire GL, Simon J, Magro A, Pujol B, Hemptinne J, Danchin E. 2019. Different phenotypic plastic responses to

predators observed among aphid lineages specialized on different host plants. Sci Rep. 9:9017

Serrano N, O'Farrell PH. 1997. Limb morphogenesis: connections between patterning and growth. Curr Biol. 7: R186-R195

Shang F, Ding B, Xiong Y, Dou W, Wei D, Jiang H, Wei D, Wang J. 2016. Differential expression of genes in the alate and apterous morphs of the brown citrus aphid, Toxoptera citricida. Sci Rep 2016. 6:32099

Shen Q, Zhao L, Xie G, Wei P, Yang M, Wang S, Zhang F, Tang B. 2015. Cloning three *Harmonia axyridis* (Coleoptera: Coccinellidae) heat shock protein 70 family genes: regulatory function related to heat and starvation stress. J Entomol Sci. 50:168-185

Shen Z, Liu Y, Zhu F, Cai L, Liu X, Tian Z, Cheng J, Li Z, Liu X. 2020. MicroRNA-277 regulates dopa decarboxylase to control larval-pupal and pupal-adult metamorphosis of *Helicoverpa armigera*. Insect Biochem Mol Biol. 122:103391

Shi Z, Wang S, Wang S, Zhang L, Xu Y, Guo X, Zhang F, Tang B. 2017. Effects of starvation on the carbohydrate metabolism in *Harmonia axyridis* (Pallas). Biol Open. 6:1096-1103

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

Shigenobu S, Wilson ACC. 2011. Genomic revelations of a mutualism: the pea aphid and its obligate bacterial symbiont. Cell Mol Life Sci. 68:1297-1309

Shukla E, Thorat LJ, Nath BB, Gaikwad SM. 2015. Insect trehalase: physiological significance and potential applications. Glycobiology. 25:357-367

Siegler MVS, Jia XX. 1999. Engrailed negatively regulates the expression of cell adhesion molecules connectin and neuroglian in embryonic *Drosophila* nervous system. Neuron. 22:265-276

Silva AX, Bacigalupe LD, Luna-Rudloff M, Figueroa CC. 2012. Insecticide resistance mechanisms in the green peach aphid *Myzus persicae* (Hemiptera: Aphididae) II: Costs and benefits. PLoS One. 7:e36810

Silva CPM, Terra RW, Ferreira C. 2004. The role of carboxyl, guanidine and imidazole groups in catalysis by a midgut trehalase purified from an insect larva. Insect Biochem. Mol. Biol. 34:1089-1099

Silva-Oliveira G, De Paula IF, Medina JM, Alves-Bezerra M, Gondim KC. 2021. Insulin receptor deficiency reduces lipid synthesis and reproductive function in the insect *Rhodnius prolixus*. Biochim Biophys Acta Mol Cell Biol Lipids. 1866:158851

Simmonds A, Hughes S, Tse J, Cocquyt S, Bell J. 1997. The effect of dominant vestigial alleles upon vestigial-mediated wing patterning during development of *Drosophila melanogaster*. Mech Dev. 67:17-33

Simola D, Graham RJ, Brady CM, Enzmann BL, Desplan C, Ray A, Zwiebel LJ, Bonasio R, Reinberg D, Liebig J, Berger SL. 2016. Epigenetic (re)programming of caste-specific behavior in the ant *Camponotus floridanus*. Science. 351:aac6633

Simons AM. 2011. Modes of response to environmental change and the elusive empirical evidence for bet hedging. Proc Royal Soc B. 278:1601-1609

Simpson SJ, Sword GA, Lo N. 2011. Polyphenism in insects. Curr Biol. 21:R738-R749

Sims RJ III, Nishioka K, Reinberg D. 2003. Histone ly-sine methylation: A signature for chromatin function. Trends Genet. 19:629–639

Singh S, Mishra G, Omkar. 2020. Maternal body size and age govern reproduction and offspring phenotype in the zig-zag ladybird beetle (*Menochilus sexmaculatus*). Can J Zool. 99: 97-105

Skúlason S, Smith TB. 1995. Resource polymorphisms in vertebrates. Trends Ecol Evol. 10:366–370

Smith CR, Toth AL, Suarez AV, Robinson GE. 2008. Genetic and genomic analyses of the division of labour in insect societies. Nat Rev Genet. 9:735-748

Smith TB, Skúlason S. 1996. Evolutionary significance of resource polymorphisms in fishes, amphibians, and birds. Ann Rev Ecol Syst. 27:111–133.

Smykal V., Daimon T, Kayukawa T, Takaki K, Shinoda T, Jindra M. (2014b). Importance of juvenile hormone signaling arises with competence of insect larvae to metamorphose. Developmental Biol. 390:221–230

Soares MPM, Pinheiro DG, Freitas FCP, Simões ZLP, Bitondi MMG. 2021. Transcriptome dynamics during metamorphosis of imaginal discs into wings and thoracic dorsum in *Apis mellifera* castes. BMC Genom. 22:756

Solano PJ, Mugat B, Martin D, Girard F, Huibant J, Ferraz C, Jacq B, Demaille J, Maschat F. 2003. Genome-wide identification of in vivo *Drosophila* engrailed-binding DNA fragments and related target genes. Development. 130:1243-1254

Song J. Guo W, Jiang F, Kang L, Zhou S. 2013. Argonaute 1 is indispensable for juvenile hormone mediated oogenesis in the migratory locust, *Locusta migratoria*. Insect Biochem Mol Biol. 43:879-887

Song L, Gao Y, Li J, Ban L. 2018. iTRAQ-based comparative proteomic analysis reveals molecular mechanisms underlying wing dimorphism of the pea aphid, *Acyrthosiphon pisum*. Front Physiol. 9:1016

Song X, Huang F, Liu J, Li C, Gao S, Wu W, Zhai M, Yu X, Xiong W, Xie J, Li B. 2017. Genome-wide DNA methylomes from discrete developmental stages reveal the predominance of non-CpG methylation in *Tribolium castaneum*. DNA Res. 24:445-457

Song Y, Fee L, Lee TH, Wharton RP. 2007. The molecular chaperone Hsp90 is required for mRNA localization in *Drosophila melanogaster* embryos. Genetics. 176:2213-2222

Southamer R, Breeuwer JA, Hurst GD. 1999. *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. Annu Rev Microbiol. 53:71-102

Søvik E, Bloch G, Ben-Shahar Y. 2015. Function and evolution of microRNAs in eusocial Hymenoptera. Front Genet. 6:193

Spacht DE, Teets NM, Denlinger DL. 2018. Two isoforms of Pepck in *Sarcophaga bullata* and their distinct expression profiles through development, diapause, and in response to stresses of cold and starvation. J Insect Physiol. 111:41-46

Spaulding AW, von Dohlen CD. 2001. Psyllid endosymbionts exhibit patterns of co-speciation with hosts and destabilizing substitutions in ribosomal RNA. Insect Mol Biol. 10:57-67

Srinivasan DG, Brisson JA. 2012. Aphids: a model for polyphenism and epigenetics. Genet Res Int. 2012:431531

Standage DS, Berens AJ, Glastad KM, Severin AJ, Brendel VP, Toth AL. 2016. Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. Mol Ecol. 25:1769-1784

Stark R, Brown G. 2011. DiffBind: differential binding analysis of ChIP-Seq peak data. available online at: http://bioconductor.org/packages/release/bioc/html/DiffBind.html

Stearns S.C. 1992. The evolution of life histories. Oxford University Press, Oxford.

Stepanik VA, Harte PJ. 2012. A mutation in the E(Z) methyltransferase that increases trimethylation of histone H3 lysine 27 and causes inappropriate silencing of active Polycomb target genes. Dev Biol. 364:249-258

Stomp M, van Dijk MA, van Overzee HMJ, Wortel MT, Sigon CAM, Egas M, Hoogveld H, Gons HJ, Huisman J. 2008. The timescale of phenotypic plasticity and its impact on competition in fluctuating environments. Am Nat. 172:169-185

Strassburger K, Lutz M, Müller S, Teleman AA. 2021. Ecdysone regulates *Drosophila* wing disc size via a TORC1 dependent mechanism. Nat. Commun. 12:6684

Sultan SE, Horgan-Kobelski T, Nichols LM, Riggs CE, Waples RK. 2013. A resurrection study reveals rapid adaptive evolution within populations of an invasive plant. Evol Appl. 6:266-278

Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 9:465-476

Svanbäck R, Pineda-Krch M, Doebeli M. 2009. Fluctuating population dynamics promotes the evolution of phenotypic plasticity. Am Nat. 174:176–189

Tabadkani SM, Ahsaei SM, Hosseininaveh V, Nozari J. 2013. Food stress prompts dispersal behavior in apterous pea aphids: do activated aphids incur energy loss? Physiol Behav. (110-111):221-225

Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H, Shinkai Y. 2002. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev. 16:1779-1991

Takemura M, Maoka T, Koyanagi T, Kawase N, Nishida R, Tsuchida T, Hironaka M, Ueda T, Misawa N. 2021. Elucidation of the whole carotenoid biosynthetic pathway of aphids at the gene level and arthropodal food chain involving aphids and the red dragonfly. BMC Zool. 6:19

Tamaru H. 2010. Confining euchromatin/heterochromatin territory: jumonji crosses the line. Genes Dev. 14:1465-1478

Tang Y, He H, Qu X, Cai Y, Ding W, Qiu L, et al. (2020). RNA Interference-mediated knockdown of the transcription factor Krüppel homologue 1 suppresses vitellogenesis in *Chilo Suppressalis*. Insect Mol. Biol. 29:183–192

Tanimoto H, Itoh S, Dijke P, Tabata T. 2000. Hedgehog creates a gradient of DPP activity in *Drosophila* wing imaginal discs. Mol Cell. 5:59-71

Tariel J, Plénet S, Luquet E. 2020. How do developmental and parental exposures to predation affect personality and immediate behavioural plasticity in the snail *Physa acuta*? Proc Royal Soc B. 287:1941

Taylor SC, Nadeau K, Abbasi M, Lachance C, Nguyen M, Fenrich J. 2019. The ultimate qpcr experiment: producing publication quality, reproducible data the first time. Trends Biotechnol. 37:761-774

Teleman AA, Cohen SM. 2000. Dpp Gradient Formation in the *Drosophila* wing imaginal disc. Cell 103:971–980.

Teleman AA, Cohen SM. 2006. *Drosophila* lacking microRNA miR-278 are defective in energy homeostasis. Genes Dev 20:417-422

Thomas GH, Zucker J, Macdonald SJ, Sorokin A, Goryanin I, Douglas AE. 2009. A fragile metabolic network adapted for cooperation in the symbiotic bacterium *Buchnera aphidicola*. BMC Syst Biol. 3:24

Thompson ME, Muller MN, Sabbi K, Machanda ZP, Otali E, Wrangham RW. 2016. Faster reproductive rates trade off against offspring growth in wild chimpanzees. Proc Natl Acad Sci U S A. 113:7780-7785

Tian S, Monteiro A. 2022. A transcriptomic atlas underlying developmental plasticity of seasonal forms of *Bicyclus anynana* butterflies. Mol Biol Evol. 39:msac126

Tomoyasu Y, Arakane Y, Kramer K.J, Denell RE. 2009. Repeated co-options of exoskeleton formation during wing-to-elytron evolution in beetles. Curr Biol. 19:2057-2065

Tsuchida T, Koga R, Fukatsu T. 2004. Host plant specialization governed by facultative symbiont. Science. 303:1989

Tsuji H and Kawada K. 1987. Development and degeneration of wing buds and indirect flight muscles in the pea aphid *Acyrtosiphon pisum* (Harris). J Appl Entomol Zool. 31:247-252

Turner CM, Adler PN. 1995. Morphogenesis of *Drosophila* pupal wings in vitro. Mech Dev. 52:247-255

van Buskirk J, Steiner UK. 2009. The fitness costs of developmental canalization and plasticity. J Evol Biol. 22:852-860

Van Munster MV, Dullemans AM, Verbeek M, van den Heuvel JFJM, Reinbold C, Brault V, Clérivet A, van der Wilk F. 2003. Characterization of a new densovirus infecting the green peach aphid *Myzus persicae*. J Invertebr Pathol. 84:6-14

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 7:003401- 003411

Vellichirammal NN, Gupta P, Hall TA, Brisson JA. 2017. Ecdysone signaling underlies the pea aphid transgenerational wing polyphenism. Proc Natl Acad Sci U S A. 114:1419-1423

Vigoreaux JO, Saide JD, Valgeirsdottir K, Pardue ML. 1993. Flightin, a novel myofibrillar protein of *Drosophila* stretch-activated muscles. J Cell Biol. 121:587-598

Vilmos P, Bujna A, Szuperák M, Havelda Z, Várallyay É, Szabad J, Kucerova L, Somogyi K, Kristó I, Lukácsovich T, Jankovics F, Henn L, Erdélyi M. Viability, longevity, and egg production of *Drosophila melanogaster* are regulated by the miR-282 microRNA. Genetics. 192:469-480

Vissers JH, Nicassio F, van lohuizen M, Di Fiore PP, Citterio E. 2008. The many faces of ubiquitinated histone H2A: insights from the DUBs. Cell Div. 3:8

Waddington, CH. 1968. Towards a theoretical biology. Nature. 218:525-527

Wagner DL, Liebherr JK. 1992. Flightlessnessin insects. Trends Ecol Evol. 7:216-220

Walker DW, Muffat J, Rundel C, Benzer S. 2006. Overexpression of a *Drosophila* homolog of Apolipoprotein D leads to increased stress resistance and extended lifespan. Curr Biol. 16:674-679

Walsh TK, Brisson JA, Robertson HM, Gordon K, Jaubert-Possamai S, Tagu D, Edwards OR. 2010. A functional DNA methylation system in the pea aphid, *Acyrthosiphon pisum*. Insect Mol Biol. 19:215-228

Wang G, Zhou J, Gou Y, Quandahor P, Liu C. 2021. Trehalose and glucose levels regulate feeding behavior of the phloem-feeding insect, the pea aphid Acyrthosiphon pisum Harris. Sci Rep. 11:15864

Wang H, Liu Z, Wang Y, Ma L, Zhang W, Xu B. 2020. Genome-wide differential DNA methylation in reproductive, morphological, and visual system differences between queen bee and worker bee (*Apis mellifera*). Front Genet. 11:770

Wang M, Wang Y, Chang M, Wang X, Shi Z, Raikhel AS, Zou Z. 2022. Ecdysone signaling mediates the trade-off between immunity and reproduction via suppression of amyloids in the mosquito Aedes aegypti. PLoS Pathog. 18:e1010837

Wang X, Chen Z, Feng Z, Zhu J, Zhang Y, Liu T. 2019. Starvation stress causes body color change and pigment degradation in *Acyrthosiphon pisum*. Front Physiol. 10:197

Wang X, Wheeler D, Avery A, Rago A, Choi J, Colbourne JK, Clark AG, Werren JH. 2013. Function and evolution of DNA methylation in *Nasonia vitripennis*. PLoS Genet. 9:e1003872

Wang Y, Jorda M, Jones P.L, Maleszka R, Ling X, Roberston H.M, Mizzen CA, Peinado MA, Robinson GE. 2006. Functional CpG methylation system in a social insect. Science. 314:645-647

Weber B, Zicola J, Oka R, Stam M. 2016. Plant enhancers: A call for discovery. Trends Plant Sci. 21:974-987

Wedd L, Kucharski R, Maleszka R. 2022. DNA methylation in honey bees and the unresolved questions in insect methylomics. Adv Exp Med Biol. 1389:159-176

Weibe KF, Elebute OO, LeMoine CMR, Cassone BJ. 2020. A Day in the Life: Identification of developmentally regulated micrornas in the colorado potato beetle (*Leptinotarsa decemlineata*; Coleoptera: Chrysomelidae). J Econ Entomol. 113:1445-1454

Weisser WW, Braendle C, Minoretti N. 1999. Predator-induced morphological shift in the pea aphid. Proc Royal Soc B. 266:1175

Weisser WW, Stadler B. 1994. Phenotypic plasticity and fitness in aphids. Eur J Entomol. 91:71-78

Wernegreen JJ, Moran NA. 1999. Evidence for genetic drift in endosymbionts (Buchnera): analyses of protein-coding genes. Mol Biol Evol. 16:83-97

Wesneat DF, Sherman PW, Morton ML. 1990. The ecology and evolution of extra-pair copulations in birds. In: Current ornithology, vol. 7 (Power DM, ed). New York: Plenum Press; 331-369.

West-Eberhard MJ. 2003. Developmental Plasticity and Evolution. Oxford University Press: New York, NY, USA.

Wheeler DE. 1986. Developmental and physiological determinants of caste in social hymenoptera: Evolutionary Implications. Am Nat. 128:13-34

Whitehead LF, Douglas AE. 1993. Populations of symbiotic bacteria in the parthenogenetic pea aphid (*Acyrthosiphon pisum*) Symbiosis. Proc Royal Soc B. 254:29-32

Wightman B, Ha I, Ruvkun G. 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. Cell. 75:855-862

Wilcox JL, Dunbar HE, Wolfinger RD, Moran NA. 2003. Consequences of reductive evolution for gene expression in an obligate endosymbiont. Mol Microbiol. 48:1491-1500

Will T, Schmidtberg H, Skaljac M, Vilcinskas A. 2017. Heat shock protein 83 plays pleiotropic roles in embryogenesis, longevity, and fecundity of the pea aphid *Acyrthosiphon pisum*. Dev Genes Evol. 227:1-9

Williams JA, Paddock SW, Vorwerk K, Carroll SB. 1994. Organization of wing formation and induction of a wing-patterning gene at the dorsal/ventral compartment boundary. Nature. 368:299-305

Willis CG, Ruhfel B, Primack RB, Miller-Rushing AJ, Davis CC. 2008. Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. Proc Natl Acad Sci U S A. 105:17029-17033

Wilson ACC, Ashton PD, Calevro F, Charles H, Colella S, Febvay G, Jander G, Kushlan PF, Macdonald SJ, Schwartz JF, Thomas GH, Douglas AE. 2010. Genomic insight into the amino acid relations of the pea aphid, *Acyrthosiphon pisum*, with its symbiotic bacterium *Buchnera aphidicola*. Insect Mol Biol. 19:249-258

Wood J, Schwer B, Verdin E, Helfand S. 2018. SIRT4 is a mitochondrial regulator of metabolism and lifespan in *Drosophila melanogaster*. Innov Aging. 2:91

Wu Ct, Morris JR. 2001. Genes, genetics, and epigenetics: A correspondence. Science. 293:1103-1105

Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation. 2:100141

Wyles JS, Kunkel JG, Wilson AC. 1983. Birds, behavior, and anatomical evolution. Proc Natl Acad Sci U S A. 80:4394–4397

Xiang H, Zhu J, Chen Q, Dai F, Li X, Li M, Zhang H, Zhang G, Li D, Dong Y, Zhao L, Lin Y, Cheng D, Yu J, Sun J, Zhou X, Ma K, He Y, Zhao Y, Guo S, Ye M, Guo G, Li Y, Li R, Zhang X, Ma L, Kristiansen K, Guo Q, Jiang J, Beck S, Xia Q, Wang W, Wang J. 2010. Single base–resolution methylome of the silkworm reveals a sparse epigenomic map. Nat Biotechnol. 28:516-520

Xiao M, Li J, Li W, Wang Y, Wu F, Xi Y, Zhang L, Ding C, Luo H, Li Y, Peng L, Zhao L, Peng S, Xiao Y, Dong S, Cao J, Yu W. 2017. MicroRNAs activate gene transcription epigenetically as an enhancer trigger. RNA Biol. 14:1326-1334

Xu G, Lyu H, Yi Y, Peng Y, Feng Q, Song Q, Gong C, Peng X, Palli S.R, Zheng S. 2021. Intragenic DNA methylation regulates insect gene expression and reproduction through the MBD/Tip60 complex. iScience. 24:102040

Xu G, Yi Y, Lyu H, Gong C, Feng Q, Song Q, Peng X, Liu L, Zheng S. 2020. DNA methylation suppresses chitin degradation and promotes the wing development by inhibiting Bmara-mediated chitinase expression in the silkworm, *Bombyx mori*. Epigenetics Chromatin. 13:34

Xu G, Zhang J, Lyu H, Song Q, Feng Q, Xiang H, Zheng S. 2018. DNA methylation mediates BmDeaf1-regulated tissue- and stage-specific expression of BmCHSA-2b in the silkworm, Bombyx mori. Epigenetics Chromatin. 11:32

Xu K, Yang W, Tian Y, Wu Y, Wang J. 2015. Insulin signaling pathway in the oriental fruit fly: The role of insulin receptor substrate in ovarian development. Gen Comp Endocrinol. 216:125-133

Xu L, Zhang J, Zhan A, Wang Y, Ma X, Jie W, Cao Z, Omar MAA, He K, Li F. 2020. Identification and analysis of MicroRNAs associated with wing polyphenism in the brown planthopper, *Nilaparvata lugens*. Int J Mol Sci. 21:9754

Xu T, Han J, Zhang J, Chen J, Wu N, Zhang W, Zhang Y. 2016. Absence of protoheme IX farnesyltransferase CtaB causes virulence attenuation but enhances pigment production and persister survival in MRSA. Front Microbiol. 7:1625

Xu X, Liu L, Zhang Q, Wu J. 2011. Morph-specific differences in biochemical composition related to flight capability in the wing-polyphenic *Sitobion avenae*. Entomol Exp Appl. 138:128-136

Yan H, Bonasio R, Simola DF, Liebig J, Berger SL, Reinberg D. 2015. DNA methylation in social insects: how epigenetics can control behavior and longevity. Annu Rev Entomol. 60:435-452

Yang C, Pan H, Liu Y, Zhou X. 2014. Selection of Reference Genes for Expression Analysis Using Quantitative Real-Time PCR in the Pea Aphid, *Acyrthosiphon pisum* (Harris) (Hemiptera, Aphidiae). PLoS One. 9:e110454

Yang C, Pospisilik JA. 2019. Polyphenism – A window into gene-environment interactions and phenotypic plasticity. Front Genet. 10:132

Yang J, Kalhan SC, Hanson RW. 2009. What is the metabolic role of phosphoenolpyruvate carboxykinase? J Biol chem. 284:27025-27029

Yang M, Wei Y, Jiang F, Wang Y, Guo X, He J, Kang L. 2014. MicroRNA-133 inhibits behavioral aggregation by controlling dopamine synthesis in locusts. PLoS Genet. 10:e1004206

Yang X, Liu X, Xu X, Li Z, Li Y, Song D, Yu T, Zhu F, Zhang Q, Zhou X. 2014. Gene expression profiling in winged and wingless cotton aphids, *Aphis gossypii* (Hemiptera: Aphididae). Int J Biol Sci. 10:257-267

Yao B, Li Y, Wang Z, Chen L, Poidevin M, Zhang C, Lin L, Wang F, Bao H, Jiao B, Lim J, Cheng Y, Huang L, Phillips BL, Xu T, Duan R, Moberg KH, Wu H, Jin P. 2018. Active N 6-Methyladenine demethylation by DMAD regulates gene expression by coordinating with polycomb protein in neurons. Mol Cell. 71:848-857

Yasugi T, Yamada T, Nishimura T. 2017. Adaptation to dietary conditions by trehalose metabolism in *Drosophila*. Sci Rep. 7:1619

Yegnasubramanian S, Wu Z, Haffner MC, Esopi D, Aryee MJ, Badrinath R, He TL, Morgan JD, Carvalho B, Zheng Q, Marzo AMD, Irizarry RA, Nelson WG. 2011. Chromosome-wide mapping of DNA methylation patterns in normal and malignant prostate cells reveals pervasive methylation of gene-associated and conserved intergenic sequences. BMC Genet. 12:313

Yin Y, Zhu Y, Mao J, Gundersen-Rindal DE, Liu C. 2021. Identification and characterization of microRNAs in the immature stage of the beneficial predatory bug *Arma chinensis Fallou* (Hemiptera: Pentatomidae). Arch Insect Biochem Physiol. 107:e21786

Yu G, Wang L, He Q. 2015. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 31:2382-2383

Yuan X, Feng W, Imhof A, Grummt I, Zhou Y. 2007. Activation of RNA polymerase I transcription by cockayne syndrome group B protein and histone methyltransferase G9a. Mol Cell. 27:585-595

Yue Y, Yang RL, Wang WP, Zhou QH, Chen EH, Yuan GR, et al. (2018). Involvement of Met and Kr-H1 in JH-Mediated reproduction of female *Bactrocera dorsalis* (Hendel). Front Physiol. 9:482

Zakhari S. 2013. Alcohol metabolism and epigenetics changes. Alcohol Res. 1:6-16

Zecca M, Struhl G. 2021. A unified mechanism for the control of *Drosophila* wing growth by the morphogens Decapentaplegic and Wingless. PLoS Biol. 19:e3001111

Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science. 5980:916-919

Zeng J, Yi SV. 2010. DNA methylation and genome evolution in honeybee: gene length, expression, functional enrichment covary with the evolutionary signature of DNA methylation. Genome Biol Evol. 2:770-780

Zera AJ and Harshman LG. 2001. The physiology of life-history trade-offs in animals. Annu Rev Ecol Syst. 32:95-126

Zera AJ, Denno RF. 1997. Physiology and ecology of dispersal polymorphism in insects. Annu Rev Entomol. 42:207-230

Zera AJ, Mole S, Rokke K. 1994. Lipid, carbohydrate and nitrogen content of long-and short-winged *Gryllus firmus*: implications for the physiological cost of flight capability. J Insect Physiol. 40:1037-1044

Zhang C, Brisson JA, Xu H. 2019. Molecular mechanisms of wing polymorphism in insects. Annu Rev Entomol. 64:297-413

Zhang C, Sun L, Xie J, Cao C. 2022. RNAi-based functional analysis of bursicon genes related to wing expansion in gypsy moths. J Insect Physiol. 139:104398

Zhang F, Li X, Zhang Y, Coates B, Zhou X, Cheng D. 2015. Bacterial symbionts, *Buchnera*, and starvation on wing dimorphism in English grain aphid, *Sitobion avenae* (F.) (Homoptera: Aphididae). Front Physiol. 6:155

Zhang J, Fu S, Chen S, Chen H, Liu Y, Liu X, Xu H. 2021. Vestigial mediates the effect of insulin signaling pathway on wing-morph switching in planthoppers. PLoS Genet. 17:e1009312

Zhang J, Wen D, Li EY, Palli SR, Li S, Wang J, Liu S. 2021. MicroRNA miR-8 promotes cell growth of corpus allatum and juvenile hormone biosynthesis independent of insulin/IGF signaling in *Drosophila melanogaster*. Insect Biochem Mol Biol. 136:103611

Zhang J, Xing Y, Li Y, Yin C, Ge C, Li F. 2015. DNA methyltransferases have an essential role in female fecundity in brown planthopper, *Nilaparvata lugens*. Biochem Biophys Res Commun. 464:83-88

Zhang L, Qiu L, Yang H, Wang H, Zhou M, Wang S, Tang B. 2017. Study on the effect of wing bud chitin metabolism and its developmental networ genes in brown planthopper *Nilaparvata lugens*, by knockdown of TRE gene. Front Physiol. 26:750

Zhang L, Wang M, Li X, Wang X, Jia C, Yang X, Feng R, Yuan M. 2018. A small set of differentially expressed genes was associated with two color morphs in natural populations of the pea aphid *Acyrthosiphon pisum*. Gene. 651:23-32

Zhang Q, Dou W, Song Z, Jin T, Yuan G, De Schutter K, Smagghe G, Wang J. 2020. Identification and profiling of *Bactrocera dorsalis* microRNAs and their potential roles in regulating the developmental transitions of egg hatching, molting, pupation and adult eclosion. Insect Biochem Mol Biol. 127:103475

Zhang R, Chen J, Jiang L, Qiao G. 2019. The genes expression difference between winged and wingless bird cherry-oat aphid *Rhopalosiphum padi* based on transcriptomic data. Sci Rep. 9:4754

Zhang T, Song W, Li Z, Qian W, Wei L, Yang Y, Wang W, Zhou X, Meng M, Peng J, Xia Q, Perrimon N, Cheng D. 2018. Krüppel homolog 1 represses insect ecdysone biosynthesis by directly inhibiting the transcription of steroidogenic enzymes. Proc Natl Acad Sci U S A. 115:3960–5.

Zhang X, Luo D, Pflugfelder GO, Shen J. 2013. Dpp signaling inhibits proliferation in the *Drosophila* wing by Omb-dependent regional control of bantam. Development 140:2917–2922

Zhang Y, Liu T, Meyer C.A, Eeckhoute J, Johnson D.S, Bernstein B.E, Nusbaum C, Myers R.M, Brown M, Li W, Liu X.S. 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9:R137

Zhang Y, Ma Z, Zhou H, Chao Z, Yan S, Shen J. 2021. Nanocarrier-delivered dsRNA suppresses wing development of green peach aphids. Insect Sci. 29: 669-682

Zhang Y, Wang L, Wu K, Wyckhuys K.A.G, Heimpel G.E. 2008. Flight performance of the soybean aphid, *Aphis glycines* (Hemiptera: Aphididae) under different temperature and humidity regimens. Environ Entomol. 37:301-306

Zhang Z, Feng J, Pan C, Lv X, Wu W, Zhou Z, Liu F, Zhang L, Zhao Y. 2013. Atrophin–Rpd3 complex represses Hedgehog signaling by acting as a corepressor of CiR. J Cell Biol. 203:575-583

Zhou X, Ye Y, Ogihara MH, Takeshima M, Fujinaga D, Liu C, Zhu Z, Kataoka H, Bao Y. 2020. Functional analysis of ecdysteroid biosynthetic enzymes of the rice planthopper, *Nilaparvata lugens*. Insect Biochem Mol Biol. 123:103428

Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. 2007. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet. 39:61-69

Zwaan B, Bijlsma R, Hoekstra RF. 1995. Direct selection on life span in *Drosophila melanogaster*. Evolution. 49:649-659

Zwier MV, Verhulst EC, Zwahlen RD, Beukeboom LW, van de Zande L. 2012. DNA methylation plays a crucial role during early Nasonia development. Insect Mol Biol. 21:129-138